

RESEARCH ARTICLE

Full screening and accurate subtyping of HLA-A*02 alleles through group-specific amplification and mono-allelic sequencing

Shengli Song¹, Miaomiao Han¹, Han Zhang, Yuanxia Wang and Hong Jiang

HLA-A*02 is the most prevalent and polymorphic major histocompatibility complex (MHC) allele family in humans. Functional differences have been revealed among subtypes, demanding further subtyping of HLA-A*02 in basic and clinical settings. However, the fast growing polymorphisms render traditional primer- or probe-based typing methods impractical and result in increasing ambiguities in direct sequence-based typing. In this study, we combined group-specific amplification and mono-allelic sequencing to design and validate a simple scheme for the complete screening and accurate subtyping of all 540 reported HLA-A*02 alleles. This scheme could be performed in routine labs to facilitate studies with an interest in HLA-A*02.

Cellular & Molecular Immunology (2013) 10, 490–496; doi:10.1038/cmi.2013.33; published online 19 August 2013

Keywords: genotyping techniques; group-specific amplification; HLA-A*02 antigen; mono-allelic sequencing; sequence-based typing

INTRODUCTION

HLA-A*02 (A*02) is the most prevalent major histocompatibility complex (MHC) class I allele family in humans, presenting at high frequencies in all ethnic populations.¹ However, A*02 is the most polymorphic MHC allele family identified, with 540 alleles, and an additional allele A*02:415, recorded in the IMGT/HLA database (release 3.12.0).² Early studies have established that even a single amino-acid difference in the peptide-binding groove of the A*02 molecule might be sufficient to dramatically change the nature of bound peptides.^{3,4} A recent study of the crystal structures of three A*02 alleles revealed that minor residue changes between different alleles induce significant alterations in the MHC-peptide interface and introduce conformational changes in the p3–p8 peptide region.⁵ Structural and functional differences underlie the differential disease outcome in individuals of different A*02 subtypes,⁶ and there is a need for high-resolution A*02 subtyping of individuals for the design of personalized vaccines and the development of disease treatments.⁷

A number of DNA-based techniques have been used for HLA typing, and the most commonly used techniques include polymerase chain reaction with sequence-specific primers

(PCR-SSP) or sequence-specific oligonucleotide probes (PCR-SSOP) and direct sequence-based typing (SBT). The main advantages of PCR-SSP and PCR-SSOP methods are that these methods are fast, generating useful data for discriminating the alleles of a small family.⁸ However, the dependence on certain polymorphic sites for each allele or allele group identification confines these primer- or probe-based methods to low or intermediate resolutions when applied to studies of large allele families.⁹ In contrast, SBT technology detects the polymorphic regions from conserved priming sites through direct sequencing, facilitating both high throughput and high resolution, and presenting a powerful technique for new allele identification. However, the main drawback of SBT is the ambiguities resulting from the *cis/trans* assignment of base calls in heterozygous samples.⁹ One technical approach to the issue of ambiguity is to harness the power of next-generation sequencing methods;^{9,10} however, this option is an expensive choice for routine work.

In the case of A*02 subtyping, a number of methods have been established based on PCR-SSP^{11–13} or PCR-SSOP.^{14,15} However, considering the rapidly growing number of both A*02 and non-A*02 alleles in recent years, it is impractical to

Shanghai Institute of Immunology, Shanghai Jiao Tong University School of Medicine, Shanghai, China

¹Song and Han contributed equally to this study.

Correspondence: Dr H Jiang, Shanghai Institute of Immunology, Shanghai Jiao Tong University School of Medicine, 280 South Chongqing Road, Shanghai 200025, China.

E-mail: hongjiang24@gmail.com

Received: 20 May 2013; Revised: No; Accepted: 13 June 2013

set up a system with a large number of overlapping primers or probes and an even larger number of reactions or hybridizations to completely analyze each sample. The rapid growth in allele number reflects an even more rapid growth in the complexity of allele ambiguities in direct SBT.² To generate a simple and accurate method for subtyping A*02 alleles, we developed a strategy combining group-specific amplification and mono-allelic sequencing to analyze all reported congenetic human MHC class I alleles. After validating this technique using a panel of reference samples, we successfully applied this scheme in another study to screen individuals of particular A*02 subtypes.

MATERIALS AND METHODS

Bioinformatics analysis

Multiple alignment files (in MSF format) of genomic and coding DNA sequences of MHC class I loci (HLA-A, -B, -C, -E, -F, -G, -H, -J, -K and -L) were downloaded from the FTP directory of the IMGT/HLA database (<ftp://ftp.ebi.ac.uk/pub/databases/ipd/imgt/hla/>, release 3.12.0) and analyzed using ClustalX software (<http://www.clustal.org/download/>, version 2.1). A phylogenetic tree was generated using the program package PHYLIP (<http://evolution.genetics.washington.edu/phylip.html>, version 3.69), employing the neighbor-joining algorithm. The tree was viewed using FigTree software (<http://tree.bio.ed.ac.uk/software/figtree/>, version 1.4.0).

