Received: 24 April 2024 | Revised: 7 September 2024 | Accepted: 20 September 2024

DOI: 10.1111/vox.13747

ORIGINAL ARTICLE

Vox Sanguinis ISBT International Society of Blood Transfusion

A novel approach to simultaneous genotyping of human platelet antigen systems and human leucocyte antigen class I loci using PacBio long-read sequencing

Peizhe Zhao[1] | Qilu Lyu[2] | Yi Xu[1] | Yajun Liang[1] | Yunxiang Wu[1] | Qing Li[1] | Hua Wang[1] | Yao Yuan[3] | Runjun He[3] | Weiyi Fu[3] | Demei Zhang[1] | Yujie Kong[4,5]

[1]Department of Blood Transfusion, Taiyuan Blood Center, Taiyuan, China

[2]Institute for Clinical and Experimental Transfusion Medicine, Medical Faculty of Tuebingen, Tübingen, Germany

[3]Xi'An Haorui Gene Technologies Ltd, Xi'An, China

[4]Department of Laboratory, The First Affiliated Hospital of Chengdu Medical College, Chengdu, China

[5]School of Clinical Medicine, Chengdu Medical College, Chengdu, China

Correspondence
Demei Zhang, Department of Blood Transfusion, Taiyuan Blood Center, No. 185, Yingze West St, Wanbailin District, Taiyuan, Shanxi, China.
Email: 13834532717@126.com

Yujie Kong, Department of Laboratory, the First Affiliated Hospital of Chengdu Medical College, Chengdu, Sichuan 610599, China.
Email: 645539512@qq.com

Funding information
Health Commission of Shanxi Province, Grant/Award Number: 2021XM13; Weigao Scientific Research Foundation of CSBT, Grant/Award Number: CSBT-MWG-2023-01; The First Affiliated Hospital of Chengdu Medical College, Grant/Award Number: CYFY-GQ52

Abstract

Background and Objectives: Accurate human leucocyte antigen (HLA) and human platelet antigen (HPA) typing is essential for establishing a blood platelet donor bank to deal with refractoriness in patients undergoing multiple platelet transfusions. Current methods, such as Sanger and next-generation sequencing, encounter difficulties in haplotyping. Herein, the aim of this study was to establish a method for HLA and HPA typing based on the long read sequencing.

Study Design and Methods: The HPA and HLA class I genotypes of 268 platelet donors from the Taiyuan Blood Center, China were identified using long-read sequencing on the PacBio platform. Allele frequencies for HPA systems and HLA class I genes were calculated, and genetic variability within HPA system genes was analysed.

Results: Polymorphisms were identified in 8 of the 35 HPA systems (HPA-1 to HPA-6w, HPA-15 and HPA-21w), with the frequencies of the 'b' allele at 0.0187, 0.0709, 0.4086, 0.0075, 0.0149, 0.0317, 0.4310 and 0.0019, respectively. The alleles with the highest frequencies at the HLA-A, HLA-B and HLA-C loci are HLA-A02:01, B51:01, B46:01 and C06:02, respectively. Additionally, several genetic patterns in HPA systems were identified, including the c.166-1029C>T variant, which was found exclusively in samples carrying the HPA-1b allele.

Conclusion: This study developed a targeted long-read sequencing method characterized by high throughput and simultaneity, capable of resolving allele ambiguities for effective HLA class I genotyping in establishing a platelet donor bank.

Peizhe Zhao and Qilu Lyu contributed equally to this work and shared the first authorship.

Keywords
HLA, HPA, long-read sequencing, platelet transfusion

Highlights

- Herein, we report the polymorphism of human leucocyte antigen (HLA) and human platelet antigen (HPA) systems of blood donors in Taiyuan Blood Center (North of China).
- Our findings identify previously undocumented haplotype patterns in HPA systems.
- We developed a targeted long-read sequencing method characterized by high throughput and simultaneity, capable of resolving allele ambiguities for HLA class I genotyping in establishing a platelet donor bank.

INTRODUCTION

Platelets are a critical haematological component with diverse functions, including roles in thrombosis, haemostasis and immune regulation [1]. The human platelet membrane displays various antigens, such as blood group antigens, human leucocyte antigen (HLA) class I proteins, human platelet antigens (HPA) and cluster of differentiation 36 (CD36) antigens, which are critical for platelet transfusion medicine [2–5]. Patients receiving platelet transfusions that are not properly matched for HLA, HPA and/or CD36 are at higher risk of developing alloantibodies, as evidenced in cases of platelet transfusion refractoriness (PTR) [6]. PTR is characterized by a recurrent suboptimal response to platelet transfusions, marked by a lower-than-expected increase in post-transfusion platelet counts. HLA class I, HPA and CD36 antibodies are the most common contributors to PTR in patients [6]. Thus, HLA and/or HPA-matched platelet transfusions are recommended to reduce PTR [7].

