A frequent nonsense mutation in exon 1 across certain HLA-A and HLA-B alleles in leukocytes of patients with acquired aplastic anemia

Hiroki Mizumaki,[1] Kazuyoshi Hosomichi,[2] Kohei Hosokawa,[1] Takeshi Yoroidaka,[1] Tatsuya Imi,[1] Yoshitaka Zaimoku,[1] Takamasa Katagiri,[3] Mai Anh Thi Nguyen,[1] Dung Cao Tran,[1] Mahmoud Ibrahim Yousef Elbadry,[1,4] Kazuhisa Chonabayashi,[5] Yoshinori Yoshida,[5] Hiroyuki Takamatsu,[1] Tatsuhiko Ozawa,[6] Fumihiro Azuma,[7] Hiroyuki Kishi,[6] Yoichi Fujii,[8] Seishi Ogawa,[8,9] Atsushi Tajima[2] and Shinji Nakao[1]

[1]Department of Hematology, Kanazawa University, Kanazawa, Japan; [2]Department of Bioinformatics and Genomics, Graduate School of Advanced Preventive Medical Sciences, Kanazawa University, Kanazawa, Japan; [3]Clinical Laboratory Sciences, Kanazawa University Graduate School, Kanazawa, Japan; [4]Department of Internal Medicine, Division of Hematology, Faculty of Medicine, Sohag University, Sohag, Egypt; [5]Center for iPS Cell Research and Application, Kyoto University, Kyoto, Japan; [6]Department of Immunology, University of Toyama, Toyama, Japan; [7]HLA Laboratory, Japanese Red Cross Kanto-Koshinetsu Block Blood Center, Kotoku, Japan; [8]Department of Pathology and Tumor Biology, Graduate School of Medicine, Kyoto University, Kyoto, Japan and [9]Department of Medicine, Center for Hematology and Regenerative Medicine (HERM), Karolinska Institute, Stockholm, Sweden

 doi:10.3324/haematol.2020.247809

Received: February 16, 2020.

Accepted: May 14, 2020.

Pre-published: May 21, 2020.

Correspondence: *SHINJI NAKAO* - snakao8205@staff.kanazawa-u.ac.jp

Supplemental information

A frequent nonsense mutation in exon 1 across certain HLA-A and -B alleles in leukocytes of patients with acquired aplastic anemia

Hiroki Mizumaki[1], Kazuyoshi Hosomichi[2], Kohei Hosokawa[1], Takeshi Yoroidaka[1], Tatsuya Imi[1], Yoshitaka Zaimoku[1], Takamasa Katagiri[3], Mai Anh Thi Nguyen [1], Dung Cao Tran[1], Mahmoud Ibrahim Yousef Elbadry[1,4], Kazuhisa Chonabayashi[5], Yoshinori Yoshida[5], Hiroyuki Takamatsu[1], Tatsuhiko Ozawa[6], Fumihiro Azuma[7], Hiroyuki Kishi[6], Yoichi Fujii[8], Seishi Ogawa[8,9], Atsushi Tajima[2] and Shinji Nakao[1]

[1]Department of Hematology, Kanazawa University, Kanazawa, Japan;

[2]Department of Bioinformatics and Genomics, Graduate School of Advanced Preventive Medical Sciences, Kanazawa University, Kanazawa, Japan;

[3]Clinical Laboratory Sciences, Kanazawa University Graduate School, Kanazawa, Japan;

[4]Department of Internal Medicine, Division of Hematology, Faculty of Medicine, Sohag University, Egypt;

[5]Center for iPS Cell Research and Application, Kyoto University, Kyoto, Japan;

[6]Department of Immunology, University of Toyama, Toyama, Japan;

[7]HLA Laboratory, Japanese Red Cross Kanto-Koshinetsu Block Blood Center, Kotoku, Japan;

[8]Department of Pathology and Tumor Biology, Graduate School of Medicine, Kyoto University, Kyoto, Japan;

[9]Department of Medicine, Center for Hematology and Regenerative Medicine (HERM), Karolinska Institute, Stockholm, Sweden

Correspondence: Shinji Nakao, M.D, PhD.

Department of Hematology, Graduate School of Medical Sciences, Kanazawa University, 13-1 Takaramachi, Kanazawa, Ishikawa, 920-8641, Japan.

E-mail: snakao8205@staff.kanazawa-u.ac.jp

Phone: +81-76-265-2274

FAX: +81-76-234-4252

SUPPLEMENTAL METHODS

SUPPLEMENTAL TABLES

SUPPLEMENTAL FIGURES

with 6pLOH.

Supplemental Figure 4. Correlation between the variant allele frequencies (VAF) determined by the ddPCR and those determined by target sequencing with the next generation sequencer.

Supplemental Figure 5. HLA-class I allele lacking granulocytes detected in 20 patients of the cohort 1.

Supplemental Figure 6. Determination of the cut-off allele frequency that can be detected by the ddPCR.

SUPPLEMENTAL METHODS

Patients

Twenty Japanese AA patients with HLA-LLs who did not have *HLA-B*40:02* allele were analyzed for the presence of loss-of-function mutations in HLA alleles. A total of 353 Japanese AA patients including the 20 patients (severe and very severe, n=151; non-severe, n=202; male, n=167; female, n=186; median age, 63 [range, 11–93] years) were further analyzed for the prevalence and clinical significance of $Exon1^{mut}$ in AA between 2010 and 2018. Of the 353 patients, 265 had been included in our previous studies (1, 2).

HLA typing and definition of AA severity

All patients were genotyped for HLA-A, HLA-B, HLA-C, and HLA-DRB1 alleles using the polymerase chain reaction (PCR) sequence-specific oligonucleotide method and 6pLOH was detected by GeneChip 500K arrays (Affymetrix) or droplet digital polymerase chain reaction (ddPCR) using a QX200 AutoDG Droplet Digital PCR System (Bio-Rad, Hercules, USA), as previously described (1, 2). HLA haplotypes of patients were estimated from the database of the Japanese population, which includes 18,604 healthy individuals, as previously described (3). Informed consent was obtained from all patients for the HLA-genotyping and genetic analyses. The study protocols

were approved by the ethical committee of Kanazawa University Institute of Medical, Pharmaceutical, and Health Sciences. Severe AA was diagnosed when at least two of the following criteria were met; the neutrophil count was $< 0.5 \times 10^9$/L, the platelet count was $< 20 \times 10^9$/L and the reticulocyte count was $< 20 \times 10^9$/L. Very severe AA was defined as a neutrophil count $< 0.2 \times 10^9$/L in addition to the criteria for severe AA (4). The response criteria have been described previously (5).

