

Clinical utility of next-generation sequencing for inherited bone marrow failure syndromes

Hideki Muramatsu, MD, PhD¹, Yusuke Okuno, MD, PhD¹, Kenichi Yoshida, MD, PhD², Yuichi Shiraishi, PhD³, Sayoko Doisaki, MD, PhD¹, Atsushi Narita, MD, PhD¹, Hirotochi Sakaguchi, MD, PhD¹, Nozomu Kawashima, MD, PhD¹, Xinan Wang, MD, PhD¹, Yinyan Xu, MD, PhD¹, Kenichi Chiba, PhD³, Hiroko Tanaka, PhD³, Asahito Hama, MD, PhD¹, Masashi Sanada, MD, PhD^{2,4}, Yoshiyuki Takahashi, MD, PhD¹, Hitoshi Kanno, MD, PhD⁵, Hiroki Yamaguchi, MD, PhD⁶, Shouichi Ohga, MD, PhD⁷, Atsushi Manabe, MD, PhD⁸, Hideo Harigae, MD, PhD⁹, Shinji Kunishima, PhD⁴, Eiichi Ishii, MD, PhD¹⁰, Masao Kobayashi, MD, PhD¹¹, Kenichi Koike, MD, PhD¹², Kenichiro Watanabe, MD, PhD¹³, Etsuro Ito, MD, PhD¹⁴, Minoru Takata, MD, PhD¹⁵, Miharuru Yabe, MD, PhD¹⁶, Seishi Ogawa, MD, PhD², Satoru Miyano, PhD^{3,17} and Seiji Kojima, MD, PhD¹

Purpose: Precise genetic diagnosis of inherited bone marrow failure syndromes (IBMFS), a heterogeneous group of genetic disorders, is challenging but essential for precise clinical decision making.

Methods: We analyzed 121 IBMFS patients using a targeted sequencing covering 184 associated genes and 250 IBMFS patients using whole-exome sequencing (WES).

Results: We achieved successful genetic diagnoses for 53 of 121 patients (44%) using targeted sequencing and for 68 of 250 patients (27%) using WES. In the majority of cases (targeted sequencing: 45/53, 85%; WES: 63/68, 93%), the detected variants were concordant

with, and therefore supported, the clinical diagnoses. However, in the remaining 13 cases (8 patients by target sequencing and 5 patients by WES), the clinical diagnoses were incompatible with the detected variants.

Conclusion: Our approach utilizing targeted sequencing and WES achieved satisfactory diagnostic rates and supported the efficacy of massive parallel sequencing as a diagnostic tool for IBMFS.

Genet Med advance online publication 19 January 2017

Key Words: Fanconi anemia; inherited bone marrow failure; next-generation sequencing; target sequencing; whole-exome sequencing

INTRODUCTION

Inherited bone marrow failure syndromes (IBMFS) are part of a heterogeneous disease category involving a family history in which at least one hematopoietic cell lineage is decreased in the bone marrow. IBMFS consist of more than 25 defined disease entities, including Fanconi anemia (FA), Diamond-Blackfan anemia (DBA), and dyskeratosis congenita (DC).¹ Certain IBMFS have been associated with an increased risk of secondary malignancies. The diagnosis is based on hematological

and physical findings with the aid of several disease-specific diagnostic tests such as the chromosomal breakage test for FA and molecular diagnosis with conventional Sanger sequencing for a very limited number of the causative genes. With recent advances in clinical molecular studies that revealed a considerable amount of pathognomonic molecular lesions in IBMFS,²⁻⁴ the role of genetic tests has become more important in the diagnosis of these diseases. However, because clinical and laboratory findings can overlap among different IBMFS, the selection

The first three authors contributed equally to this work.

¹Department of Pediatrics, Nagoya University Graduate School of Medicine, Nagoya, Japan; ²Department of Pathology and Tumor Biology, Graduate School of Medicine, Kyoto University, Kyoto, Japan; ³Laboratory of DNA Information Analysis, Human Genome Center, Institute of Medical Science, The University of Tokyo, Tokyo, Japan; ⁴Department of Advanced Diagnosis, Clinical Research Center, National Hospital Organization Nagoya Medical Center, Nagoya, Japan; ⁵Department of Transfusion Medicine and Cell Processing, Tokyo Women's Medical University, Tokyo, Japan; ⁶Department of Hematology, Nippon Medical School, Tokyo, Japan; ⁷Department of Pediatrics, Yamaguchi University Graduate School of Medicine, Ube, Japan; ⁸Department of Pediatrics, St. Luke's International Hospital, Tokyo, Japan; ⁹Department of Hematology and Rheumatology, Tohoku University Graduate School, Sendai, Japan; ¹⁰Department of Pediatrics, Ehime University Graduate School of Medicine, Ehime, Japan; ¹¹Department of Pediatrics, Hiroshima University Hospital, Hiroshima, Japan; ¹²Department of Pediatrics, Shinshu University School of Medicine, Matsumoto, Japan; ¹³Department of Hematology/Oncology, Shizuoka Children's Hospital, Shizuoka, Japan; ¹⁴Department of Pediatrics, Hirosaki University Graduate School of Medicine, Hirosaki, Japan; ¹⁵Laboratory of DNA Damage Signaling, Department of Late Effects Studies, Radiation Biology Center, Kyoto University, Kyoto, Japan; ¹⁶Department of Cell Transplantation and Regenerative Medicine, Tokai University Hospital, Isehara, Japan; ¹⁷Laboratory of Sequence Analysis, Human Genome Center, Institute of Medical Science, The University of Tokyo, Tokyo, Japan. Correspondence: Seiji Kojima (kojimas@med.nagoya-u.ac.jp)