HLA is categorized into three classes: class I, class II and class III. Within the HLA class I region, the classical and highly polymorphic HLA-A, HLA-B and HLA-C genes coexist with the non-classical and less polymorphic HLA-E, HLA-F and HLA-G genes [8, 9]. However, HLA-C is frequently excluded from platelet donor banks in many centres, likely due to its relatively unclear clinical significance and associated cost considerations [8, 10]. Nevertheless, studies have indicated that HLA-C may contribute to the development of PTR, highlighting the importance of considering its potential role in PTR [11]. Moreover, given the latest HPA database, 35 HPAs have been identified, located on Integrin Subunit Beta 3 (*ITGB3*), *Glycoprotein Ib Platelet Subunit Alpha (GP1BA)*, *Integrin Subunit Alpha 2b (ITGA2B)*, *Integrin Subunit Alpha 2 (ITGA2)*, *Glycoprotein Ib Platelet Subunit Beta (GP1BB)*, *Glycoprotein IX (GP9)* and Cluster of Differentiation 109 (*CD109*) [12]. Comprehensive studies and databases focusing on HPA genotypes across diverse populations are crucial for elucidating the distribution and frequency of various HPA alleles. Such insights would greatly enhance our understanding of platelet transfusion strategies and aid in the prevention of alloimmunization-related complications.

HLA and HPA systems have a high polymorphism in human genes, and genotyping methods such as Sanger sequencing and next-generation sequencing (NGS) are used to identify donor HLA and HPA genotypes. Sanger sequencing, although regarded as the gold standard for gene analysis, encounters challenges in HLA genotyping due to ambiguities and is not well-suited for high-throughput applications [13, 14]. NGS also has some limitations, like short reads resulting in phase analysis difficulties in HLA genotyping, it can pose significant challenges for NGS sequencing due to limited capability in detecting large structural variations, such as large insertions, deletions and inversions. Short read lengths may not fully cover the entire variant region, leading to the inability to identify or misclassify these variants in the haplotype phasing analysis [15]. Recent advancements in third-generation sequencing (TGS) technologies have revolutionized molecular detection, offering tremendous sequencing capacity with unprecedented depth and accuracy [16]. Recently, PacBio [17] and Nanopore [18] techniques have emerged, offering the advantage of sequencing full-length haplotype sequences for human genes and addressing the limitations of Sanger sequencing and NGS.

Despite the advancements, limited data exists on the application of long-read sequencing in platelet donor bank establishments. In this study, we utilized the PacBio platform for typing HLA-A, HLA-B, HLA-C and HPA-related genes to establish an extendedly typed platelet donor bank. Additionally, a comprehensive analysis of the sequences of these genes was conducted.

MATERIALS AND METHODS

Sample collection

In this study, we included 268 unrelated and randomly selected platelet donors who had each donated at least three times at the Shanxi Province Blood Center, China. All participants were of Han ethnicity and provided informed consent. For each donor, 5 mL of peripheral whole blood was collected using ethylenediaminetetraacetic acid (EDTA) as an anticoagulant.

Genomic DNA extraction

Genomic DNA was extracted from the samples using the QIAamp DNA Blood Mini Kit (QIAGEN, Valencia, CA, USA), following the

