Philippine Journal of Biochemistry and Molecular Biology
2020; 1(1): 29-36
http:// http://www.psbmb.org/pjbmb.html
doi: 10.5555/pjbmb.ph.2020.01.01.29

PJBMB

Short Communication

Identification of Class I HLA Alleles in Anonymized Cell Therapy Specimens through Real-Time PCR with Melt-Curve Analysis

Joanne Jennifer E. Tan[1], Maria Teresa A. Barzaga[2] and Francisco M. Heralde III[2,3]

[1]College of Medicine, University of the Philippines Manila
[2]Molecular Diagnostics and Cellular Therapeutics Laboratory, Lung Center of the Philippines
[3]Department of Biochemistry and Molecular Biology, College of Medicine, University of the Philippines Manila

Email address:
jetan1@up.edu.ph

To cite:
Tan, JJE; Barzaga, MTA; Heralde, FIIIM.2020. Identification of Class I HLA Alleles in Anonymized Cell Therapy Specimens through Real-Time PCR with Melt-Curve Analysis. PJBMB. Vol. 1, No. 1, 2020, pp. 29-36. doi: 10.5555/pjbmb.ph.2020.01.01.29

Received: 09 19, 2019; **Accepted**: 11 18, 2020; **Published**: 12 06, 2020

Abstract: Accurate human leukocyte antigen (HLA) typing is crucial for allogeneic dendritic cell anti-cancer vaccination, where at least a seven out of eight HLA match to the medium resolution allele level is required. Molecular methods can provide medium to high resolution typing but are expensive and time-consuming. A modified method to facilitate faster and more efficient medium resolution identity matching for some common alleles in Filipinos was developed using real-time PCR with melt-curve analysis. The most common Filipino HLA Class I alleles identified from available databases were: A*02, A*24, B*15 and C*07. Primers specific to these alleles were designed. DNA were extracted from 17 de-identified unstimulated stem cell specimens from the Molecular Diagnostics and Cellular Therapeutics Laboratory of the Lung Center of the Philippines (MDCTL-LCP). Comparison of the melt-curve profile was able to determine some degree of HLA allele identity. Distinct alleles from the samples that were identified through sequencing were A*03:01 and C*07:32. Higher resolution typing was not possible for some alleles due to highly similar sequences in the amplified region. The study thus explored the use of real-time PCR and melting curve analysis in describing some common HLA Class I alleles in Filipinos. Further replication and more specific primers will be needed to establish the melting curve profiles for HLA identification use.

.

Keywords: allogeneic dendritic cell vaccine; HLA Class I typing; real-time PCR with melt-curve analysis; A*03:01 and C*07:32 alleles

1. INTRODUCTION

Human Leukocyte Antigens (HLA) are cell surface proteins that present antigenic peptides to generate immune defense reactions (Choo, 2007; Cruz, 2017). The HLA loci are the most polymorphic genes present in the whole genome, which ensure that few individuals are highly similar, and that the population is well-equipped to deal with all types of immune attacks (Choo, 2007). There are two classes of HLA based on their structure and function: HLA Class I and Class II. Class I molecules, the most common ones being HLA-A, -B, and -C, are expressed on virtually all nucleated cells, whereas class II molecules, which include HLA-DR, -DQ and -DP, are only expressed on “immune competent” cells, such as dendritic cells, B lymphocytes, and macrophages (Allard et al., 2014; Leddon et al., 2010).

HLA typing is primarily performed to prevent graft rejection and graft vs. host disease, which can happen when HLA types do not match during cell or tissue transplantation (Choi et al. 2009). There are two main categories of HLA typing: serological-based and molecular-based (Paunic et al., 2012). HLA typing by serology is the oldest and most common method in routine clinical setting (Erlich et al., 2012). However, serology may be unreliable and only provides low

resolution typing, which will give a limited detection of HLA polymorphism (Erlich et al., 2012). With the development of molecular techniques, medium to high resolution methods of HLA typing are now available (Paunic et al., 2012). Molecular techniques, or DNA-based histocompatibility testing, utilize the polymerase chain reaction (PCR) which is the general method used to amplify specific regions of DNA.

The most common PCR-based HLA typing techniques include sequence-specific oligonucleotide probes (SSO), sequence-specific primers (SSP), and sequence-based typing (SBT) (Perng, et al., 2012). PCR-SSO is a relatively a high-throughput and inexpensive method, and is usually used for large-scale, low-resolution HLA allele analysis. SSP provides medium to high resolution and is typically used on samples that have failed to be analyzed by SSO, since it is more expensive and not ideal for many samples. SBT has the highest resolution and is the only way to directly sequence and identify new alleles, although it is also the most expensive. Real-time PCR is now also being used in HLA typing since it involves minimal hands-on time and thus subject to less human error, is less expensive compared to other methods, and does not require post-PCR processing, therefore reducing the risk of contamination (Gersuk & Nepom, 2006). SBT is regarded as the gold standard in HLA typing (Perng, et al., 2012).