Submitted 15 September 2016; accepted 16 October 2016; advance online publication 19 January 2017. doi:10.1038/gim.2016.197

of targeted genes for genetic diagnosis is difficult. Therefore, a comprehensive genetic diagnostic approach is warranted for patients with a suspicious diagnosis of IBMFS due to the increased number of genes to be analyzed.

Next-generation sequencing (NGS) encompasses a broad range of techniques that enable the simultaneous sequencing of a massive amount of nucleic acid molecules.⁵ These complementary approaches include targeted gene sequencing, whole-exome sequencing (WES), and whole-genome sequencing (WGS), each with distinct advantages and disadvantages. Compared with WES/WGS, targeted gene sequencing is a relatively inexpensive approach for the identification of pathogenic mutations in more than 100 genes and has relatively higher sequence coverage. NGS has been used for research studies regarding certain IBMFS, but its clinical utility is limited. Only two studies used targeted gene sequencing with NGS technology to diagnose patients with distinct IBMFS.^{6,7} The majority of these patients had previously undergone extensive genetic testing, all of which yielded negative results. Targeted gene sequencing efficiently identified causative mutations in both studies, which supported the utilization of NGS for genetic screening of patients with IBMFS.

In this study, we developed a targeted gene sequencing platform with 184 genes that were specifically designed for the diagnosis of IBMFS. This platform was used as a first-line diagnostic test for 121 patients clinically diagnosed with IBMFS. In addition, we performed diagnostic WES for 250 patients with IBMFS who could not be genetically diagnosed using conventional genetic tests.

MATERIALS AND METHODS

Patients for targeted sequencing

A total of 121 consecutive patients who were clinically diagnosed with IBMFS at Nagoya University Hospital and other nationwide institutions were included in this study. None of the patients had previously undergone genetic tests. Clinical diagnoses included DBA ($n = 26$), DC ($n = 13$), FA ($n = 22$), Shwachman-Diamond syndrome (SDS, $n = 6$), severe congenital neutropenia (SCN, $n = 7$), other anemia ($n = 21$), other neutropenia ($n = 3$), other thrombocytopenia ($n = 13$), and other bone marrow failure (BMF) (cytopenia with ≥ 2 lineages; $n = 10$) (Table 1, Supplementary Table S1 online and Supplementary Figure S1 online).

Whole-exome sequencing of patients

In 2011, we established a government-supported nationwide program, the Research on Measures for Intractable Diseases Project of the Ministry of Health, Labor, and Welfare, for rare inherited blood disorders. As of December 2013, a total of 733 patients were registered with the program (Supplementary Table S1 online) and had clinical diagnoses including FA ($n = 117$), DBA ($n = 110$), congenital hemolytic anemia (HA; $n = 261$), DC ($n = 62$), congenital dyserythropoietic anemia (CDA; $n = 21$), congenital sideroblastic anemia (CSA; $n = 34$), congenital amegakaryocytic thrombocytopenia (CAMT; $n = 10$),

hereditary hemophagocytic lymphohistiocytosis (HLH; $n = 65$), SCN ($n = 47$), and unclassified IBMFS ($n = 6$). Conventional genotyping such as Sanger sequencing had confirmed the clinical diagnoses of 262 patients; however, no causative candidate germline variants were identified ($n = 267$) or no genotyping was performed ($n = 204$) for the remaining 471 patients. Of these 471 patients, 250 patients with FA ($n = 73$), DBA ($n = 61$), HA ($n = 44$), DC ($n = 29$), CDA ($n = 12$), CSA ($n = 9$), CAMT ($n = 7$), HLH ($n = 6$), SCN ($n = 3$), and unclassified IBMFS ($n = 6$) were enrolled in the present study (Table 2, Supplementary Tables S2 and S3 online). Most of these patients (182/250, 73%) underwent various genetic tests with negative results before WES analysis, whereas the remaining 68 (27%) patients were not evaluated with genetic tests other than WES.

Sample preparation and next-generation sequencing

Written informed consent was obtained from patients or their legal guardians. This study was approved by the ethics committees of Nagoya University Graduate School of Medicine and Graduate School of Medicine, Kyoto University.