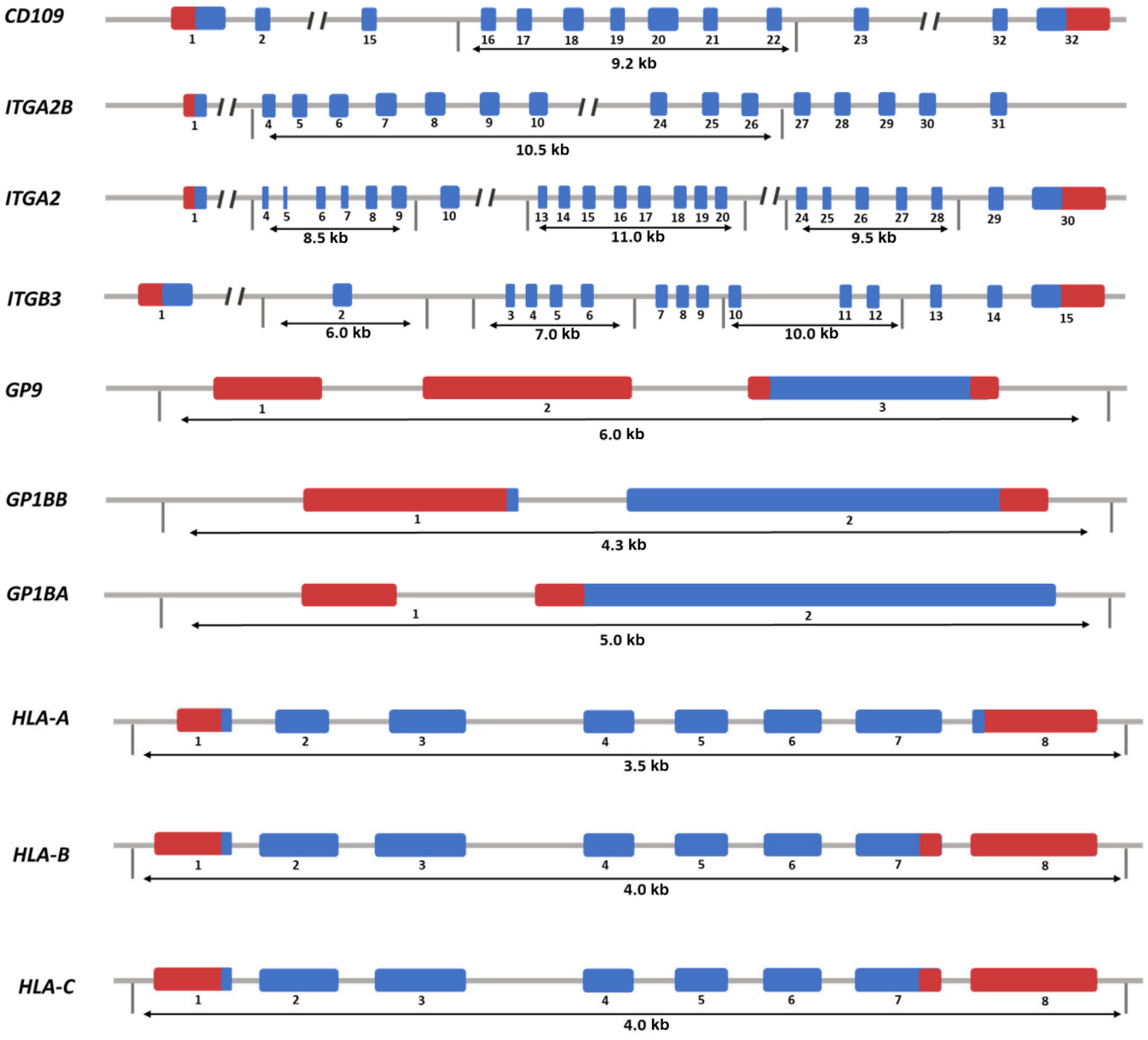

FIGURE 1 LR-PCR amplification of Long-Range PCR HLA and HPA core region. The region indicated by the arrow represents the amplified area; Red and blue squares represent the untranslated regions and open reading frame.

manufacturer's instructions. The DNA was quantified using a Qubit 4.0 Fluorometer (Thermo Fisher Scientific, Wilmington, DE, USA) with a Double-stranded DNA (dsDNA) Deoxynucleotide Triphosphates HS Assay Kit (Thermo Fisher Scientific, Wilmington, DE, USA). The final concentration of the DNA was adjusted to 30 ng/μL.

Long-Range PCR (LR-PCR) Circular consensus sequencing amplification for PacBio SMRT sequencing

In this study, a blood group gene panel was designed, encompassing 10 platelet transfusion-related genes, including the entire genes of *HLA-A*, *HLA-B*, *HLA-C*, *GP1BA* and *GP9*, *GP1BB*, and the core regions of *ITGB3*, *ITGA2B*, *ITGA2* and *CD109*, covering HPA 1-35w. The amplification regions are depicted in Figure 1. The primers are listed in Table S1.

KOD FX Neo (Toyobo, Osaka, Japan) was utilized according to the instructions for LR-PCR. The reaction system consists of KOD Neo FX Buffer 12.5 μL, Deoxynucleotide Triphosphates (dNTPs) coding sequence (2 mM) 5 μL, total primers 1 μL, genomic DNA 6 μL and KOD Neo FX 0.5 μL. The PCR cycling conditions were set: denaturation at 94°C for 2 min each cycle, following 30 cycles of 98°C for 12 s and 68°C for 12 min for denaturation and annealing and the last step with 68°C for 10 min. PCR products were then purified using 0.6× Ampure PB beads (Pacific Biosciences, Menlo Park, CA, USA).

Library preparation and PacBio SMRT sequencing

Single-molecule real-time (SMRT) libraries were prepared using a one-step method that integrates DNA damage repair, end-repair and adapter ligation reactions to generate pre-sequencing libraries with unique barcode adapters. A 10 μL reaction master mixture was prepared, consisting of 4 μL of PCR product, 5 μM barcoded adapter (Integrated DNA Technologies), 1× T4 DNA ligase buffer (Enzymatics), 1 mM ATP (New England Biolabs), 200 μM dNTP (New England Biolabs), 2.5 units of T4 polynucleotide kinase (Enzymatics), 0.75 units of T4 DNA-polymerase (Enzymatics) and 180 units of T4 DNA ligase (Enzymatics).

For SMRT library construction, 120–250 ng of purified PCR amplicon was combined with the enzyme mixture and incubated at 37°C for 20 min, 25°C for 15 min and 65°C for 10 min. Exonuclease I and Exonuclease III were subsequently added to remove un-ligated products. The final pre-library was purified using 0.6× Ampure PB beads. For multiple sample sequencing, pre-libraries were pooled based on equal masses and subjected to two rounds of purification with 0.45× Ampure PB beads.

The final library was prepared with sequencing enzymes and primers using Sequel Binding Kit 2.2 and Internal Control Kit 1.0 (Pacific Biosciences, Menlo Park, CA, USA). DNA-polymerase complexes at a concentration of 150 pM were then loaded onto the Sequel II platform (Pacific Biosciences, Menlo Park, CA, USA) for sequencing with a 20-h run time.