Ethics statement

Reference genomic DNA samples were used in accordance with the Terms and Conditions of Use Statement For Material Requested From The Fred Hutchinson Cancer Research Center IHWG Cell and Gene Bank. The volunteers were enrolled with written informed consent according to the protocol approved through the Ethics Committee of Shanghai Jiao Tong University School of Medicine.

Reference and population samples

A panel of reference genomic DNA samples of known HLA-A types was obtained from the Fred Hutchinson Cancer Research Center IHWG Cell and Gene Bank (Table 1). A random population of 126 healthy Chinese individuals in Shanghai was enrolled in this study. The genomic DNA was prepared from

the peripheral blood mononuclear cells of volunteers using a simple and reliable procedure (Supplementary Protocol 1). Briefly, the cells were resuspended at 5×10^3 cells/ μ l in 10 mM Tris-HCl (pH 8.0) containing 100 ng/ μ l proteinase K (Merck, Darmstadt, Germany) and heated at 56 °C for 1 h, followed by incubation at 95 °C for 10 min on a thermocycler.

Primers, reaction combinations and amplification conditions

Primers and related information are listed in Table 2. The references are indicated for primers that have been previously described. The reaction combinations and related information are detailed in Table 3. An internal control reaction, specific for a 333-bp region of the *B2M* gene, was included in each reaction in main- and sub-screening to verify that the PCR reaction functioned properly.

The 100- μ l PCR reaction mixture contained 10 mM Tris-HCl (pH 8.3), 50 mM KCl, 1.5 mM MgCl₂, 200 μ M of each dNTP, 0.2 μ M of each primer, 2.5 units *Taq* DNA Polymerase Hot Start Version (TaKaRa, Dalian, China) and 10 μ l template DNA. The template DNA was either 25 ng/ μ l purified genomic DNA or 5×10^3 cells/ μ l lysates for reactions using cycling protocol CP1 (below) or 1 : 200 diluted CP1 products for the CP2 (below) reactions. The reaction volumes were 10 μ l for screening and 50 μ l for sequencing or cloning.

The PCR reactions were performed on a TProfessional Thermocycler (Biometra, Göttingen, Germany). The following cycling protocols were used: CP1: 95 °C 5 min, 5 cycles of (95 °C 1 min, 68 °C 1 min, 72 °C t_E), 5 cycles of (95 °C 1 min, 63 °C 1 min, 72 °C t_E), 25 cycles of (95 °C 30 s, 58 °C 45 s, 72 °C t_E) and 72 °C 5 min; and CP2: 95 °C 2 min, 5 cycles of (95 °C 30 s, 70 °C 1 min, 72 °C t_E), 5 cycles of (95 °C 30 s, 65 °C 1 min, 72 °C t_E), 10 cycles of (95 °C 30 s, 60 °C 45 s, 72 °C t_E) and 72 °C 5 min; where t_E represents elongation time, as indicated with each reaction set listed in Table 3. The PCR products (5 μ l each sample) were analyzed on a 1% agarose gel stained with ethidium bromide.

Amplicon purification, sequencing and allele assignment

The PCR amplicons were purified using the Cycle-Pure Kit (Omega Bio-Tek, Norcross, GA, USA) for direct sequencing, or using a cloning-sequencing strategy, the purified amplicons

Table 1 Reference genomic DNA samples from the Fred Hutchinson Cancer Research Center IHWG Cell and Gene Bank

IHWG no.	Sample ID	HLA-A types in database	Subtyping or grouping results in this study
IHW09213	KLO	A*02:08/A*01:01	A*02:08/NA2
IHW09220	XLI-ND	A*02:10/A*30:01	A*02:10/NA2
IHW09377	FH5	A*02:01:01:01/A*29:02:01	A*02:01:01:01/NA2
IHW09112	CHA, AJ	A*03:01:01:01/A*24:03:01	NA2/NA2
IHW09211	IDF	A*69:01/A*66:01	NA2/NA2
IHW09363	GRC-187	A*68:01:02/A*31:01:02	NA2/NA2
IHW09064	AMALA	A*02:17:01	A*02:17:02
IHW09126	WATANABE	A*02:01:01/A*02:07	A*02:01:01:01/A*02:07:01
IHW09446	FH58	A*02:01:01:01/A*02:20:02	A*02:01:01:01/A*02:20:02

Abbreviation: IHWG, The International Histocompatibility Working Group.