Genomic DNA was extracted from peripheral blood, bone marrow, and Epstein-Barr virus-transformed lymphoblastoid cell lines (EBV-LCL) using a QIAamp DNA Blood Mini Kit (Qiagen, Hilden, Germany). For targeted sequencing, the target region of the genomic DNA was enriched using SureSelect custom bait (Agilent Technologies, Santa Clara, CA) covering the exons and 10 bases surrounding the exons of 184 genes (Supplementary Table S4 online). For WES, genomic DNA was captured using SureSelect Human All Exon 50M, V4, or V5 Kits (Agilent Technologies). Captured genomic DNA was analyzed by massively parallel sequencing using a HiSeq 2000 or 2500 (Illumina, San Diego, CA) next-generation sequencer with a 100×2 paired-end option.

Detection of causative variants

We detected germline variants using our established pipeline (Genomon-exome, <http://genomon.hgc.jp/exome/>). In brief, sequence reads were aligned to the hg19 reference genome using the Burrows-Wheeler Aligner; variants were detected using our in-house variant caller. Variant allele frequency (VAF) > 0.2 (20%) was used as the cut-off value.

Following the guidelines published by the American College of Medical Genetics and Genomics,⁸ we removed common single-nucleotide polymorphisms (SNPs) showing minor allele frequency values of more than 1% in (i) the ESP6500 exome variant server (the National Heart, Lung, and Blood Institute Exome Sequencing Project, Seattle, WA; <http://evs.gs.washington.edu/EVS/>, as of April 2014); (ii) the 1000 genomes project;⁹ or (iii) our in-house SNP database.¹⁰ These variants were considered the causative variants that were previously reported to be pathogenic (category 1) or were otherwise highly expected to cause the associated disorders (e.g., nonsense, frameshift, and splice site variants) (category 2). Other variants of unknown significance such as missense variants without further evidence of pathogenicity were treated as

Table 1 Summary of clinical and genetic diagnosis by target sequencing

Categories of clinical diagnosis	N	Patients with genetic diagnosis	Clinical and genetic diagnosis		Identified gene mutations (n)
			Matched	Unmatched	
FA	22	15 (68%)	15	0	FANCA (11), FANCF (1), FANCG (3)
DBA	26	12 (46%)	11	1	RPL5 (3), RPS17 (3), RPS19 (4), RPS24 (1), SPTB (1)
CHA	8	4 (50%)	4	0	PIEZO1 (2), SPTB (2)
DC	13	5 (38%)	4	1	DKC1 (2), SBDS (1), TINF2 (2)
CDA	9	2 (22%)	2	0	KLF1 (1), CDAN1 (1)
CSA	4	2 (50%)	2	0	ALAS2 (1), SLC25A38 (1)
CAMT	2	1 (50%)	0	1	RUNX1 (1)
HLH	0	0	0	0	-
SCN	7	2 (29%)	2	0	ELANE (1), HAX1 (1)
SDS	6	4 (67%)	4	0	SBDS (4)
Other neutropenia	3	0	0	0	-
Other thrombocytopenia	11	3 (27%)	1	2	RUNX1 (2), VWF (1)
Other BMF	10	3 (30%)	0	3	FANCG (1), RPL5 (1), SBDS (1)
Total cohort	121	53 (44%)	45	8	

BMF, bone marrow failure; CAMT, congenital amegaryocytic thrombocytopenia; CDA, congenital dyserythropoietic anemia; CHA, congenital hemolytic anemia; CSA, congenital sideroblastic anemia; DBA, Diamond-Blackfan anemia; DC, dyskeratosis congenita; FA, Fanconi anemia; HLH, hereditary hemophagocytic lymphohistiocytosis; SCN, severe congenital neutropenia; SDS, Shwachman-Diamond syndrome.

Table 2 Summary of genetic diagnoses by whole-exome sequencing

Categories of clinical diagnosis	Total cohort		Without previous genetic tests		With previous genetic tests		Clinical and genetic diagnoses		
	N	Patients with genetic diagnosis	n	Patients with genetic diagnosis	n	Patients with genetic diagnosis	Matched	Unmatched	Identified gene mutations (n)
FA	73	35 (48%)	16	10 (63%)	57	25 (44%)	35	0	FANCG (17), FANCA (14), FANCB (1), FANCF (1), SLX4 (1), BRCA2 (1)
DBA	61	11 (18%)	0	0	61	11 (18%)	11	0	RPS26 (3), RPS7 (2), RPS19 (2), RPL5 (2), RPL35A (1), RPL11 (1)
HA	44	7 (16%)	44	7 (16%)	0	0	6	1	SPTA1 (2), SPTB (2), ANK1 (2), CDAN1 (1)
DC	29	7 (24%)	2	1 (50%)	27	6 (22%)	7	0	TERT (3), TINF2 (2), DKC1 (2)
CDA	12	3 (25%)	0	0	12	3 (25%)	1	2	CDAN1 (1), SPTA1 (1), ANK1 (1)
CSA	9	0	0	0	9	0	0	0	
CAMT	7	1 (14%)	0	0	7	1 (14%)	0	1	TINF2 (1)
HLH	6	3 (50%)	0	0	6	3 (50%)	2	1	UNC13D (1), XIAP (1), MVK (1)
SCN	3	0	0	0	3	0	0	0	
SDS	0	0	0	0	0	0	0	0	-
Other BMF	6	1 (17%)	6	1 (17%)	0	0	1	0	RUNX1 (1)
Total	250	68 (27%)	68	19 (28%)	182	49 (27%)	63	5	