Data analysis and DNA variant calling

Sequencing data were processed using SMRT Link v10.1.0 software. Raw reads were de-multiplexed, and barcode sequences were automatically identified. Subreads were analysed to generate HiFi reads using HiFi software. The HiFi reads were clustered with the Pbaa tool on the SMRT Link platform to correct random errors and enhance data quality. For HPA-related genes, filtered HiFi reads were aligned to the human reference genome (GRCh38) using pbmm2. Alignments specifically targeted Circular consensus sequencing (CCS fragments) from *GP1BA*, *GP9*, *GP1BB*, *ITGB3*, *ITGA2B*, *ITGA2* and *CD109*. The target CCS reads were then realigned using pbmm2. Single nucleotide variants (SNVs), small insertions and deletions (indels), and variant calling were performed using DeepVariant v1.2.0. The Integrative Genomics Viewer was utilized to display alignments and call genotypes.

For HLA genotyping, coding sequence (CDS) sequences were aligned with the IPD-IMGT/HLA Database, and the best-aligned reads were identified as candidate genotypes and ranked in descending order. The data were subsequently clustered using PBAA (PacBio Amplicon Analysis, Version 1.0.3) for haplotype validation, and the best match genotype was selected.

TGS accuracy verification

To validate the accuracy of our findings, we randomly selected 60 samples from the 268 analysed in this study and cross-referenced the TGS results with Luminex assay results for HLA systems (HLA-A, HLA-B and HLA-C) and TaqMan assay results for six HPA systems (HPA-1 to HPA-5 and HPA-15).

We performed HLA typing using the Luminex FLEXMAP 3D system (Luminex Corporation, Austin, TX, USA) and the LABType® SSO Typing Test (One Lambda, Inc., Canoga Park, CA, USA). HPA typing was conducted using the Applied Biosystems® 7500 Fast Real-Time PCR System (Thermo Fisher Scientific, Waltham, MA, USA) and the HPA (1–29) Antigen Typing Kit (Weihe Biotechnology Co., Ltd., Jiangsu, China).

RESULTS

Distribution of HPA system

In this study, we utilized targeted TGS to identify genetic variations within the HPA and HLA systems. All data exhibited a read quality greater than 0.99, with read numbers consistently exceeding 100. To validate the accuracy of our findings, we cross-referenced TGS results with TaqMan assay results for six HPA systems (HPA-1 to HPA-5 and HPA-15) from a cohort of 60 samples, observing a 100% concordance rate. Polymorphisms were identified in the HPA systems HPA-1 to HPA-6w, HPA-15w and HPA-21w, while the remaining 27 HPA systems only exhibited the 'aa' homozygous form. The allele frequencies of HPA-1b to HPA-6b, HPA-15b and HPA-21b were 0.0187, 0.0709, 0.4086, 0.0075, 0.0149, 0.0317, 0.4310 and 0.0019, respectively, and all SNV distributions conformed to Hardy–Weinberg equilibrium ($p > 0.05$). The genotypes and allele frequencies of the systems exhibiting polymorphisms are summarized in Table 1. Additionally, we analysed SNVs within other exonic regions of HPA genes. A total of 49 SNVs were identified in the exon sequencing regions of the *ITGB3*, *ITGA2B*, *ITGA2*, *CD109*, *GP1BA* and *GP9* genes (Table S2). Notably, two of these SNVs were novel and not existed in the dbSNV database (The GenBank IDs of the two novel SNVs: 2854037 and 2,853,304). Specifically, c.689C>T of the *ITGA2B* gene and c.134C>T of the *GP9* gene could result in amino acid changes (Table S2).

Genetic diversity patterns within HPA system gene

Despite extensive research into the polymorphism of HPA loci across various populations, the complete sequence patterns of each HPA-related gene have not been elucidated using haplotype sequencing. In this study, we initiated an investigation focused on the sequence patterns of HPA-related alleles within *ITGA2*. Our data indicate no linkage between HPA-1b and HPA-4b, as well as between HPA-6b and HPA-21wb. Intriguingly, we observed that the c.166-1029C>T variant was present exclusively in samples carrying the HPA-1b allele, suggesting it could serve as an auxiliary marker for HPA-1b screening (Table 2). A similar correlation was found between c.1959C>T and the HPA-21wa allele (Table 2). Besides, HPA-5b exhibits a distinct pattern

TABLE 1 HPA genotypes and allele frequencies in Chinese platelet blood donors ($n = 268$).

	Genotype number			Allele frequency		Hardy–Weinberg equilibrium test	
HPA system	Aa	Ab	Bb	a	b	χ^2	*p*-Value
HPA1	258	10	0	98.13%	1.87%	0.096	>0.05
HPA2	233	32	3	92.91%	7.09%	2.349	>0.05
HPA3	98	121	49	59.14%	40.86%	1.159	>0.05
HPA4	264	4	0	99.25%	0.75%	0.015	>0.05
HPA5	260	8	0	98.51%	1.49%	0.061	>0.05
HPA6	251	17	0	96.83%	3.17%	0.287	>0.05
HPA15	84	137	47	56.90%	43.10%	0.478	>0.05
HPA21	267	1	0	99.81%	0.19%	0.001	>0.05

Abbreviation: HPA, human platelet antigen.