Table 2 Primers used for screening and subtyping of HLA-A*02 alleles

Primer ^a	Sequences (5'–3')	Intron/exon	Nucleotide location ^b	Ref.
For main-screening				
78C-FW1 ^c	CTCGCCCCAGGCTCC	Intron 1, exon 2	(120–130), 74–78	12
78C-FW2 ^c	CCTCGTCCCCAGGCTCC	Intron 1, exon 2	(119–130), 74–78	
78T-FW	CCTCGTCCCCAGGCTCT	Intron 1, exon 2	(119–130), 74–78	19–22
A-I3-RV	GTCCCAATTGTCTCCCTC	Intron 3	(83–65)	24 ^d
77G-FW	TCCTCGTCCCCAGGCTG	Intron 1, exon 2	(118–130), 74–77	
538A-RV	GTGCTTGGTGGTCTGAGCT	Exon 3	556–538	
hB2M-FW	CCGATATTCCTCAGGTACTC			12 ^d
hB2M-RV	ACACAACCTTCAGCAGCTTAC			12 ^d
For sub-screening				
318A-FW	GACGGGGAGACACGGAAA	Exon 2	301–318	
555A-RV	GCCGCCTCCCACTTGT	Exon 3	570–555	
149C-FW	CCCGCTTCATCGCC	Exon 2	135–149	
440G-RV	CGGAGGAAGCGCCC	Exon 3	453–440	
412A-FW	GTTCTCACACCATCCAGATA	Exon 3	393–412	
555G-RV	GCCGCCTCCCACTTGC	Exon 3	570–555	
For sequencing or cloning				
A-I1-FW	CCTCTGYGGGGAGAAGCAA	Intron 1	(28–46)	24 ^d
A-I4-RV	CAGAGAGGCTCCTGCTTTC	Intron 4	(54–36)	
Seq-I2-RV	TCGGACCCGGAGACTGTG	Intron 2	(81–64)	
Seq-I2-FW	GTTTCATTTTCAGTTTAGGCCA	Intron 2	(148–169)	
Seq-I3-FW	GGTGTCTGTCCATTCTC	Intron 3	(507–524)	
For ambiguity resolving				
A-F5-FW	GATTCCCCAACTCCGCAG	5'-flanking region	(–198–(–181))	
78T-RV	GAAGAAATACCTCATGGAGTGA	Exon 2	99–78	
Seq-I1-RV	CTCATGGAGTGAGAGCCTGG	Exon 2, intron 1	89–74, (130–127)	

^a Group- or allele-specific primers are named (polymorphic position and nucleotide)-FW/RV, locus-specific primers are named (locus)-(intron/exon)-FW/RV and sequencing primers are referred to as Seq-(Intron/Exon)-FW/RV.

^b The exon nucleotides are numbered according to the alignment of coding DNA sequences of HLA-A alleles in the IMGT/HLA database.² The 5'-flanking region and intron passages are shown in the brackets and numbered according to Kotsch *et al.*²⁵

^c 78C-FW1 and 78C-FW2 were synthesized separately and mixed at a 1 : 1 mole ratio before use and designated as 78C-FW.

^d With 5' nucleotide truncations from or additions to previously reported sequences.

Table 3 Reaction combinations for the screening and subtyping of HLA-A*02 alleles

Reaction	Primer combination ^a	Target alleles ^b	Cycling protocol ^c	Amplicon size (bp)
For main-screening				
78C/I3	78C-FW, A-I3-RV	GNA2 group	CP1(1 min)	881 or 882
78T/I3	78T-FW, A-I3-RV	GA2 group		882
77G/538A	77G-FW, 538A-RV	A*02:236		688
Control	hB2M-FW, hB2M-RV	(Human <i>B2M</i> gene)		333
For sub-screening				
318A/555A	318A-FW, 555A-RV	SA2 group	CP2(30 s)	510
149C/440G	149C-FW, 440G-RV	Partial SNA2 alleles ^d		516
412A/555G	412A-FW, 555G-RV	Another partial SNA2 alleles ^e		178
For sequencing or cloning				
78C/I4	78C-FW, A-I4-RV	GNA2 group	CP1(2 min)	1728 or 1729
78T/I4	78T-FW, A-I4-RV	GA2 group		1729
I1/I4	A-I1-FW, A-I4-RV	HLA-A		1820
For ambiguity resolving				
F5/78T	A-F5-FW, 78T-RV	Partial GA2 alleles	CP1(30 s)	427

^a Primer information is listed in Table 2.

^b Group names are specified in Figure 1 and Table 4.

^c Expressed as Protocol(Elongation time); the cycling protocols are detailed in the section on 'Materials and methods'.

^d Alleles A*01:01:13, A*03:01:12, A*11:01:34, A*24:02:58, A*25:01:07, A*26:20 and A*68:03:02.