BMF, bone marrow failure; CAMT, congenital amegaryocytic thrombocytopenia; CDA, congenital dyserythropoietic anemia; CSA, congenital sideroblastic anemia; DBA, Diamond-Blackfan anemia; DC, dyskeratosis congenita; FA, Fanconi anemia; HA, hemolytic anemia; HLH, hereditary hemophagocytic lymphohistiocytosis; SCN, severe congenital neutropenia; SDS, Shwachman-Diamond syndrome.

nondiagnostic in this study. For the specific pathogenicity of each variant, we used the Human Genome Mutation Database (<http://www.hgmd.cf.ac.uk/>, as of March 2014) and performed an extensive search of the literature in PubMed (<http://www.ncbi.nlm.nih.gov/pubmed>). The results of the literature search were discussed with physicians who were experts in that particular disease area and genetic diagnoses were made based on the mode of inheritance of each disease. Finally, causative variants were validated by Sanger sequencing.

Using targeted sequencing data, we performed copy-number analysis as described previously.¹¹ In brief, the coverage of each exon normalized by the mean coverage of the entire sample was compared with that of 12 unrelated reference samples. Exons exhibiting normalized coverage greater than 3 standard deviations from the coverage of reference samples were determined to be candidates for copy-number alterations. All candidate exons were visually inspected using the Integrative Genomics Viewer.

RESULTS

Genetic diagnosis by targeted sequencing

Capture-based targeted sequencing covered 99.4% of the target region in 184 genes with more than 20 independent reads. With this coverage, our in-house pipeline detected 227 (201–267) coding variants per patient, of which common SNPs with >1% minor allele frequency accounted for 97% of all detected variants. In total, we identified 69 variants that were considered to be in category 1 (i.e., previously reported alleles) or category 2 (i.e., previously unknown but highly probable variants within known causative genes for each disease subtype). In addition, we were able to identify pathognomonic copy-number aberrations in 11 patients (*FANCA* ($n = 5$; UPN-1028, -1082, -1084, -1372, and -1373), *RPS17* ($n = 3$; UPN-1174,

-1186, and -1304), *RUNX1* ($n = 1$, UPN-1222), *SBDS* ($n = 1$, UPN-1212), and *SPTB* ($n = 1$, UPN-1350) (Figure 1).

We calculated genetic diagnostic rates by our targeted sequencing pipeline based on the estimated mode of inheritance. For instance, we genetically diagnosed patients harboring homozygous or compound heterozygous mutations of known causative genes for autosomal recessive diseases such as FA. Patients with a clinical diagnosis of FA (15/22, 68%) and SDS (4/6, 67%) achieved relatively high genetic diagnostic rates. The genetic diagnosis was achieved in approximately half of the patients with DBA (12/26, 46%) and DC (5/13, 38%). Genetic diagnostic rates of other categories were as follows: SCN (2/7, 29%), other anemia (9/21, 43%), other neutropenia (0/3, 0%), other thrombocytopenia (4/13, 33%), and other BMF (3/10,