Detection and FACS sorting of HLA-LLs

Peripheral blood (PB) samples were stained with anti-HLA-A allele-specific and lineage-specific mAbs to detect leukocytes that lack particular HLA-A or HLA-B alleles (HLA-allele-lacking leukocytes: HLA-LLs). We used mAbs specific for each of HLA-B13, HLA-B60, and HLA-B61 that we developed (2) and HLA-Bw6, as well as mAbs specific for HLA-A, in order to detect HLA-A-positive ($HLA\text{-}A^+$) and HLA-B negative ($HLA\text{-}B^-$) granulocytes, when patients possessed these HLA-B alleles. The mAbs used in this study are summarized in Supplemental Table 1. HLA-LLs were detected using BD FACSCantoII (BD Biosciences, San Jose, CA, USA) and analyzed with the Flowjo (version 10.1) software program (Tree Star, Ashland, OR, USA).

From the PB samples of patients with HLA-LLs, paired fractions, including

granulocytes which lack the HLA class I allele (HLA[-] granulocytes) and granulocytes which retain the HLA class I allele (HLA[+] granulocytes), and $CD3^+$ T cells were sorted using a fluorescence-activated cell sorter (BD FACSAria™ Fusion, BD Biosciences, Franklin Lakes, NJ, USA). Sorted granulocyte populations and control cells ($CD3^+$ T cells or buccal mucosa cells) were subjected to DNA extraction (Supplemental Figure 2). All PB samples from the patients with HLA-LLs were collected during treatment.

Next generation sequencing of HLA class I genes

Nucleotide sequences of HLA-A and HLA-B in sorted granulocytes of the 20 HLA-LL(+) AA patients were determined by using the sequence capture method with a next generation sequencer (NGS). The sequence capture method based on hybridization between DNA of an adapter-ligated library (KAPA Hyper Prep Kit, Roche Diagnostics, Westfield, IN, USA) and a biotinylated DNA probe (SeqCap EZ choice kit, Roche Diagnostics, Westfield, IN, USA) custom-designed based on target sequences of HLA genes.(6) Paired-end sequence reads (350 bp read 1 and 250 bp read 2 in length) were obtained by using the MiSeq sequencer (illumina, San Diego, CA, USA). Somatic mutations were detected as difference from reference sequence with low variant allele

frequency (VAF <1%), and covering with at least 1000 reads. All of the mutations were validated by ultra-deep sequencing of HLA-A and HLA-B locus-specific long-range PCR amplicons (2, 7). The median read depths for HLA-A and HLA-B amplicons sequencing in the 20 patients were 43,031 (23,689-95,332) and 35,267 (19,378-59,846), respectively. HLA alleles were assigned using the Omixon Target software program (version 2.0.0, Omixon, Budapest, Hungary) for read alignment and genotype calling (reference sequence, IMGT/HLA Database 3.26.0).

Novel ddPCR assay for detecting *Exon1^{mut}*

We designed ddPCR assays to precisely detecting *Exon1^{mut}* in the PB using a QX200 AutoDG Droplet Digital PCR System (Bio-Rad Laboratories, Hercules, CA, USA), according to the manufacturer`s instructions. Briefly, we designed two different sets of primer pairs complementary to the consensus sequences of HLA-A and HLA-B alleles, and locked nucleic acid (LNA)-based probes with a non-fluorescent quencher (Iowa Black®FQ[IBFQ]**)** complementary to wild-type (WT) and mutant-specific (MT) sequences, which were labeled with different fluorochromes (hexachlorofluorescein (HEX) for WT and carboxyfluoresein (FAM) for MT) using PrimerQuest and

BioPhysics (Integrated DNA Technologies, Coralville, IA, USA) (Supplemental Table 2).

The reaction mixtures for ddPCR consisted of the 2 LNA probes, 1 set of the primer pairs and 100 ng of genomic DNA, 5 μl of 4×ddPCR Multiplex Supermix (Bio-Rad), and deionized water in a final volume of 20 μl. Cycling conditions were as follows; 95°C/10 min for enzyme activation, followed by 40 cycles of 94°C/30 sec and 59.5°C/1 min, and 98°C/10 min for enzyme deactivation. Droplets were measured with the QX200 droplet reader and were analyzed with Quantasoft software ver 1.7.4 (Bio-Rad), according to the manufacturer's instructions. The fractional abundance of the mutant allele was obtained by dividing the number of copies of the mutant allele (FAM) by the total number of copies of the WT allele (HEX) plus the mutant allele (FAM). Results were considered evaluable when the number of accepted droplets per well was at least 10,000.

Deep sequencing using unique molecular identifiers

In order to confirm the presence of HLA allelic mutations at very low frequencies (VAF<1%), we performed targeted deep sequencing for HLA class I genes by using

sequence capture method with biotin-labelled oligo probes designed for all entire HLA class I genes and a unique molecular identifiers (UMI) for adapters.

To reduce the sequence read errors and increase the sensitivity of rare mutation detection, we used xGen® Dual Index UMI Adapters (Integrated DNA Technologies, Coralville, IA, USA), which reduces amplification errors and enables the accurate detection of true variants at as low as 0.2% VAF, in place of standard Illumina dual index adaptors in this assay (8). Enriched DNA fragments for entire HLA genes were sequenced by paired-end sequencing (2×150 bp) using an illumine HiSeq 4000 (Illumina, San Diego, CA, USA). To ensure accurate variant calling, we used three thresholds, at least 10,000× raw sequencing depth, at least 1,000× collapsed sequencing depth after UMI error correction, and at least 6 reads with alternative allele. The sequence reads carrying $Exon1^{mut}$ were used for HLA typing as HLA allele-specific tagSNP detection.

Next-generation sequencing data processing

Paired-end illumine reads were mapped to the reference genome (GRCh37) using Burrows-Wheeler Aligner (bwa) v.0.7.12.(9) bwa-generated SAM files were converted to the BAM format, then sorted and indexed using SAM tools v.1.2 (10). Duplicated

reads were marked with Picard v.1.52 (https://github.com/broadinstitute/picard). After alignment of reads, the heuristic somatic mutation caller, VarScan 2 (11), was used to detect somatic mutation. The mutations were reviewed using Unified Genotyper in the Genome Analysis Toolkit (GATK) v3.4 (12) and the alignment data from granulocytes, T cells and buccal mucosa cells were visually compared via IGV (13). The somatic mutations were given functional annotation and COSMIC (https://cancer.sanger.ac.uk/cosmic) mutation ID meaning by ANNOVAR (14).