^e Alleles A*01:01:13, A*03:01:12, A*03:89, A*11:01:34 and A*30:08.

were cloned into a pMD19-T vector (TaKaRa, Dalian, China), and the colony PCR products of two separate clones and of the original amplicons were subjected to sequencing. The allele assignment was accomplished through a nucleotide BLAST search in the IMGT/HLA database² using the combined sequences of exons 2–3 or 2–4 from the sequencing results.

RESULTS

Bioinformatics analysis and scheme design: mono-allelic sequencing based on group-specific amplification

To design a scheme for subtyping A*02 alleles, we first analyzed the nucleotide sequences of the 540 reported A*02 alleles. A multiple alignment of the sequences of exons 2 and 3 was performed using ClustalX, and a phylogenetic tree was generated (Supplementary Figure 1). We observed 197 parallel branches originating from the common allele group A*02:01:01G, implying the independent evolution of these latter alleles/groups. As it is impractical to determine numerous alleles/groups using conventional primer- or probe-based methods, such as PCR-SSP and PCR-SSOP, we employed another commonly used technology, SBT. However, when direct SBT is employed to analyze heterozygous samples, a large number of ambiguous allele combinations result, reflecting the similarities among A*02 alleles. Indeed, there are 17 235 ambiguous combinations containing A*02 alleles.²

Because the major proportion of A*02⁺ individuals are heterozygous for A*02 alleles,¹⁶ we employed mono-allelic sequencing followed by the specific-amplification of A*02 alleles to overcome the ambiguity in direct SBT. The sequences of all

540 A*02 alleles were analyzed in the context of all alleles reported in both HLA-A and other congenetic MHC class I loci. Intriguingly, a single nucleotide alteration (from C to T at position 78), which distinguishes most A*02 alleles from others, was observed at the 5'-end of exon 2, and this feature is relatively conserved among all class I alleles analyzed (Table 4). This feature was explored as a general rule for discrimination between A*02 and non-A*02 alleles through group-specific amplification using reactions 78C/I3 and 78T/I3 (Table 3). For further discrimination of specific alleles outside the general grouping rule, group- or allele-specific reactions 77G/538A, 318A/555A, 149C/440G and 412A/555G (Table 3) were designed based on sequence analysis. Thus, we developed a simple scheme for the complete screening of A*02 alleles through group-specific amplification in haploid genomes (Figure 1).

Next, we applied this screening scheme to diploid genomes, and subtyping strategies were developed for each case (Table 5). The following findings are noteworthy:

1. Sequencing with primers Seq-I2-RV and Seq-I2-FW identified the sequences associated with exons 2 and 3, respectively. When required, exon 4 was sequenced using the primer Seq-I3-FW. For the five alleles of the two groups (group (A*02:01:01:01, A*02:01:01:02L, A*02:01:01:03) and group (A*02:17:01, A*02:17:02)), differences between group members existed only in the 5'-flanking region, exon 1 or intron 1, and additional sequencing of the amplicons of reaction F5/78T (Table 3) using Seq-I1-RV is required to resolve this ambiguity.