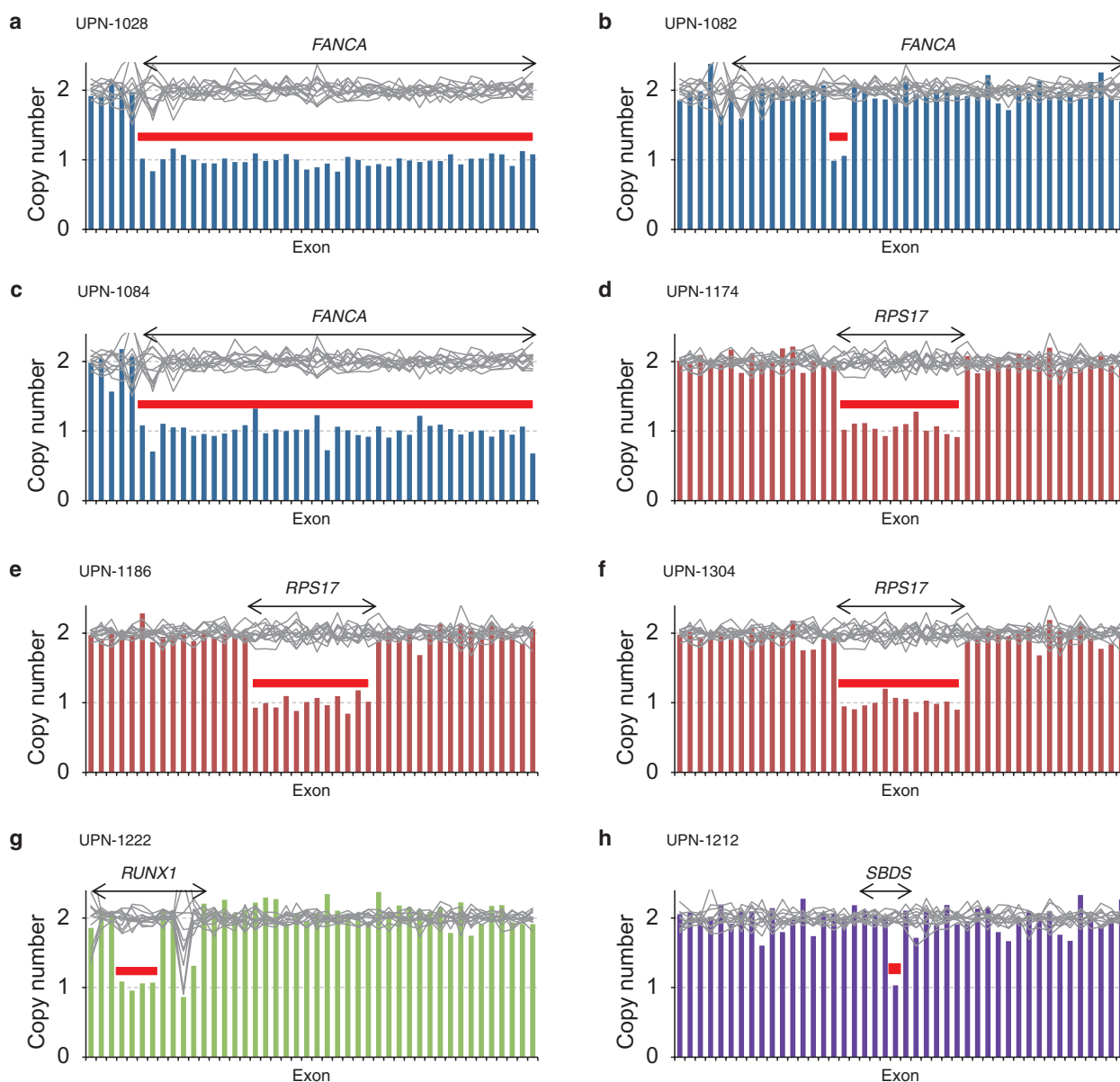


Figure 1 Copy-number analysis. The estimated copy number of each exon was based on the number of reads obtained from the targeted sequence. Analysis revealed pathogenic heterozygous gene deletion in eight patients. Each bar represents an exon and the y-axis represents the estimated copy numbers. Gray lines indicate estimated copy numbers obtained from 12 unrelated samples. Red bars indicate estimated deleted regions.

30%). In total, our targeted sequencing pipeline genetically diagnosed 53 of 121 patients (44%) (Table 1, Supplementary Table S1 online). No patient had two or more genetic diagnoses.

Genetic diagnosis by whole-exome sequencing

WES covered more than 80% (56%–91% with SureSelect 50M) to 90% (79–98% with SureSelect V4 or V5) of the coding region with more than 20 independent reads (Supplementary Figure S2 online). At this coverage level, our in-house pipeline detected 19,574 (13,811–21,945) coding variants per patient. Common SNPs with >1% minor allele frequency accounted for 97% of all detected variants.

We identified 64 category 1 variants (previously reported alleles) and 23 category 2 variants (previously unknown but highly probable variants within known causative genes for each disease subtype) (Supplementary Table S3 online). These 87 variants established genetic diagnoses for 68 (27%) patients (Table 2). Diagnostic efficacy was comparable between patients with or without prior genetic testing (26 and 27%, respectively). No patient received more than one genetic diagnosis.

The highest diagnostic efficacy was achieved in patients with FA; in these patients, one or more highly putative causative variants were detected by WES in 35 of 73 patients (48%), followed by DBA (11/61, 18%), HA (7/44, 16%), DC (7/29, 24%), CDA (3/12, 25%), and HLH (3/6, 50%). In contrast, candidate variants were detected or otherwise not identified in only a small fraction of patients with CSA (0/9), CAMT (1/7), and SCN (0/3).

Discordance between clinical diagnosis and genetic variants by targeted sequencing and whole-exome sequencing

In the majority of cases (target sequencing: 45/53, 85%; WES: 63/68, 93%), the detected variants were concordant with, and therefore supported, the clinical diagnoses. However, for the remaining 13 cases (8 patients by target sequencing and 5 patients by WES), the clinical diagnoses were incompatible with the detected variants (Table 3, Supplementary Data online).