References

1. Katagiri T, Sato-Otsubo A, Kashiwase K, et al. Frequent loss of HLA alleles associated with copy number-neutral 6pLOH in acquired aplastic anemia. Blood. 2011;118(25):6601-6609.

2. Zaimoku Y, Takamatsu H, Hosomichi K, et al. Identification of an HLA class I allele closely involved in the autoantigen presentation in acquired aplastic anemia. Blood. 2017;129(21):2908-2916.

3. Ikeda N, Kojima H, Nishikawa M, et al. Determination of HLA-A, -C, -B, -DRB1 allele and haplotype frequency in Japanese population based on family study. Tissue Antigens. 2015;85(4):252-259.

4. Bacigalupo A, Hows J, Gluckman E, et al. Bone marrow transplantation (BMT) versus immunosuppression for the treatment of severe aplastic anaemia (SAA): a report of the EBMT SAA working party. Br J Haematol. 1988;70(2):177-182.

5. Camitta BM. What is the definition of cure for aplastic anemia? Acta Haematol. 2000;103(1):16-18.

6. Hirata J, Hosomichi K, Sakaue S, et al. Genetic and phenotypic landscape of the major histocompatibilty complex region in the Japanese population. Nat Genet. 2019;51(3):470-480.

7. Hosomichi K, Jinam TA, Mitsunaga S, et al. Phase-defined complete sequencing of the HLA genes by next-generation sequencing. BMC Genomics. 2013;14:355.

8. MacConaill LE, Burns RT, Nag A, et al. Unique, dual-indexed sequencing adapters with UMIs effectively eliminate index cross-talk and significantly improve sensitivity of massively parallel sequencing. BMC Genomics. 2018;19(1):30.

9. Li H, Durbin R. Fast and accurate short read alignment with Burrows-Wheeler transform. Bioinformatics. 2009;25(14):1754-1760.

10. Li H, Handsaker B, Wysoker A, et al. The Sequence Alignment/Map format and SAMtools. Bioinformatics. 2009;25(16):2078-2079.

11. Koboldt DC, Zhang Q, Larson DE, et al. VarScan 2: somatic mutation and copy number alteration discovery in cancer by exome sequencing. Genome Res. 2012;22(3):568-576.

12. McKenna A, Hanna M, Banks E, et al. The Genome Analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. Genome Res. 2010;20(9):1297-1303.

13. Robinson JT, Thorvaldsdottir H, Winckler W, et al. Integrative genomics viewer. Nat Biotechnol. 2011;29(1):24-26.

14. Wang K, Li M, Hakonarson H. ANNOVAR: functional annotation of genetic variants from high-throughput sequencing data. Nucleic Acids Res. 2010;38(16):e164.

SUPPLEMENTAL TABLES

Supplemental Table 1. The monoclonal antibodies used for flow cytometry

Antigen	Isotype	Conjugate	Source
CD3	Mouse IgG1	PerCP-Cy5.5	BD Biosciences
CD19	Mouse IgG1	APC-Cy7	Beckman Coulter
CD33	Mouse IgG1	APC	Beckman Coulter
HLA-A2/28	Mouse IgG2a	FITC	One Lambda
HLA-A9/24	Mouse IgG2b	FITC	One Lambda
HLA-A9/24	Mouse IgG2b	Biotin	One Lambda
HLA-A30/31	Mouse IgM	Biotin	One Lambda
Streptavidin	NA	PE	BD Biosciences
HLA-Bw6	Mouse IgG2a	FITC	One Lambda
Mouse Ig	Goat polyclonal Ig	BV421	BD Horizon
HLA-B13/B60/B61	Human/mouse chimeric IgG with constant domein of mouse IgG1	None	Original

Abbreviations: Ig, immunoglobulin; PerCP-Cy5.5, peridinin-chlorophyll proteins-Cy5.5 tandem; PE, phycoerythrin; APC, allophycocyanin; APC-Cy7, allophycocyanin-Cy7 tandem; FITC, fluorescein isothiocyanate; BV421, BD Horizon Brilliant Violet 421

Supplemental Table 2. Sequence of primers and probes

Name	Description	Sequence（5’→3’）
primer set of HLA-A	Forward Primer	CGAGGATGGCCGTCATG
	Reverse Primer1	GGACACGGAGGTGTAGAAATAC
	Reverse Primer2	GGGACACGGATGTGAAGAAATA
	Reverse Primer3	GGACACGGATGTGGAGAAATA
	Reverse Primer4	GGACACGGATGTGGTGAAATA
primer set of HLA-B	Forward Primer1	GTCTCCTCAGACACCGAGAT
	Forward Primer2	TCTCCTCAGACGCCAAGAT
	Forward Primer3	TCCTCAGACGCCGAGAT
	Reverse Primer1	CTCCCTACAGAGGCCATTTC
	Reverse Primer2	CCGGCAGAGGCCATTTC
	Reverse Primer3	CCCACAGAGGCCATTTCC
Probe set	LNA probe for WT (HEX)	CC+C+C+GAA+CC
	LNA probe for MT (FAM)	C+C+C+T+GAA+CC

Abbreviations: LNA, locked nucleic acid; FAM, carboxyfluoresein; HEX, hexachlorofluorescein; MT, mutant-specific sequences; WT, wild-type specific sequences

+C, +G, and +T indicates bases which connect to LNA.