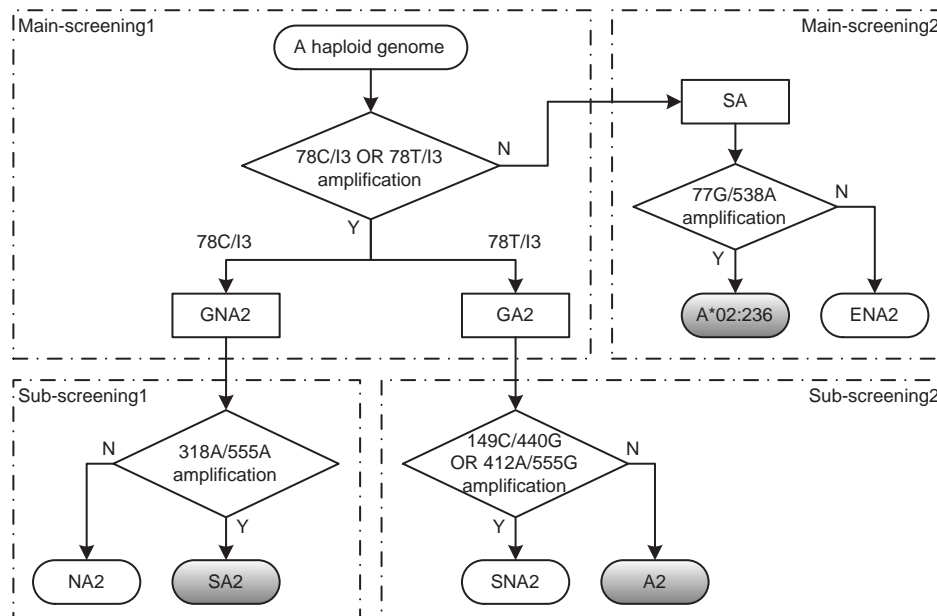


Figure 1 The scheme for full screening of HLA-A*02 alleles from haploid genomes using group-specific amplification. In Main-screening1, the reactions 78C/I3 and 78T/I3 amplify HLA-A alleles of General Non-A*02 (GNA2) and General A*02 (GA2) groups, respectively; the remaining HLA-A alleles are grouped as Special HLA-A (SA), in which reaction 77G/538A discriminates A*02:236 from Exceptional Non-A*02 (ENA2) alleles through Main-screening2. Reaction 318A/555A discriminates Special A*02 (SA2) alleles from Non-A*02 (NA2) alleles in GNA2 group through Sub-screening1, and either 149C/440G or 412A/555G or both reactions indicate the presence of Special Non-A*02 (SNA2) other than A*02 (A2) alleles in the GA2 group via Sub-screening2. The reactions are detailed in Table 3.

Table 4 MHC class I alleles sorted according to the nucleotide sequences at the 5'-end of exon 2

MHC loci		Number of alleles sorted according to the nucleotide sequences at positions 74–78 ^a			
		GCTCC	GCTCT	GCTG(T/C)	Other sequences
HLA-A	Non-A*02	1686 ^{b,c}	9 ^{e,f}	2 ^{h,i}	4 ^{h,i}
	A*02	6 ^{b,d}	533 ^{e,g}	1 ^{h,i}	0
HLA-B		2925	0	0	5
HLA-C		1781	0	0	4
HLA-E, -F, -G, -H, -J, -K and -L		108	0	0	7

According to published sequences in the IMGT/HLA database (release 3.12.0).²

^a Numbered according to the alignment of coding DNA sequences of HLA-A alleles in the IMGT/HLA database.²

^b General Non-A*02 (GNA2) alleles.

^c Non-A*02 (NA2) alleles.

^d Special A*02 (SA2) alleles.

^e General A*02 (GA2) alleles.

^f Special Non-A*02 (SNA2) alleles.

^g A*02 (A2) alleles.

^h Special HLA-A (SA) alleles.

ⁱ Exceptional Non-A*02 (ENA2) alleles.

^j Allele A*02:236.

- In case 3, the cloning-sequencing strategy is currently the most practical method to obtain accurate results, as there are still 1588 ambiguous combinations within these 533 A2 alleles;² for the rare case 2.2, we prefer the cloning-sequencing strategy, instead of setting up a PCR-SSP system, for the identification of each SA2 allele.
- In cases 3 and 2.2, amplicons from the diploid template and the colony PCR products of two separate clones were sequenced to reveal sequences on both haplotypes, facilitating corrections of possible PCR errors in some clones.
- Table 5 represents a full scheme, which can be simplified according to the purpose of a given study. For example, in a common case where individuals of certain subtypes in GA2 group are screened and subtyped, one might screen for samples of cases 1 and 3 and treat case 1 as case 1.1 directly, considering that alleles of SA2 and SNA2 groups are rare.

Validation of the scheme with reference samples

This scheme was validated through the screening and subtyping of a panel of reference samples of known HLA-A types obtained from the Fred Hutchinson Cancer Research Center IHWG Cell and Gene Bank. The screening/grouping results of the group-specific amplification of all samples were consistent with the HLA-A genotypes indicated in the database (Table 1). Representative results are shown in Figure 2. The subtyping results of the mono-allelic sequencing of all A*02⁺ samples were identical (or of higher resolutions) to the known identified A*02 subtypes, as indicated in the database. However, one exception, IHW09064, was detected in two separate experiments to be A*02:17:02 (Table 1) other than A*02:17:01 as reported.¹⁷ These two alleles differ from each other in only one nucleotide in exon 1.¹⁸ We searched for reports on the A*02:17:01 allele and found no

additional confirmation of this polymorphic position, except for the initial report identifying this allele in the cell line AMALA (i.e., IHW09064).¹⁷ Therefore, we consider that further confirmation of this allele in A*02:17:01⁺ reference cell lines is needed.

Due to the lack of reference samples for A*02:236, SA2 or SNA2 subtypes, we analyzed each of these alleles/groups based on the parental groups defined in this study, and we generated a panel of substitution templates from the reference samples (Supplementary Table 1 and Supplementary Figure 2a). These substitution templates consist of specific templates and most alike templates. The specific templates are of identical sequences at both priming sites to that of corresponding primers, respectively. Whereas most alike templates are of mismatched but most alike sequences at either priming site to that of the corresponding primer. The specificities of these primer sets were subsequently validated using these substitution templates (Supplementary Figure 2b).