UPN-1350, one of the discordant patients clinically diagnosed with DBA, harbored a known pathognomonic *SPTB* gene deletion and was genetically diagnosed as hereditary spherocytosis. Two patients with thrombocytopenia (congenital amegakaryocytic thrombocytopenia (UPN-1222) and idiopathic thrombocytopenic purpura (ITP); UPN-1277) were reallocated as familial platelet disorder with propensity to myeloid malignancy (FPD/AML) with *RUNX1* gene alterations¹² (microdeletion (UPN-1222, Figure 1) and p.P330fs (UPN-1277)). Three patients had IBMFS but not otherwise specified obtained genetic diagnoses as a result of our analysis (UPN-1324, DBA; UPN-1348, SDS; and UPN-1355, FA). Similarly, genetic sequencing analysis corrected the diagnoses of two other patients (UPN-1289 (from DC to SDS) and UPN-1098 (from chronic ITP to von Willebrand disease (VWD) 2B)). In UPN-355, the initial clinical diagnosis was HA, but the detected variants (*CDAN1* p.N599S and p.P146fs) supported the diagnosis of CDA. Conversely, in UPN-216 and UPN-485, the suspected clinical diagnosis of CDA was revised to

Table 3 Genetic diagnoses unmatched with clinical diagnoses

UPN	Sequencing	Clinical Dx	Genetic Dx	Gene mutation (1)	Zygosity	Gene mutation (2)	Zygosity
UPN-1350	Target	DBA	Hereditary spherocytosis	<i>SPTB</i>	Deletion		Hetero
UPN-1289	Target	DC	SDS	<i>SBDS</i>	p.K33E		Hetero
UPN-1098	Target	Chronic ITP	VWD 2B	<i>VWF</i>	p.A1461D		Hetero
UPN-1222	Target	CAMT	FPD/AML	<i>RUNX1</i>	Deletion		Hetero
UPN-1277	Target	ITP	FPD/AML	<i>RUNX1</i>	p.P330fs		Hetero
UPN-1324	Target	IBMFS, NOS	DBA	<i>RPL5</i>	p.Y219X		Hetero
UPN-1348	Target	IBMFS, NOS	SDS	<i>SBDS</i>	p.K62X		Hetero
UPN-1355	Target	IBMFS, NOS	FA	<i>FANCG</i>	p.Q356X		Hetero
UPN-355	WES	HA	CDA	<i>CDAN1</i>	p.N599S		Hetero
UPN-216	WES	CDA	HA	<i>SPTA1</i>	p.R28H		Hetero
UPN-485	WES	CDA	HA	<i>ANK1</i>	p.R935X		Hetero
UPN-83	WES	CAMT	DC	<i>TINF2</i>	p.R276X		Hetero
UPN-312	WES	HLH	Hyper IgD syndrome	<i>MVK</i>	Splice site (c.227-1G>A)		Hetero

CAMT, congenital amegakaryocytic thrombocytopenia; DBA, Diamond-Blackfan anemia; DC, dyskeratosis congenita; Dx, diagnosis; FA, Fanconi anemia; FPD/AML, familial platelet disorder with propensity to myeloid malignancy; Hetero, heterozygous; ITP, idiopathic thrombocytopenic purpura; IBMFS, inherited bone marrow failure syndrome; NOS, not otherwise specified; SDS, Shwachman-Diamond syndrome; UPN, unique patient number; VWD, von Willebrand disease; WES, whole-exome sequencing.

HA as a result of identification of mutations (*SPTA1* p.R28H in UPN-216 and *ANK1* p.R935X in UPN-485) that are typically associated with HA. In these cases, the inconsistency was most likely due to an overlap of clinical phenotypes between HA and CDA; both of which are characterized by hemolysis. In UPN-83, a patient clinically diagnosed with possible CAMT, the genetic diagnosis of DC was reached with causative variants in *TINF2* (p.R276X). In UPN-312, the clinical diagnosis of HLH was revised to hyperimmunoglobulin D syndrome, a type of hyperinflammatory syndrome, based on the causative variants in *MVK* (p.A147T and c.227-1G>A).

DISCUSSION

Recent advances in genetic research identified a large number of causative genes of IBMFS and reinforced the need for a comprehensive genetic diagnostic system in both clinical practice and research.⁶ Here, we developed a molecular diagnostic system using WES and a targeted sequencing pipeline covering 184 associated genes for IBMFS. We were successful in providing genetic diagnoses for 53 of 121 patients (44%) by targeted sequencing (Table 1) and for 68 of 250 patients (27%) by WES (Table 2). Although the possibility of concomitant diagnoses remains, 13 patients with discordant clinical and genetic diagnoses clearly demonstrated the clinical value of next-generation sequencing (Table 3).