Supplemental Table 3. Patient characteristics

UPN	Age at diagnosis	Sex	Disease duration, year	Severity	Treatment	Time after IST, year	Response to IST	% of GPI(-) cells at diagnosis	
								Granulocytes	RBCs
211	73	F	1.7	NSAA	CSA	1.6	CR	0.061	0.036
335	18	M	1.7	VSAA	ATG+CSA+EPAG	1.7	NR	0.024	0.009
338	54	M	0.8	SAA	ATG+CSA+EPAG	0.7	NR	0.079	0.050
248	61	F	2.8	SAA	ATG+CSA	2.8	PR	0.012	0.000
262	75	M	0.3	SAA	CSA	0.3	PR	0.103	0.047
349	68	F	14.3	NSAA	CSA	14.3	CR	0.001	0.001
344	25	F	15.8	NSAA	CSA	15.7	CR	0.405	0.207
339	47	M	11.8	NSAA	AS	11.6	CR	0..000	0.000
333	53	M	3.9	NSAA	CSA	3.9	PR	3.499	0.407
309	51	F	3.1	VSAA	Romi+AS	NE	CR	0.000	0.000
334	81	M	10.3	SAA	ATG+CSA	10.3	CR	0.001	0.000
336	20	M	0.2	SAA	ATG+CSA	0.1	CR	1.833	0.085
210	80	M	2.1	NSAA	CSA＋EPAG	0.9	NR	7.33	0.700
340	42	M	12.6	SAA	CSA	12.6	CR	0.180	0.340
353	52	M	6.8	SAA	CSA	NE	CR	12.324	0.524
348	70	M	1.3	NSAA	CSA+EPAG+AS	1.3	PR	0.706	0.069
235	35	M	0.4	NSAA	ATG+CSA	0.3	CR	0.101	0.092
346	55	F	24.8	NSAA	None	NE	NE	0.172	1.039
242	68	M	0.5	NSAA	ATG+CSA	0.3	CR	0.058	0.035
285	69	M	0.2	SAA	ATG+CSA+EPAG	0.2	NR	0.078	0.008

Supplemental Table 3. Patient characteristics (continued)

UPN	% of HLA (-) granulocytes in total granulocytes		A	B	C	DRB1	A	B	C	DRB1
	Type of granulocytes	%	haplotype (1)				haplotype (2)			
211	A2(+)B4001(-)G	9.3	02:01	40:01	03:03	15:02	31:01	51:01	14:02	08:02
335	A2(-)G	90.0	02:01	40:03	03:04	14:05	03:01	44:02	05:01	13:01
338	A2(-)G	31.9	02:07	54:01	07:04	12:01	24:02	15:18	01:02	04:05
248	A2(+)B1301(-)G	29.8	02:01	13:01	03:04	12:02	24:02	52:01	12:02	15:02
262	A2(-)G	47.5	02:06	39:01	07:02	15:01	11:01	39:01	01:02	15:01
349	A31(-)G	99.4	31:01	48:01	03:04	09:01	02:06	39:01	07:02	15:01
344	A2(-)G	95.4	02:06	07:02	07:02	01:01	01:01	37:01	06:02	15:01
339	A2(-)G	80.1	02:06	39:01	07:02	08:05	24:02	15:18	07:04	04:01
333	A24(+)Bw6(-)G	74.4	24:02	54:01	01:02	04:01	01:01	37:01	06:02	10:01
309	A2(-)G	99.8	02:06	54:01	01:02	15:01	24:02	46:01	01:03	09:01
334	A2(+)B4001(-)G	22.0	24:02	40:01	07:02	12:01	11:01	67:01	07:02	15:01
336	A2(-)G	6.4	02:06	39:01	08:03	15:01	24:02	54:01	07:02	04:05
210	A24(-)G	30.1	24:02	54:01	01:02	04:05	02:06	35:01	03:03	15:02
340	A31(-)G	99.2	31:01	51:01	14:02	14:05	24:02	07:02	07:02	01:01
353	A2(-)G	2.4	02:06	48:01	08:03	16:02	26:02	15:01	03:03	15:01
348	A24(-)G	38.3	24:02	52:01	12:02	15:02	02:06	39:01	07:02	15:01
235	A2(-)G	2.9	02:06	59:01	01:02	04:05	24:02	07:02	07:02	01:01
346	A31(-)G	99.6	31:01	35:01	03:03	09:01	26:01	39:01	07:02	08:03
242	A2(-)G	20.5	02:06	07:02	07:02	15:02	33:03	44:03	14:03	13:02
285	A2(-)G	19.3	02:06	39:01	07:02	15:01	24:02	52:01	12:02	15:02

Abbreviations: UPN, unique patient number; F, female; M, male; GPI-AP(-), glycosylphosphatidylinositol-anchored proteins deficient;

NSAA, non-severe aplastic anemia; SAA, severe aplastic anemia; VSAA, very severe aplastic anemia; PNH, paroxysmal nocturnal hemoglobinuria; CsA, cyclosporine; ATG, antithymocyte globulin; AS, anabolic steroids; EPAG, eltrombopag; Romi, romiplastim; IST, immunosuppressive therapy; CR, complete response; PR, partial response; NR, no response; NE, not evaluated; G, granulocytes.