Screening and subtyping of the HLA-A*02 alleles in a random population in Shanghai

We subsequently applied this scheme to the molecular screening and subtyping of A*02 alleles in a random Chinese population in Shanghai. The scheme yielded accurate subtyping results for each sample tested. Among the 126 individuals involved, 65 (51.6%) individuals were heterozygous and 11 (8.7%) individuals were homozygous for A*02. Among these 87 A*02 alleles, the top three frequent subtypes were A*02:01:01:01 (44.8%), A*02:07:01 (25.3%) and A*02:06:01 (16.1%).

DISCUSSION

In this study, we designed and validated a simple scheme for the complete screening and accurate subtyping of A*02

Table 5 HLA-A*02 subtyping strategies based on group-specific amplification

Case	Reaction patterns ^a		Grouping ^b	Subsequent treatments ^c
Main-screening1(gDNA) ^c				
	78C/I3	78T/I3		
Case 1	+	+	GNA2/GA2	(Below)
Case 2	+	–	GNA2/GNA2, GNA2/SA	(Below)
Case 3	–	+	GA2/GA2, GA2/SA	Cloning-sequencing(I1/I4, (Seq-I2-RV, Seq-I2-FW, Seq-I3-FW))
Case 4	–	–	SA/SA	(Below)
For case 1: Sub-screening1(78C/I3) and Sub-screening2(78T/I3) ^c				
	318A/555A	149C/440G OR 412A/555G		
Case 1.1	–	–	NA2/A2	PCR(78T/I4), Sequencing(78T/I4, (Seq-I2-RV, Seq-I2-FW, Seq-I3-FW))
Case 1.2	–	+	NA2/SNA2	
Case 1.3	+	–	SA2/A2	Combining treatments for cases 1.1 and 1.4
Case 1.4	+	+	SA2/SNA2	PCR(78C/I4), Sequencing(78C/I4, (Seq-I2-RV, Seq-I2-FW))
For case 2: Main-screening2(gDNA) and Sub-screening1(78C/I3) ^c				
	77G/538A	318A/555A		
Case 2.1	–	–	NA2/NA2, NA2/ENA2	
Case 2.2	–	+	NA2/SA2, SA2/SA2, SA2/ENA2	Cloning-sequencing(I1/I4, (Seq-I2-RV, Seq-I2-FW))
Case 2.3	+	–	A*02:236/NA2	
Case 2.4	+	+	A*02:236/SA2	PCR(78C/I4), Sequencing(78C/I4, (Seq-I2-RV, Seq-I2-FW))
For case 4: Main-screening2(gDNA) ^c				
	77G/538A			
Case 4.1	–		ENA2/ENA2	
Case 4.2	+		ENA2/A*02:236, A*02:236/ A*02:236	PCR(I1/I4), Sequencing(I1/I4, Seq- I2-RV)

Abbreviation: gDNA, genomic DNA.

^a Reactions are detailed in Table 3.

^b Group names are specified in Figure 1 and Table 4.

^c Expressed as Function(Parameter), e.g., Sub-screening1(78C/I3) means Sub-screening1 operation with amplicons of reaction 78C/I3 as templates; PCR(I1/I4) means PCR operation according to reaction I1/I4; Sequencing(I1/I4, Seq-I2-RV) means sequencing of amplicons of reaction I1/I4 with primer Seq-I2-RV; and Cloning-sequencing(I1/I4, (Seq-I2-RV, Seq-I2-FW, Seq-I3-FW)) means cloning of amplicons of reaction I1/I4 and sequencing of colony PCR products of two separate clones and the original amplicons of reaction I1/I4 with primers Seq-I2-RV, Seq-I2-FW and Seq-I3-FW.

alleles. This scheme was based on the observation that a single nucleotide alteration near the 5'-end of exon 2 distinguishes most A*02 alleles from other MHC class I alleles. The group-specific amplification of 'general' A*02 or non-A*02 HLA-A alleles was achieved through specific priming at this polymorphic residue and a locus-specific site using an amplification refractory mutation system. Special alleles outside the general grouping rule were further discriminated through additional group- or allele-specific amplifications based on sequence analysis to completely screen all of the A*02 alleles reported to date.

We noted that the polymorphic residue mentioned has been widely explored for the specific amplification of A*02 alleles;^{11,12,19–22} however, only 14–37 A*02 alleles were identified at those times, and the number of non-A*02 alleles was small. In this study, through the complete analysis of all 7071 reported congenetic MHC class I alleles, we validated this feature as a general rule for A*02 and non-A*02 discrimination. For some special alleles reported, a small number of

additional reactions were designed as patches for the general rule. We predict that as more alleles are identified, only some modifications in these patches will be needed to maintain the relevance of the current scheme.