The diagnostic rate of our targeted sequencing platform was significantly higher than those demonstrated in a previous study by Zhang *et al.*⁶ (53/121 (44%) vs. 17/85 (20%); $P < 0.001$), thus reflecting the difference between patient characteristics and genetic regions covered. Although we used a non-biased approach analyzing 121 consecutive patients who were clinically suspected to have IBMFS without preceding genetic screening, Zhang *et al.* resequenced patients who remained unclassified after a conventional genetic workup using a sequencing platform with a smaller gene number (85 genes) and a lower coverage rate ($>10\times$ coverage in 98.2% of bases).⁶ Nine patients in our study had diagnostic variants in four genes that were not included in the gene panel of Zhang *et al.* (*PIEZO1* ($n = 2$), *RPS17* ($n = 3$), *SPTB* ($n = 3$), and *VWF* ($n = 1$)). Ghemla *et al.* reported similar diagnostic rates (59/158, 37%) using a multiplex polymerase chain reaction (PCR) platform covering 72 genes⁷; however, the detection rate of copy-number variants was significantly lower than that of our study (2/158 (2%) vs. 11/121 (9%); $P = 0.003$), thus reflecting the superiority of the capture-based platform over multiplex PCR in copy-number analysis. In addition, 10 patients in our study had diagnostic variants in 7 genes that were not included in the gene panel utilized by Ghemla *et al.* (*ALAS2* ($n = 1$), *CDANI* ($n = 1$), *KLF1* ($n = 1$), *PIEZO1* ($n = 2$), *SLC25A38* ($n = 1$), *SPTB* ($n = 3$), and *VWF* ($n = 1$)).

Unsuspected genetic diagnosis has a positive effect on the clinical management of patients, particularly for gene mutations conferring cancer predisposition. Patients with most categories of IBMFS, including DBA,¹³ DC,¹⁴ FA,¹⁵ SDS,¹⁶ and SCN,¹⁷ have a significantly higher probability of hematological

malignancies than the general population. In addition, we identified three *RUNX1* gene alterations (mutation/deletion) among four patients with thrombocytopenia in the targeted sequencing cohort who were diagnosed with FPD/AML. All three patients developed thrombocytopenia during infancy, suggesting that FPD/AML should be included in the differential diagnosis of infantile thrombocytopenia. In total, we genetically diagnosed 43 patients with cancer predisposition in the targeted sequencing cohort (DBA, $n = 12$; DC, $n = 4$; FA, $n = 16$; SDS, $n = 6$; SCN, $n = 2$; FPD/AML, $n = 3$) and 55 patients in the WES cohort (DBA, $n = 11$; DC, $n = 8$; FA, $n = 35$; FPD/AML, $n = 1$). These patients need to be continually and regularly evaluated by complete blood counts and physical examination to screen for the development of hematological malignancies. In addition, appropriate genetic counseling and familial genetic screening are mandatory.

The diagnostic rate of WES was inferior to that of target sequencing (53/121 (44%) vs. 68/250 (27%); $P = 0.002$), mainly due to the conventional genetic testing conducted before enrollment in the WES cohort. For example, the highest diagnostic rate was demonstrated for FA (15/22, 68%) in the target sequencing cohort, reflecting the ability of the chromosomal breakage test to achieve a precise clinical diagnosis. In contrast, the diagnostic efficacy of WES in patients with FA and prior genetic testing tended to be lower (43%, 25/57), whereas WES in patients with FA without prior genetic testing showed diagnostic rates (62%, 10/16) comparable to that of target sequencing. In addition, our target sequencing platform was able to identify copy-number variants in 11 of 121 patients (9%); however, we could not perform a reliable copy-number analysis in the WES cohort due to relatively low coverage.

Compared to WES, targeted sequencing can achieve similar results at a lower cost for sequencing and computing resources. Although we did not identify any incidental, medically actionable genetic discovery that was not associated with the targeted hematological disease categories in our WES cohort, targeted sequencing could decrease the risk of incidental genetic discovery, which may cause serious ethical problems in a clinical setting.¹⁸

Future accumulation of data regarding genotype–phenotype correlation and functional studies of variants with unknown significance will promote the accuracy of genetic testing for IBMFS. The diagnostic yield of our WES analysis (27%) suggested the insufficiency of current knowledge about missense mutations. A combination of WES and array comparative genomic hybridization, RNA sequencing, and capture sequencing of intron lesions may complement efforts to find small deletion and splicing defects caused by missense, synonymous, and deep-in-intron variants in known causative genes.¹⁹ In addition, WES/WGS applications will certainly identify novel IBMFS causative genes, which should continue to increase the genetic diagnostic rate of next-generation sequencing.^{20,21} In this context, periodic reanalysis of the results is desirable. Furthermore, a meticulous combination of clinical judgment and analysis with genetic information is required.

In summary, we analyzed 371 IBMFS patients with two next-generation sequencing platforms and successfully diagnosed 53 of 121 (44%) and 68 of 250 (27%) patients using target sequencing and WES, respectively. Our results demonstrate the efficacy of massive parallel sequencing as a diagnostic tool for IBMFS in clinical practice.