TABLE 2 High incidence SNV with HPA systems.

HPA	SNV	Number of the individuals (A/N)
HPA-1b	c.166-1209C>T	10/10
HPA-6wa	c.1533A>G, c.1545G>A	139/519
HPA-21wa	c.1959C>T	534/535
HPA-5b	c.1602 + 344G>T, c.1602 + 548 A>G, c.1603-808A>G, c.2235 + 790 A>C, c.2236–219 T>C, c.2236–102 A>T, c.2236–62 T>C, c.2236–53 T>C, c.2347-77A>C, c.2429 + 21 T>C, c.2376 C>T, c.2484 G>A	8/8
HPA-3b	c.2187 + 34 delCAGGGGCTC, c.2187 + 260G>A, c.2188-7C>G, c. 2267 + 107A>G, c.2267 + 134G>A	216/219
HPA-15a	c.1903-288C>T; c.1903-275delGA; c.1903-218G>T; c.1903-146G>A; c.1903-119 T>A; c.1963 + 96 T>C; c.1963 + 343delATATTTTTGAA; c.1964-138C>T; c.2105 + 140G>T; c.2223 + 69delT; c.2223 + 205C>T; c.2223 + 376C>G; c.2223 + 473delT; c.2223 + 547A>G; c.2224-710C>T; c.2337 + 57A>G; c.2337 + 579insACTCCTGGCC; c.2337 + 852C>>T; c.2338-542 T>G; c.2390A>G; c.2533G>A; c.2556 + 68insT; c.2556 + 289C>T; c.2557-240G>A	243/305
HPA-15a	c.1963 + 584A>G; c.2223 + 361C>G; c.2337 + 649A>G	62/305
HPA-2b	39 bp repeats region(3–4 repeats)	31/38

A/N represents the detected mutations that occur *A* times in *N* HPA allele haplotypes.
Abbreviations: HPA, human platelet antigen; SNV, single nucleotide variant.

characterized by variants c.2568G>A, c.2376C>T and c.2484G>A, along with several intronic variants (Table 2).

For the *ITGA2B* gene, the HPA-3b allele was observed to have a strong association with a series of intronic mutations, including c.2187 + 34_2187 + 42*del*CAGGGGCTC, c.2187 + 260G>A, c.2188-7C>G, c.2267 + 107A>G and c.2267 + 134G>A (Table 2). Within the *CD109* gene, the HPA-15 allele displays two distinct patterns. The first pattern is characterized by a high incidence of the c.2533G>A variant and a series of intronic mutations associated with one HPA-15a allele. The second pattern is defined by the presence of the c.1963 + 585A>G, c.2229 + 361C>G and c.2337 + 649A>G variants (Table 2). Previous studies have reported the presence of one to four 39-base pair (bp) tandem repeats in the macroglycopeptide region of GPIbα [19, 20]. Our findings suggest that the HPA-2b allele is frequently associated with three to four repeats of this 39-bp sequence (Table 2), while HPA-2a samples contained 1–2 repeats of the 39-bp sequence.

Distribution for HLA class I loci

Our study comprehensively examined the full-length sequences of the HLA-A, HLA-B and HLA-C genes in 268 donors. To validate the performance of targeted TGS, we randomly selected 60 samples from these donors and performed HLA typing using the LABType® SSO Typing Test. The results showed 100% concordance between the two methods at both the first and second field levels of HLA typing. This study also identified previously unreported full-length genomic sequences for the alleles A33:03, B51:39, C01:85 and C15:26, which were not present in the current IPD-IMGT database. Our data indicate that A02:01 is the most prevalent HLA-A subtype, with a frequency of 21.22%. For HLA-B, the most common alleles are B51:01 and B46:01, each with a frequency of 15.30%. At the HLA-C locus, C06:02 is the most frequent allele, with a frequency of 16.39%. Detailed distributions of polymorphisms for HLA-A, HLA-B and HLA-C loci are provided in Table 3.

TABLE 3 The number and frequency of the HLA-A, HLA-B and HLA-C alleles.