Importantly, the differential amplification of A*02 and non-A*02 HLA-A alleles discriminates A*02 heterozygous individuals from homozygous individuals in diploid genomic samples and facilitates accurate subtyping using mono-allelic sequencing in heterozygous samples. Because most A*02⁺ individuals are heterozygous for the A*02 allele, this scheme is highly efficient in the majority of cases. For A*02 homozygous samples, we suggest that the cloning-sequencing strategy is currently the most practical method for accurate subtyping. Compared with conventional procedures,²³ a minor modification was made in our cloning-sequencing strategy, sequencing colony PCR products, instead of amplicons or plasmid preparations from liquid cultures, which only required one more day than direct mono-allelic sequencing with heterozygous samples.

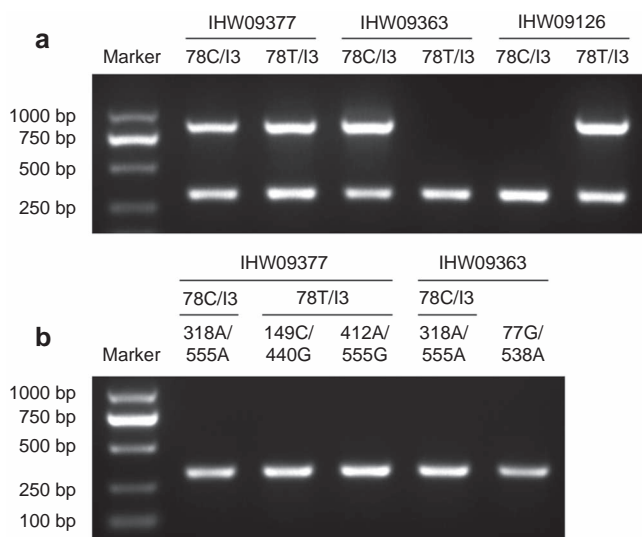


Figure 2 Representative results of screening reactions with reference samples. Reference genomic DNA samples IHW09377 (A*02 heterozygous), IHW09363 (non-A*02) and IHW09126 (A*02 homozygous) served as templates in Main-screening1 (a) and in subsequent reactions according to each case (b). The PCR products were analyzed on a 1% agarose gel stained with ethidium bromide. The analyzed amplicons are indicated as Template/Reaction above each lane. The reactions are detailed in Table 3, and the treatment strategies are outlined in Table 5. Representative results of three independent experiments.

DNA sequencing has currently become a routine technique that is available and affordable in-house or from technical service companies. Therefore, the sequencing-based scheme designed in this study provides a feasible method for accurate A*02 subtyping in routine labs lacking sophisticated platforms, such as in HLA labs. Considering the prevalence and significance of A*02 alleles in the design of vaccines and in the development of disease treatments as well as the structural and functional differences between different subtypes, this simple subtyping scheme might be valuable for facilitating basic and clinical studies with interests in HLA-A*02.

ACKNOWLEDGEMENTS

The authors would like to thank Jueqin Yang for assistance with sample preparation. The authors would also like to thank the Fred Hutchinson Cancer Research Center IHWG Cell and Gene Bank for providing reference genomic DNA samples. This work was supported through grants from the National Natural Science Foundation of China (NSF-30830093) and the National Key Program (973) for Basic Research of China (2009CB522409) to HJ.

Supplementary Information accompanies the paper on *Cellular & Molecular Immunology* website.

- 1 Imanishi T, Akaza T, Kimura A, Tokunaga K, Gojobori T. Allele and haplotype frequencies for HLA and complement loci in various ethnic groups. In: Tsuji K, Aizawa M, Sasazuki T (ed.) *HLA 1991: Proceedings of the Eleventh International Histocompatibility Workshop and Conference*. Oxford: Oxford University Press, 1992: 1065–1220.
- 2 Robinson J, Halliwell JA, McWilliam H, Lopez R, Parham P, Marsh SG. The IMGT/HLA database. *Nucleic Acids Res* 2013; **41**: D1222–D1227.