SUPPLEMENTARY MATERIAL

Supplementary material is linked to the online version of the paper at <http://www.nature.com/gim>

ACKNOWLEDGMENTS

This work was supported by the Research on Measures for Intractable Diseases Project from Ministry of Health, Labor, and Welfare and a grant-in-aid from the Ministry of Health, Labor, and Welfare of Japan (H23-TA012). The authors thank all the clinicians, patients, and families who made this study possible by providing samples. The authors also thank Yoshie Miura, Yuko Imanishi, Akitaka Sugishita, and Hiroe Namizaki for their valuable assistance. The authors acknowledge the Division for Medical Research Engineering, Nagoya University Graduate School of Medicine for technical support of the NGS. The authors acknowledge the Human Genome Center, Institute of Medical Science, University of Tokyo for providing supercomputing resources (<http://sc.hgc.jp/shirokane.html>). Finally, the authors are grateful to Enago (<http://www.enago.jp>) for the English-language review.

DISCLOSURE

S.K. received a research grant from Sanofi K.K. The other authors declare no conflict of interest.

REFERENCES

1. Sakaguchi H, Nakanishi K, Kojima S. Inherited bone marrow failure syndromes in 2012. *Int J Hematol* 2013;97:20–29.
2. Longerich S, Li J, Xiong Y, Sung P, Kupfer GM. Stress and DNA repair biology of the Fanconi anemia pathway. *Blood* 2014;124:2812–2819.
3. Ruggero D, Shimamura A. Marrow failure: a window into ribosome biology. *Blood* 2014;124:2784–2792.
4. Townsley DM, Dumitriu B, Young NS. Bone marrow failure and the telomeropathies. *Blood* 2014;124:2775–2783.
5. Bamshad MJ, Ng SB, Bigham AW, et al. Exome sequencing as a tool for Mendelian disease gene discovery. *Nat Rev Genet* 2011;12:745–755.
6. Zhang MY, Keel SB, Walsh T, et al. Genomic analysis of bone marrow failure and myelodysplastic syndromes reveals phenotypic and diagnostic complexity. *Haematologica* 2015;100:42–48.
7. Ghemlas I, Li H, Zlateska B, et al. Improving diagnostic precision, care and syndrome definitions using comprehensive next-generation sequencing for the inherited bone marrow failure syndromes. *J Med Genet* 2015;52:575–584.
8. Richards CS, Bale S, Bellissimo DB, et al.; Molecular Subcommittee of the ACMG Laboratory Quality Assurance Committee. ACMG recommendations for standards for interpretation and reporting of sequence variations: revisions 2007. *Genet Med* 2008;10:294–300.
9. Abecasis GR, Auton A, Brooks LD, et al.; 1000 Genomes Project Consortium. An integrated map of genetic variation from 1,092 human genomes. *Nature* 2012;491:56–65.
10. Kunishima S, Okuno Y, Yoshida K, et al. ACTN1 mutations cause congenital macrothrombocytopenia. *Am J Hum Genet* 2013;92:431–438.
11. Sakaguchi H, Okuno Y, Muramatsu H, et al. Exome sequencing identifies secondary mutations of SETBP1 and JAK3 in juvenile myelomonocytic leukemia. *Nat Genet* 2013;45:937–941.
12. Song WJ, Sullivan MG, Legare RD, et al. Haploinsufficiency of CBFA2 causes familial thrombocytopenia with propensity to develop acute myelogenous leukaemia. *Nat Genet* 1999;23:166–175.
13. Vlachos A, Rosenberg PS, Atsidaftos E, Alter BP, Lipton JM. Incidence of neoplasia in Diamond Blackfan anemia: a report from the Diamond Blackfan Anemia Registry. *Blood* 2012;119:3815–3819.
14. Alter BP, Giri N, Savage SA, Rosenberg PS. Cancer in dyskeratosis congenita. *Blood* 2009;113:6549–6557.
15. Kutler DI, Singh B, Satagopan J, et al. A 20-year perspective on the International Fanconi Anemia Registry (IFAR). *Blood* 2003;101:1249–1256.
16. Myers KC, Bolyard AA, Otto B, et al. Variable clinical presentation of Shwachman-Diamond syndrome: update from the North American Shwachman-Diamond Syndrome Registry. *J Pediatr* 2014;164:866–870.
17. Welte K, Zeidler C. Severe congenital neutropenia. *Hematol Oncol Clin North Am* 2009;23:307–320.
18. Amendola LM, Dorschner MO, Robertson PD, et al. Actionable exomic incidental findings in 6503 participants: challenges of variant classification. *Genome Res* 2015;25:305–315.
19. Chandrasekharappa SC, Lach FP, Kimble DC, et al.; NISC Comparative Sequencing Program. Massively parallel sequencing, aCGH, and RNA-Seq technologies provide a comprehensive molecular diagnosis of Fanconi anemia. *Blood* 2013;121:e138–e148.
20. Wang R, Yoshida K, Toki T, et al. Loss of function mutations in RPL27 and RPS27 identified by whole-exome sequencing in Diamond-Blackfan anaemia. *Br J Haematol* 2015;168:854–864.
21. Hira A, Yoshida K, Sato K, et al. Mutations in the gene encoding the E2 conjugating enzyme UBE2T cause Fanconi anemia. *Am J Hum Genet* 2015;96:1001–1007.