HLA-A	Number	Frequency	HLA-B	Number	Frequency	HLA-C	Number	Frequency
A*02:01	101	18.84%	B*51:01	41	7.65%	C*06:02	78	14.55%
A*24:02	76	14.18%	B*46:01	41	7.65%	C*01:02	66	12.31%
A*11:01	68	12.69%	B*13:02	40	7.46%	C*07:02	46	8.58%
A*02:06	36	6.72%	B*15:01	34	6.34%	C*08:01	45	8.40%
A*30:01	32	5.97%	B*54:01	29	5.41%	C*04:01	41	7.65%
A*02:07	32	5.97%	B*40:01	29	5.41%	C*03:03	40	7.46%
A*03:01	31	5.78%	B*52:01	24	4.48%	C*03:04	39	7.28%
A*31:01	27	5.04%	B*48:01	24	4.48%	C*15:02	32	5.97%
A*01:01	27	5.04%	B*40:06	21	3.92%	C*14:02	27	5.04%
A*26:01	20	3.73%	B*15:11	16	2.99%	C*12:02	24	4.48%
A*33:03	14	2.61%	B*13:01	16	2.99%	C*08:03	12	2.24%
A*68:01	8	1.49%	B*50:01	15	2.80%	C*05:01	12	2.24%
A*23:01	8	1.49%	B*35:01	15	2.80%	C*01:03	11	2.05%
A*02:05	8	1.49%	B*44:03	14	2.61%	C*03:02	10	1.87%
A*02:03	8	1.49%	B*44:02	14	2.61%	C*14:03	6	1.12%
A*32:01	7	1.31%	B*37:01	13	2.43%	C*08:22	6	1.12%
A*02:10	5	0.93%	B*07:02	13	2.43%	C*07:04	6	1.12%
A*29:01	4	0.75%	B*15:18	11	2.05%	C*15:05	5	0.93%
A*24:20	4	0.75%	B*58:01	10	1.87%	C*12:03	5	0.93%
A*03:02	4	0.75%	B*38:02	9	1.68%	C*07:01	5	0.93%
A*33:01	3	0.56%	B*51:02	8	1.49%	C*07:06	4	0.75%
A*11:02	3	0.56%	B*40:02	8	1.49%	C*08:02	3	0.56%
A*24:07	2	0.37%	B*08:01	8	1.49%	C*04:03	3	0.56%
A*68:01	1	0.19%	B*57:01	6	1.12%	C*07:01	2	0.37%
A*34:01	1	0.19%	B*55:02	6	1.12%	C*04:82	2	0.37%
A*33:03	1	0.19%	B*39:01	5	0.93%	C*17:01	1	0.19%
A*30:18	1	0.19%	B*15:27	5	0.93%	C*16:02	1	0.19%
A*30:04	1	0.19%	B*07:05	5	0.93%	C*15:26	1	0.19%
A*25:01	1	0.19%	B*44:03	4	0.75%	C*15:04	1	0.19%
A*02:53 N	1	0.19%	B*35:03	4	0.75%	C*02:02	1	0.19%
A*02:48	1	0.19%	B*27:04	4	0.75%	C*01:85	1	0.19%
			B*18:01	4	0.75%			
			B*15:02	4	0.75%			
			B*49:01	3	0.56%			
			B*35:02	3	0.56%			
			B*15:07	3	0.56%			
			B*14:02	3	0.56%			
			B*67:01	2	0.37%			
			B*45:01	2	0.37%			
			B*40:03	2	0.37%			
			B*38:01	2	0.37%			
			B*15:17	2	0.37%			
			B*81:02	1	0.19%			
			B*56:03	1	0.19%			
			B*56:01	1	0.19%			
			B*55:01	1	0.19%			

TABLE 3 (Continued)

HLA-A	Number	Frequency	HLA-B	Number	Frequency	HLA-C	Number	Frequency
			B*51:39	1	0.19%			
			B*51:07	1	0.19%			
			B*41:01	1	0.19%			
			B*39:01	1	0.19%			
			B*35:08	1	0.19%			
			B*27:05	1	0.19%			
			B*15:58	1	0.19%			
			B*15:25	1	0.19%			
			B*15:21	1	0.19%			
			B*15:05	1	0.19%			

Abbreviation: HLA, human leucocyte antigen.

DISCUSSION

PTR frequently occurs in patients who receive multiple platelet transfusions, with a prevalence rate of approximately 30%–50% [6]. PTR can be classified into two types based on etiological factors: immune and non-immune. Immune-mediated PTR comprises about 20% of all PTR cases and is predominantly driven by alloantibodies against HLA class I and HPA. Of course, CD36 antibodies should not be overlooked either [21, 22]. Presently, therapeutic interventions for managing PTR associated with alloimmunization are limited. Evidence-based strategies involving HLA and/or HPA-matched or selected components and cross-matched platelets are promising for addressing this issue [6, 21, 22]. Therefore, the establishment of the platelet donor bank is essential for providing clinically matched platelets.

In this study, we established a long-read sequencing method based on the PacBio platform for HLA class I and HPA typing. We acquired exon SNVs in the targeted genes, uncovering several previously unreported SNVs. Additionally, we sequenced the length haplotypes of *GP9*, *GP1BB* and *GP1BA* for the first time. Given that genes associated with HPA, such as *CD109*, *ITGA2B*, *ITGA2* and *ITGB3*, have lengths exceeding 10 kbp, focusing on core regions for sequencing is deemed more cost-effective.