Supplemental Table 4. Loss-of-function mutations identified in HLA class I genes

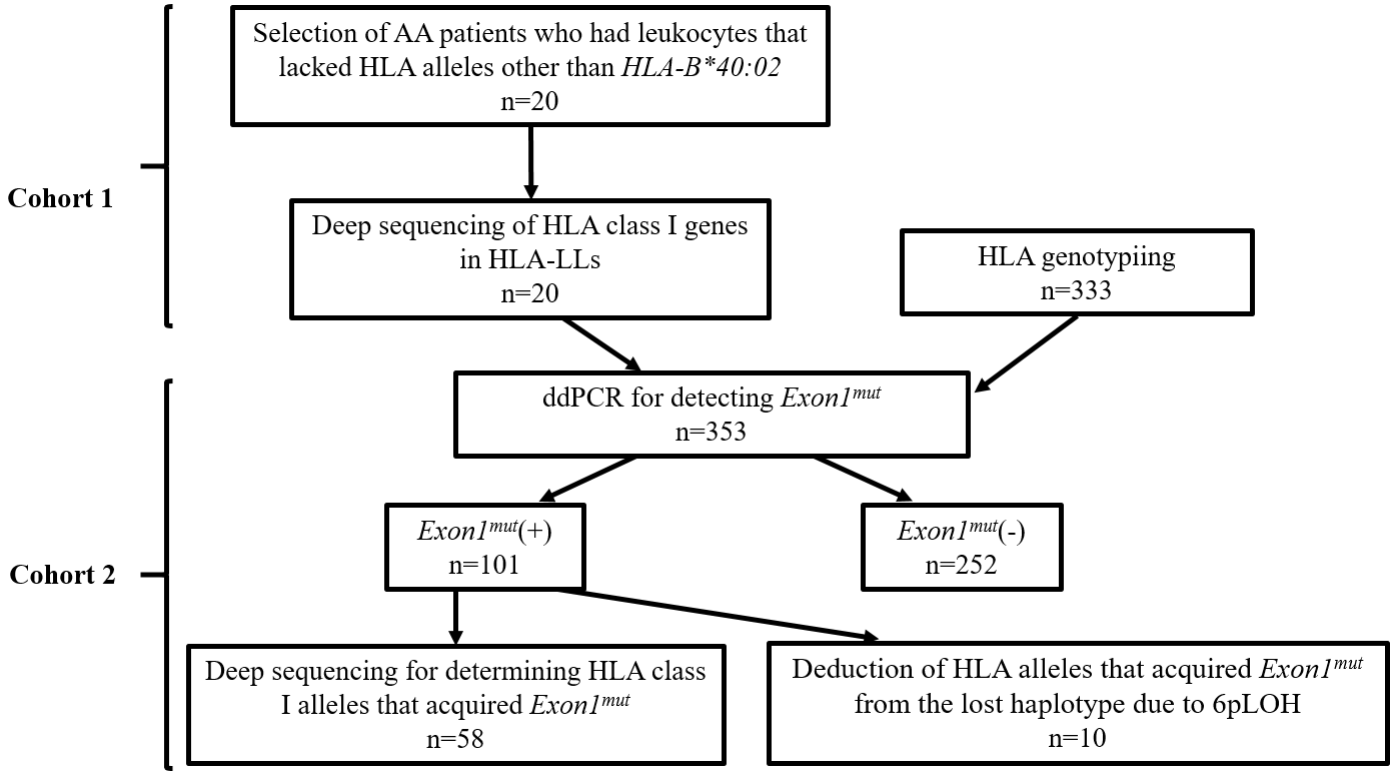
UPN	Mutation (n)	Mutated allele	Type of Mutation	Exon	Mutation (coding)	Mutation (protein)	Variant Allele Frequency (%)	COSMIC ID
211	1	B*40:01	Nonsense	1	c.19C>T	p.R7*	37.3	COSM3253080
334	3	B*40:01	Nonsense	1	c.19C>T	p.R7*	8.6	COSM3253080
			Frameshift insertion	3	c.438dupT	p.Y147Lfs	3.8	Novel
			Frameshift insertion	4	c.625dupC	p.T211Dfs	25.2	COSM5494349
262	3	A*02:06	Nonsense	1	c.19C>T	p.R7*	6.1	COSM3722056
			Missense	1	c.47C>A	p.A16D	19.6	Novel
			Nonsense	4	c.724C>T	p.Q242*	2.0	Novel
335	1	B*40:03	Frameshift deletion	5	c.907delC	p.Q303Sfs	25.4	Novel
242	3	A*02:06	Nonsense	1	c.19C>T	p.R7*	2.3	COSM3722056
			Frameshift deletion	3	c.402delC	p.L135Afs	17.8	Novel
			Frameshift insertion	3	c.556dupG	p.T187Hfs	2.9	Novel
336	3	A*02:06	Nonsense	1	c.19C>T	p.R7*	8.1	COSM3722056
			Frameshift deletion	4	c.831delG	p.E278Sfs	1.2	Novel
			Frameshift deletion	5	c.969delT	p.G324Efs	18.5	Novel
333	4	B*54:01	Nonsense	1	c.19C>T	p.R7*	7.1	COSM3253080
			start loss	1	c.1A>G	p.L2_M4del	3.3	Novel
			start loss	1	c.3G>A	p.L2_M4del	2.0	Novel
			Frameshift deletion	3	c.559_568del	p.E187Sfs	8.9	Novel
248	6	B*13:01	Nonsense	1	c.19C>T	p.R7*	15.1	COSM3253080
			Non frameshift insertion	1	c.27_28insCG	p.L10Gfs	1.6	Novel
			Frameshift deletion	2	c.209_220del	p.E70_P74delinsA	1.3	Novel
			stop gain	4	c.657dupT	p.D220fs	1.4	Novel
			Nonsense	4	c.748C>T	p.Q250*	1.9	COSM4833544
			stop gain	4	c.842dupA	p.Y280fs	1.3	Novel

339	1	A*02:06	Nonsense	1	c.19C>T	p.R7*	3.5	COSM3722056
349	3	A*31:01	Nonsense	1	c.19C>T	p.R7*	19.1	COSM3722056
			Frameshift insertion	3	c.417_418insGGAC	p.D140Gfs	1.2	Novel
			Frameshift insertion	4	c.621dupC	p.K207Qfs	1.8	COSM4765125
344	5	A*02:06	Nonsense	1	c.19C>T	p.R7*	9.3	COSM3722056
			Frameshift insertion	1	c.65dupC	p.W23Lfs	2.4	Novel
			Frameshift insertion	2	c.164_165insT	p.T55fs	1.5	Novel
			Non frameshift insertion	2	c.245_246insAGT	p.E82_Y83insV	2.0	Novel
			Nonsense	4	c.843C>A	p.Y281*	2.1	Novel
210	2	B*54:01	Frameshift deletion	1	c.19delC	p.R7Efs	1.4	Novel
			Frameshift deletion	3	c.577delC	p.R193Afs	2.7	Novel
235	3	A*02:06	Nonsense	1	c.19C>T	p.R7*	11.8	COSM3722056
			Frameshift deletion	3	c.474_475delCG	p.A159Gfs	4.5	Novel
			Frameshift deletion	4	c.772delA	p.R258Gfs	2.3	Novel
353	1	A*02:06	Nonsense	1	c.19C>T	p.R7*	31.2	COSM3722056

Red letters indicate *Exon1*^{mut} in HLA-A and –B genes.