- 3 Barouch D, Friede T, Stevanovic S, Tussey L, Smith K, Rowland-Jones S *et al*. HLA-A2 subtypes are functionally distinct in peptide binding and presentation. *J Exp Med* 1995; **182**: 1847–1856.
- 4 Sudo T, Kamikawaji N, Kimura A, Date Y, Savoie CJ, Nakashima H *et al*. Differences in MHC class I self peptide repertoires among HLA-A2 subtypes. *J Immunol* 1995; **155**: 4749–4756.
- 5 Liu J, Chen KY, Ren EC. Structural insights into the binding of hepatitis B virus core peptide to HLA-A2 alleles: towards designing better vaccines. *Eur J Immunol* 2011; **41**: 2097–2106.
- 6 Vejbaesya S, Eiermann TH, Suthipinittharm P, Banacha C, Stephens HA, Luangtrakool K *et al*. Serological and molecular analysis of HLA class I and II alleles in Thai patients with psoriasis vulgaris. *Tissue Antigens* 1998; **52**: 389–392.
- 7 Chen KY, Liu J, Ren EC. Structural and functional distinctiveness of HLA-A2 allelic variants. *Immunol Res* 2012; **53**: 182–190.
- 8 Dunckley H. HLA typing by SSO and SSP methods. *Methods Mol Biol* 2012; **882**: 9–25.
- 9 Erlich H. HLA DNA typing: past, present, and future. *Tissue Antigens* 2012; **80**: 1–11.
- 10 Erlich RL, Jia X, Anderson S, Banks E, Gao X, Carrington M *et al*. Next-generation sequencing for HLA typing of class I loci. *BMC Genomics* 2011; **12**: 42.
- 11 Krausa P, Browning MJ. A comprehensive PCR-SSP typing system for identification of HLA-A locus alleles. *Tissue Antigens* 1996; **47**: 237–244.
- 12 Gatz SA, Pohla H, Schendel DJ. A PCR-SSP method to specifically select HLA-A*0201 individuals for immunotherapeutic studies. *Tissue Antigens* 2000; **55**: 532–547.
- 13 Liang B, Zhu L, Liang Z, Weng X, Lu X, Zhang C *et al*. A simplified PCR-SSP method for HLA-A2 subtype in a population of Wuhan, China. *Cell Mol Immunol* 2006; **3**: 453–458.
- 14 Williams F, Middleton D, Savage D, Gorodezky C, Wilson DW, Fitzgerald JM *et al*. Development of PCR-SSOP for the identification of HLA-A*02 subtypes and determination of HLA-A*02 frequencies within different ethnic populations. *Tissue Antigens* 1997; **49**: 129–133.
- 15 Dalva K, Beksac M. HLA typing with sequence-specific oligonucleotide primed PCR (PCR-SSO) and use of the Luminex technology. *Methods Mol Med* 2007; **134**: 61–69.
- 16 Fleischhauer K, Zino E, Mazzi B, Severini GM, Benazzi E, Bordignon C. HLA-A*02 subtype distribution in Caucasians from northern Italy: identification of A*0220. *Tissue Antigens* 1996; **48**: 673–679.
- 17 Selvakumar A, Granja CB, Salazar M, Alasco SM, Yunis EJ, Dupont B. A novel subtype of A2 (A*0217) isolated from the South American Indian B-cell line AMALA. *Tissue Antigens* 1995; **45**: 343–347.
- 18 Fischer GF, Fae I, Frey E, Mayr WR. HLA-A*02172* adds to the heterogeneity of HLA-A*02 alleles. *Tissue Antigens* 1998; **51**: 312–314.
- 19 Browning MJ, Madrigal JA, Krausa P, Kowalski H, Allsopp CE, Little AM *et al*. The HLA-A,B,C genotype of the class I negative cell line Daudi reveals novel HLA-A and -B alleles. *Tissue Antigens* 1995; **45**: 177–187.
- 20 Krausa P, Barouch D, Bodmer JG, Browning MJ. Rapid characterization of HLA class I alleles by gene mapping using ARMS PCR. *Eur J Immunogenet* 1995; **22**: 283–287.
- 21 Krausa P, Barouch D, Bodmer JG, Hill AV, Mason C, McMichael AJ *et al*. Characterization of a novel HLA-A2 variant, A*0214, by ARMS-PCR and DNA sequencing. *Immunogenetics* 1995; **41**: 50.
- 22 Krausa P, Brywka M 3rd, Savage D, Hui KM, Bunce M, Ngai JL *et al*. Genetic polymorphism within HLA-A*02: significant allelic variation revealed in different populations. *Tissue Antigens* 1995; **45**: 223–231.
- 23 Dunn PP, Cox ST, Little AM. Sequencing protocols for detection of HLA class I polymorphism. *Methods Mol Biol* 2003; **210**: 191–222.
- 24 Cereb N, Maye P, Lee S, Kong Y, Yang SY. Locus-specific amplification of HLA class I genes from genomic DNA: locus-specific sequences in the first and third introns of HLA-A, -B, and -C alleles. *Tissue Antigens* 1995; **45**: 1–11.
- 25 Kotsch K, Wehling J, Kohler S, Blasczyk R. Sequencing of HLA class I genes based on the conserved diversity of the noncoding regions: sequencing-based typing of the HLA-A gene. *Tissue Antigens* 1997; **50**: 178–191.