The polymorphism of HPAs is extensively documented across various populations. Hong et al. [23] concluded that HPA-1 to HPA-6w, HPA-15 and HPA-21w are among the most highly polymorphic HPAs in Zhejiang platelet donors in China, consistent with our study. Although the HPA system recently expanded to HPA-35 [12], less literature reports the polymorphism of HPA-35 in the Chinese population, which limits the understanding of the role of HPA-35 on platelet donor bank establishment. Meanwhile, our analysis of 268 individuals revealed that all donors had the HPA-35 'aa' homozygous form, suggesting these low-polymorphic HPAs could be economically excluded from the platelet donor bank. Additionally, we analysed gene patterns for low-frequency HPA alleles using haplotype sequencing, providing new insights into allele linkage within HPA-related genes. Alternatively, the TaqMan assay, a rapid and efficient HPA typing method that can be completed in 2 hours, is suitable for urgent scenarios [24]. However, the accuracy of TaqMan genotyping can be significantly compromised [24], particularly when the c.1533A>G and c.1545G>A mutations in HPA-6wa haplotypes are potentially misidentified as HPA-6wb, and the c.1959C>T mutation in HPA-21a haplotypes is potentially misclassified as HPA-21wb. Furthermore, we identified intronic SNVs that are strongly associated with HPA-1b, HPA-5b, HPA-3b and HPA-15a, which could serve as useful markers for future probe-based HPA typing methods.

Correspondingly, we investigated the frequencies of HLA class I alleles in Taiyuan (Northern China) within the Chinese Han population. The most prevalent HLA class I alleles observed were HLA-A02:01, B51:01 and C06:02. We compared our results with the HLA distribution in the northern population, and the HLA-A02:01 is the predominant allele in the northern China population [25]. In contrast, A11:01 and B46:01 are the most significant class I HLAs among the population in Zhejiang province (Southern China) [8]. This supported a regional difference in HLA distribution between the northern and southern Chinese populations. From a methodological perspective, common errors in HLA genotyping can arise due to PCR amplification bias and cross-contamination between samples, resulting in the misclassification of heterozygotes as homozygotes or vice versa [26, 27]. Such inaccuracies can lead to incorrect interpretations of a patient's HLA type, potentially affecting treatment decisions, transplantation outcomes, as well as other clinical considerations. Full-length HLA gene sequencing and intronic SNV analysis are capable of confirming that a homozygous genotype comprises two distinct haplotypes. For instance, in our study, two homozygous samples containing C07:02 or C*06:02 were validated by the presence of intronic mutations c.73 + 95 T>A or c.619-115G>C, respectively, supporting the reliability of the genotyping results.

In conclusion, we have successfully established a targeted long-read sequencing method based on amplicon technology, which enables the simultaneous genotyping of all 35 HPA systems as well as HLA-A, HLA-B and HLA-C loci. This method possesses the characteristics of high throughput, simultaneity and haplotyping capability. We reported the polymorphism of HPA systems and HLA class I and analysed the gene patterns and linkage of HPA-relative genes. Overall, we evaluated the feasibility of utilizing the Pacbio platform for HPA and HLA typing, particularly for platelet donor bank establishment.

ACKNOWLEDGEMENTS

P.Z. was responsible for the study design, donor collection and experimental conduction, data analysis and the draft manuscript writing; Q.L. optimized the study design and data processing, performed the draft manuscript revision and shared the first authorship; Y.X., Y.L., Y.W., Q.L., H.W., Y.Y., R.H. and W.F. assisted in the donor collection and experiment operation. D.Z. and Y.K. supervised this project, contributed to the proposal design, data analysis and manuscript writing and shared corresponding authorship.

The work was supported by Weigao Scientific Research Foundation of CSBT (CSBT-MWG-2023-01), Health Commission of Shanxi Province (Grant No. 2021XM13) and the foundation of The First Affiliated Hospital of Chengdu Medical College (Grant No. CYFY-GQ52).

CONFLICT OF INTEREST STATEMENT

The authors declare no conflicts of interest.

DATA AVAILABILITY STATEMENT

The datasets used and/or analysed during the current study are available from the corresponding author on reasonable request.

ORCID

Peizhe Zhao https://orcid.org/0000-0003-3041-7926
Qilu Lyu https://orcid.org/0000-0002-8424-973X
Yujie Kong https://orcid.org/0000-0002-5745-3575