SUPPLEMENTAL Figures

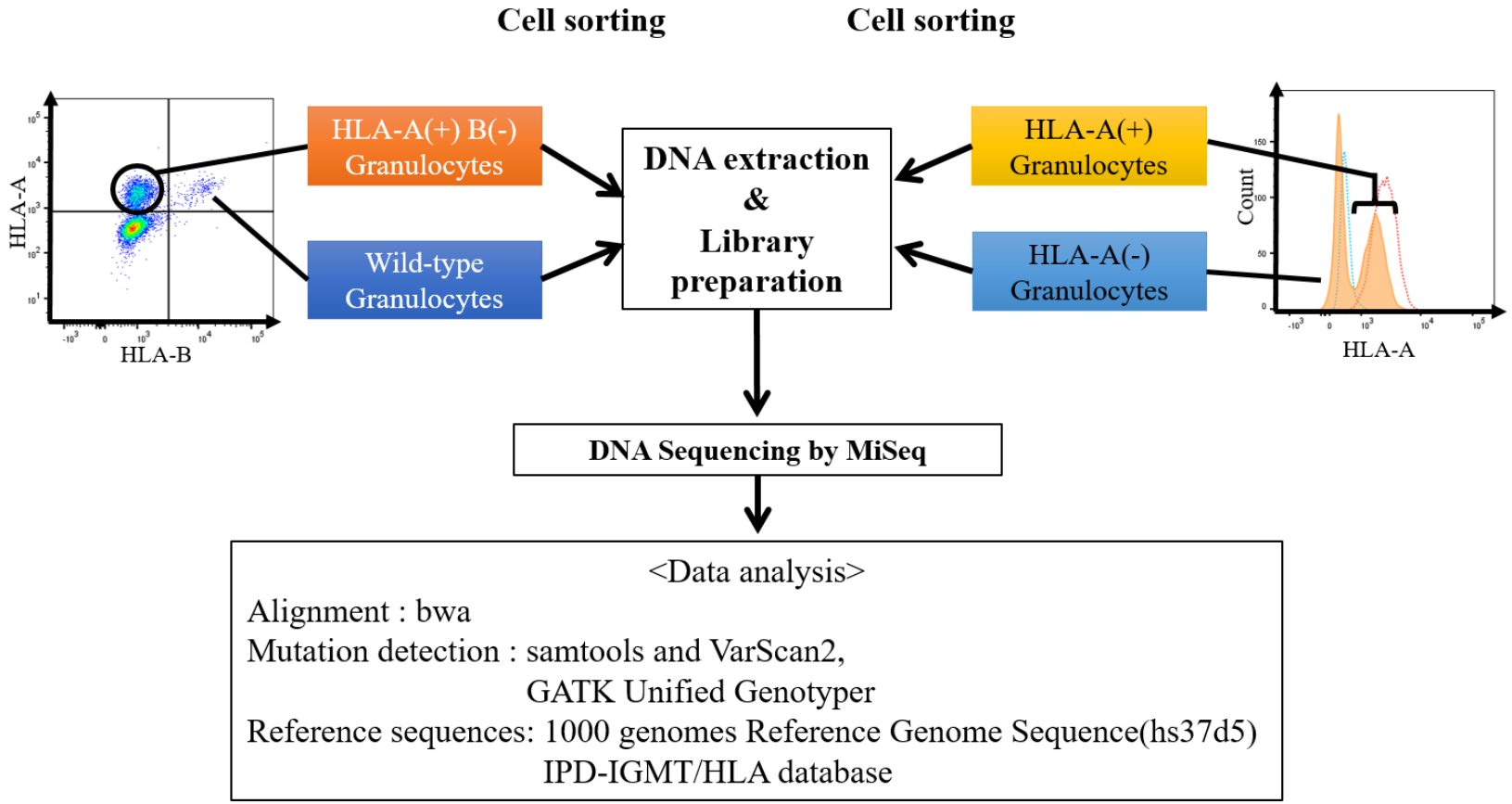
Supplemental Figure 1. Work flow diagram



Abbreviations: AA, aplastic anemia; HLA-LLs, HLA-allele-lacking leukocytes; 6pLOH, copy number neutral loss of heterozygosity of the short arm of chromosome 6.

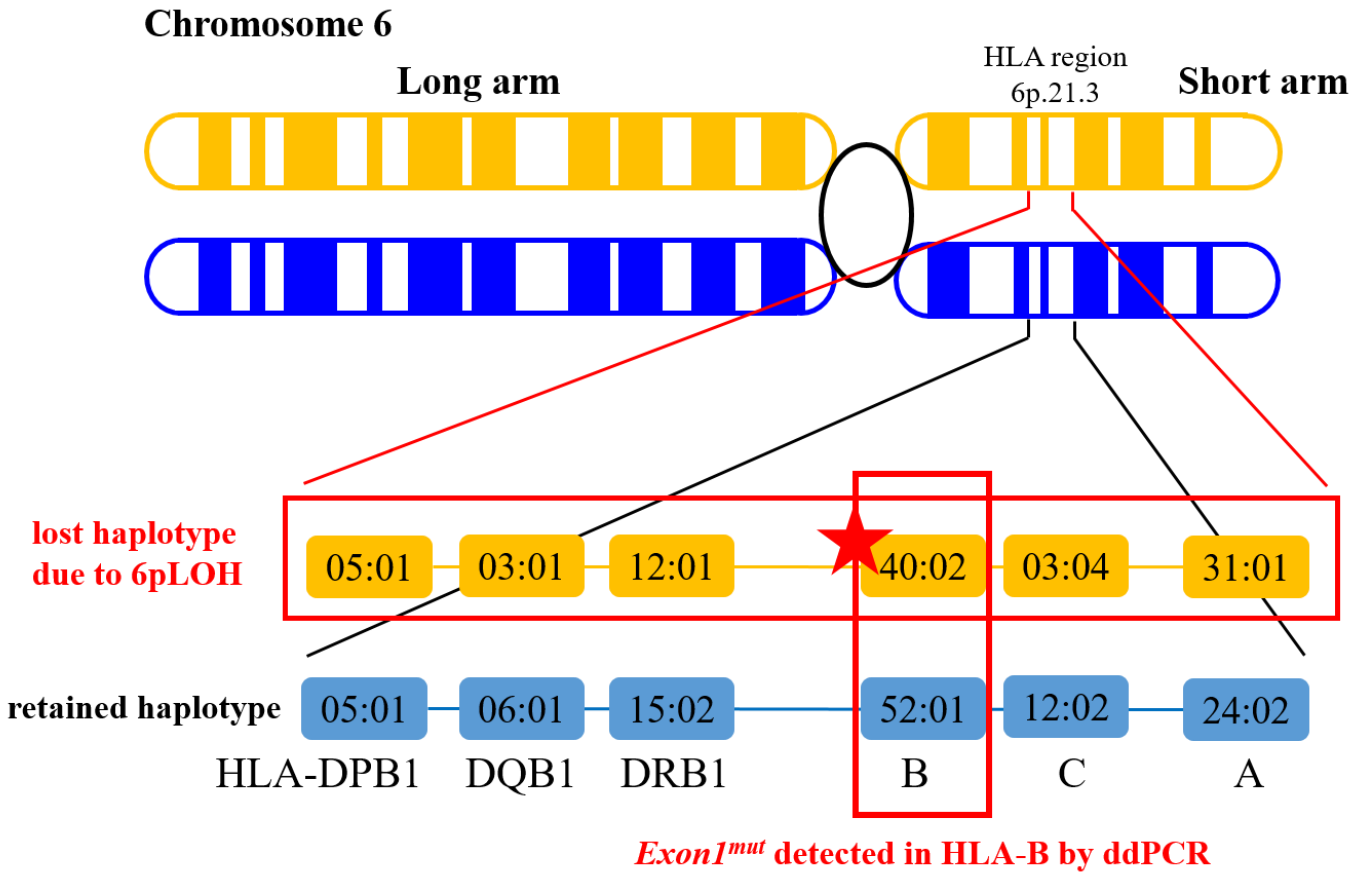
Each number denotes the number of patients studied.