REFERENCES

1. Koupenova M, Livada AC, Morrell CN. Platelet and megakaryocyte roles in innate and adaptive immunity. Circ Res. 2022;130:288–308.
2. Arinsburg SA, Shaz BH, Westhoff C, Cushing MM. Determination of human platelet antigen typing by molecular methods: importance in diagnosis and early treatment of neonatal alloimmune thrombocytopenia. Am J Hematol. 2012;87:525–8.
3. Curtis BR, Edwards JT, Hessner MJ, Klein JP, Aster RH. Blood group a and B antigens are strongly expressed on platelets of some individuals. Blood. 2000;96:1574–81.
4. Lyu Q, Peng M, Chen Q, Ji X, Wang Z, Li Q, et al. Frequency and molecular basis of CD36 deficiency among platelet donors in Kunming, China. Platelets. 2023;34:2176168.
5. Cohn CS. Platelet transfusion refractoriness: how do I diagnose and manage? Hematology Am Soc Hematol Educ Program. 2020;2020:527–32.
6. Panch SR, Guo L, Vassallo R. Platelet transfusion refractoriness due to HLA alloimmunization: evolving paradigms in mechanisms and management. Blood Rev. 2023;62:101135.
7. Rioux-Masse B, Cohn C, Lindgren B, Pulkrabek S, McCullough J. Utilization of cross-matched or HLA-matched platelets for patients refractory to platelet transfusion. Transfusion. 2014;54:3080–7.
8. Wang J, You X, He Y, Hong X, He J, Tao S, et al. Simultaneous genotyping for human platelet antigen systems and HLA-A and HLA-B loci by targeted next-generation sequencing. Front Immunol. 2022;13:945994.
9. Dendrou CA, Petersen J, Rossjohn J, Fugger L. HLA variation and disease. Nat Rev Immunol. 2018;18:325–39.
10. Xu DY, Sun CX, Yu L, He YL, Deng G, Zhang JW. Methodological study on the establishment of HLA/HPA Gene Bank of platelet donors and its clinical application. Indian J Hematol Blo. 2023;39:123–31.
11. Saito S, Ota S, Seshimo H, Yamazaki Y, Nomura S, Ito T, et al. Platelet transfusion refractoriness caused by a mismatch in HLA-C antigens. Transfusion. 2002;42:302–8.
12. Versiti. HPA database. 2020 Available from: https://www.versiti.org/hpa. Last accessed 1 Nov 2023.
13. Gerlach JA. Human lymphocyte antigen molecular typing - how to identify the 1250+alleles out there. Arch Pathol Lab Med. 2002;126:281–4.
14. Rodriguez R, Krishnan Y. The chemistry of next-generation sequencing. Nat Biotechnol. 2023;41:1709–15.
15. Klasberg S, Surendranath V, Lange V, Schofl G. Bioinformatics strategies, challenges, and opportunities for next generation sequencing-based HLA genotyping. Transfus Med Hemother. 2019;46:312–25.
16. Athanasopoulou K, Boti MA, Adamopoulos PG, Skourou PC, Scorilas A. Third-generation sequencing: the spearhead towards the radical transformation of modern genomics. Life (Basel). 2022;12:30.
17. Rhoads A, Au KF. PacBio sequencing and its applications. Genom Proteom Bioinf. 2015;13:278–89.
18. Wang Y, Zhao Y, Bollas A, Wang Y, Au KF. Nanopore sequencing technology, bioinformatics and applications. Nat Biotechnol. 2021;39:1348–65.
19. Ishida F, Furihata K, Ishida K, Yan J, Kitano K, Kiyosawa K, et al. The largest variant of platelet glycoprotein Ib alpha has four tandem repeats of 13 amino acids in the macroglycopeptide region and a genetic linkage with methionine145. Blood. 1995;86:1357–60.
20. Di Paola J, Jugessur A, Goldman T, Reiland J, Tallman D, Sayago C, et al. Platelet glycoprotein I(b)alpha and integrin alpha2 beta1 polymorphisms: gene frequencies and linkage disequilibrium in a population diversity panel. J Thromb Haemost. 2005;3:1511–21.
21. van Osch TLJ, Pongracz T, Geerdes DM, Mok JY, van Esch WJE, Voorberg J, et al. Altered Fc glycosylation of anti-HLA alloantibodies in hemato-oncological patients receiving platelet transfusions. J Thromb Haemost. 2022;20:3011–25.
22. Wu G, Zhou Y, Li L, Zhong Z, Li H, Li H, et al. Platelet immunology in China: research and clinical applications. Transfus Med Rev. 2017;31:118–25.
23. Hong X, Chen S, Ying Y, Liu Y, Xu X, He J, et al. Simultaneous genotyping of human platelet alloantigen-1 to 28bw systems by multiplex polymerase chain reaction sequence-based typing. Vox Sang. 2017;112:360–6.
24. Hui L, DelMonte T, Ranade K. Genotyping using the TaqMan assay. Curr Protoc Hum Genet. 2008;Chapter 2:Unit 2.10.
25. Cheng LH, Jin SZ, Gao SQ, Li Z, Zou HY, Wang DM, et al. Difference in HLA-A*02 allele distribution between Han populations in south and north China. Di Yi Jun Yi Da Xue Xue Bao. 2005;25:321–4.
26. Osoegawa K, Vayntrub TA, Wenda S, De Santis D, Barsakis K, Ivanova M, et al. Quality control project of NGS HLA genotyping for the 17th international HLA and Immunogenetics workshop. Hum Immunol. 2019;80:228–36.
27. Cargou M, Ralazamahaleo M, Blouin L, Top I, Elsermans V, Andreani M, et al. Evaluation of the AllType kit for HLA typing using the ion torrent S5 XL platform. HLA. 2020;95:30–9.

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

How to cite this article: Zhao P, Lyu Q, Xu Y, Liang Y, Wu Y, Li Q, et al. A novel approach to simultaneous genotyping of human platelet antigen systems and human leucocyte antigen class I loci using PacBio long-read sequencing. Vox Sang. 2024.