Supplemental Figure 2. Deep sequencing of HLA class I genes



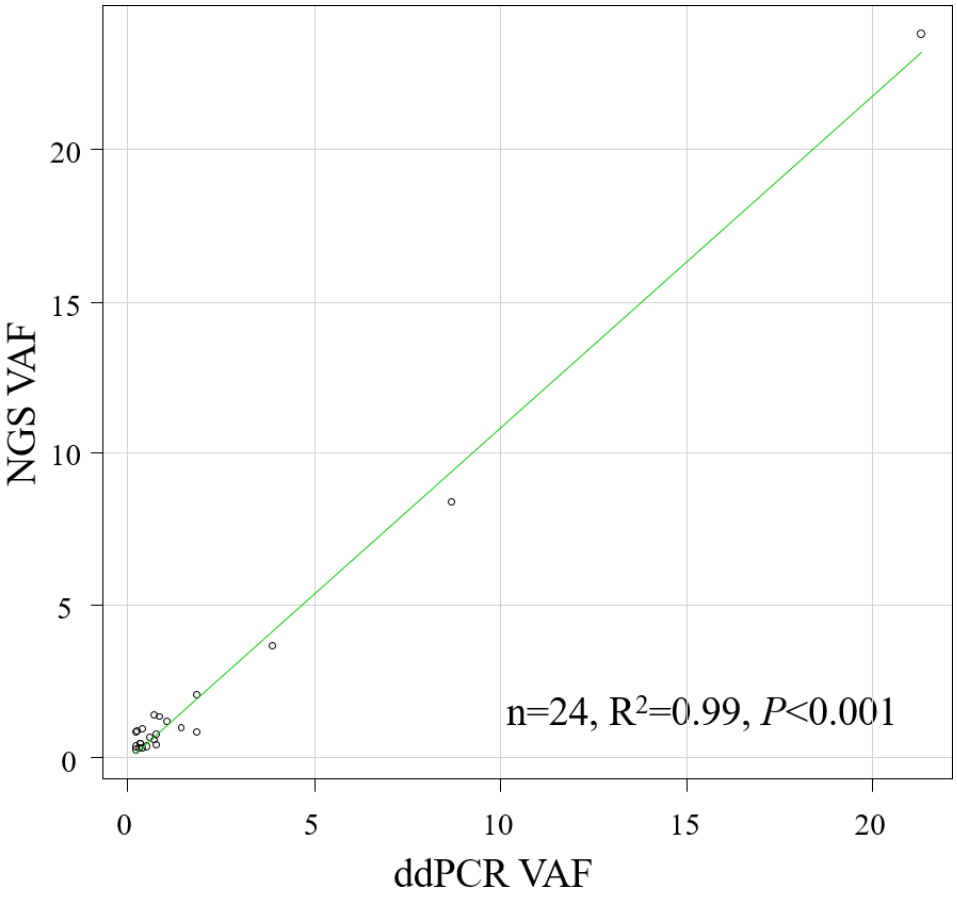
HLA(+) and HLA(-) granulocytes were sorted and subjected to targeted sequencing using a next generation sequencer (MiSeq, illumina). A representative scattergram of a patient (UPN 334) who had HLA-B4001-lacking granulocytes is shown.

Supplemental Figure 3. Estimation of the allele that acquiring *Exon1*^{mut} in a patient with 6pLOH



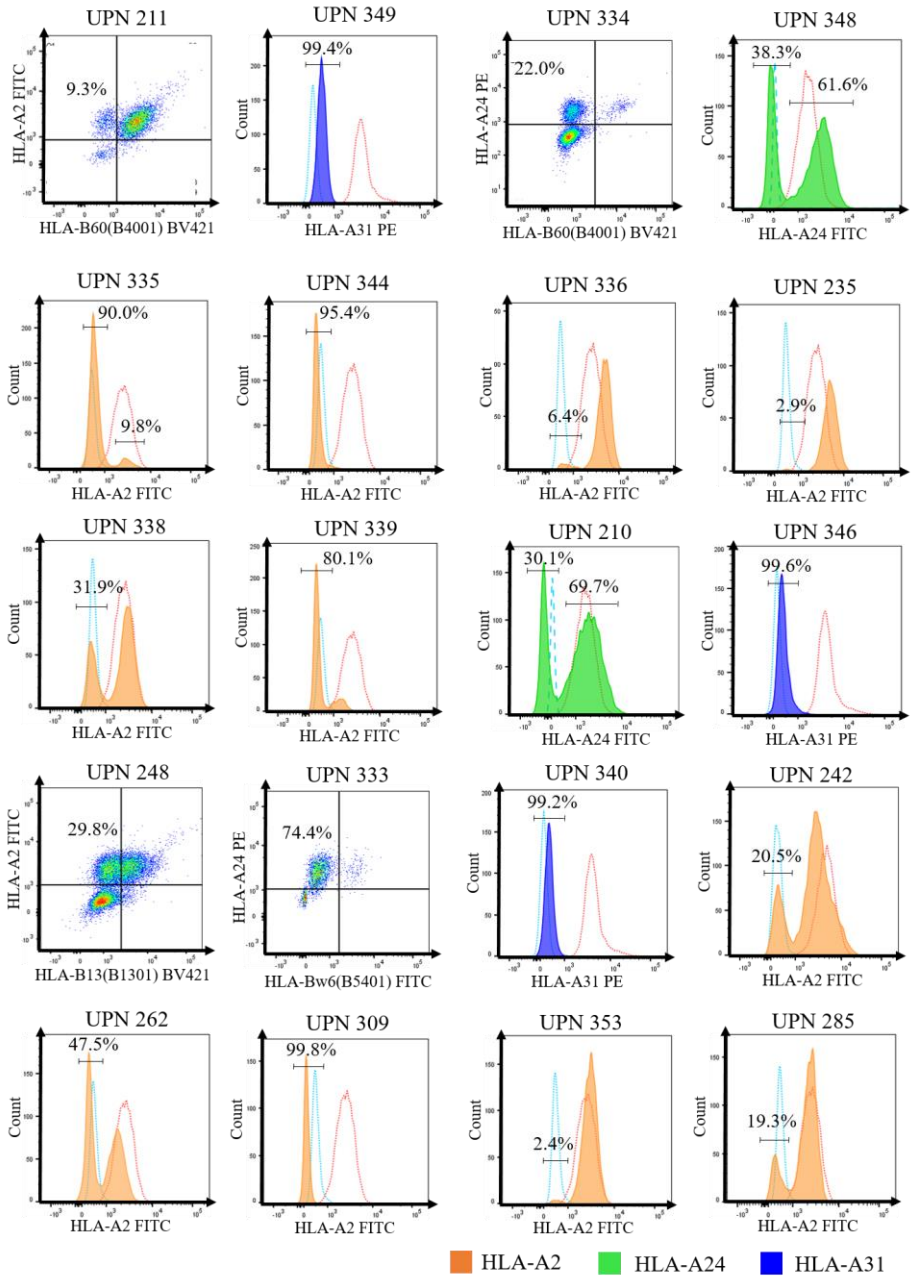
In this patient (UPN 341) whose HLA-B allele acquired *Exon1*^{mut} but could not be determined due to its very low VAF. The affected allele was estimated from an allele contained in the lost haplotype due to 6pLOH that was accompanied by *Exon1*^{mut}.

Supplemental Figure 4. Correlation between the variant allele frequencies (VAFs) determined by the ddPCR assay and those determined by deep sequencing with a next generation sequencer



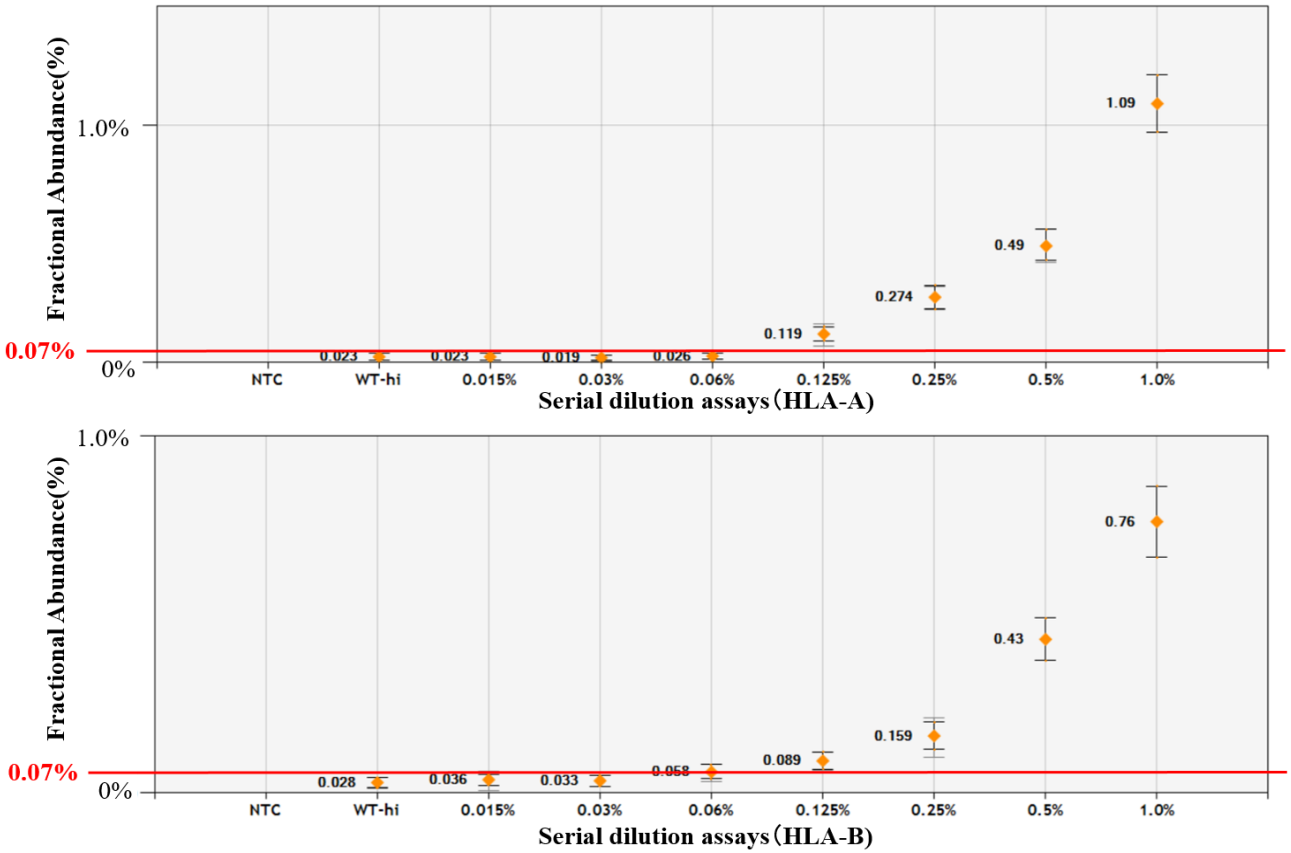
The correlation between *Exon1*^{mut}, VAFs determined by NGS with UMI and those determined by the ddPCR assay was examined using 24 different samples. The results of variant allele frequency (VAF) of the 24 samples are shown. The results of the deep sequencing with UMI were positively correlated with those of the ddPCR assay in the 24 samples (P<0.001, R^2=0.99).

Supplemental Figure 5. HLA-class I allele lacking granulocytes detected in 20 patients of the cohort 1



HLA-class I allele-lacking granulocytes and their percentages in the 20 patients are shown. The open histogram represents HLA-A expression of normal controls, including negative (blue) and positive controls (orange).

Supplemental Figure 6. Determination of the cut-off allele frequency detectable by the ddPCR



Tested samples containing a fixed amount of wild-type DNA and serial dilutions of *Exon1*^{mut} template DNA revealed a detection limit of 0.07% for both HLA-A and HLA-B. None of 24 healthy samples showed >0.042% *Exon1*^{mut} DNA (data not shown).