Anti-cancer vaccination using peptide-pulsed dendritic cells depends on the interaction of the HLA molecule and the corresponding epitope. Accurate HLA-typing is therefore very crucial for successful anti-cancer vaccination, especially in the case of allogeneic vaccination for immunocompromised patients, where at least a seven out of eight HLA match to the medium resolution allele level is required. Furthermore, characterization and classification of HLA molecules into superfamilies or supertypes in terms of peptide-binding specificities is valuable for the development of anti-cancer vaccines. Previous studies have used bioinformatics to perform hierarchical clustering and principal component analysis to classify HLA class I and class II alleles into such supertypes (Doytchinova et al., 2004). In this study, a method to facilitate faster and more efficient medium resolution identity matching for some common alleles in Filipinos using real-time PCR with melt-curve analysis was explored. Using available databases for primer design, DNA from de-identified unstimulated stem cell specimens from the Molecular Diagnostics and Cellular Therapeutics Laboratory of the Lung Center of the Philippines (MDCTL-LCP) were evaluated thru real-time PCR assay with melt-curve analysis and HLA identification was validated through sequencing.

2. METHODOLOGY

2.1 Samples and DNA extraction

Previously extracted DNA (following DNeasy Blood and Tissue Extraction Kit protocol) from 17 de-identified unstimulated stem cell specimens were obtained from the MDCTL-LCP. The specimens were derived from stored samples of patients under the stem cell therapeutics program of the Lung Center of the Philippines, who accomplished informed consent forms (ICF) specifying that their blood samples, particularly unstimulated stem cells (USCs) and dendritic cells, will be collected and separated following established clinical protocols for collection from the bone marrow, adipose tissue or through leukapheresis, and stored for an indefinite amount of time in a cryolocator containing liquid nitrogen. Access to the samples are limited to the physicians and members of the MDCTL-LCP project team, with no information regarding the identity of the patient to be disclosed. The ICF includes use of the specimens for the conduct of research that are either basic or applied, with the goal of improving LCP's molecular diagnostics and therapeutics program.
.

2. Primer Design

Selection of supertypes. The most common alleles for the Filipino population were obtained from the dbMHC database (https://www.ncbi.nlm.nih.gov/gv/mhc/). Their corresponding supertypes were then identified using the classification system of Doytchinova et al. (2004).

Primer design. Primers specific to the chosen alleles were designed using Primer-BLAST with standard settings (https://www.ncbi.nlm.nih.gov/tools/primer-blast/index.cgi) (Ye et al., 2012). Primers with amplicon lengths around 200 bp (range: 178-220) were selected arbitrarily for its intended use for real-time PCR, as longer lengths do not amplify as efficiently (Smith and Osborn, 2009). The primers were made to amplify exon 2 or 3 of the selected HLA allele, since these are the regions known to contain the most polymorphisms (total length of alleles ~1000 bp). Alleles were aligned using the IMGT/ HLA database (https://www.ebi.ac.uk/ipd/imgt/hla/index.html) (Li et al., 2015). The designed primer sequences and amplicon lengths are found in **Table 2**. Primers were also cross-checked through the NCBI database to ensure that there it will not amplify other regions in the genome.

2.3 Optimization of the Real-Time PCR Assay

DNA amplification and real-time analysis**.** The optimized PCR mix (QuantiTect SYBR Green PCR Kit cat. No. 204141) with total volume 20 µL is composed of 10 µL reaction mix, 0.75 µL each of the forward and reverse primers, 4.5 µL deionized water, and 4 µL DNA. Optimized PCR conditions include 15 min initial denaturation at 95°C, followed by 40 cycles of 5 sec denaturation at 94°C, 10 sec annealing at 55°C and 20 sec extension at 72°C. The minimum DNA template concentration used for amplification was 30 ng/uL. Cut-off Ct value based on the no-template GAPDH control was 32.94. There were no replicates done due to limited sample volumes. Housekeeping primers used were GAPDH forward, 5´-ACCCACTCCTCCACCTTTG-3´;

and GAPDH reverse, 5′-CTCTTGTGCTCTTGCTGGG-3' (Cao, et al., 2008). The real-time PCR reactions were performed using Rotor-Gene Q (2011 model) high precision Real Time thermocycler by Qiagen. Amplification was confirmed using gel electrophoresis.

Gel electrophoresis and extraction. Gel was run using 2.5% agarose in 1x TBE buffer, at 130V, 110A for 30 min. Biotium Gel Red Nucleic Acid stain and Amresco ladder EZ-Vision 100 bp were used. The concentration of agarose (2.5%) was selected according to the range of effective separation, since the lengths were around 200bp and differences were between 20-50bp. Gel extraction was done using QIAquick Gel Extraction Kit.

DNA sequencing. Samples positive for amplification from the melt-curve analysis that had clear bands were gel extracted and sent to Macrogen, Korea for sequencing. Duplicate samples from the leftover original amplicon were also sent. No concentrating methods were done for the gel extracted samples

2.4 Melt-Curve and Sequencing Results Analyses

Melt-Curve Analysis. Batch real-time PCR was performed on 17 samples using 5 primer pairs, A*02, A*24, B*15, C*07 and DRB*15. From the high-resolution melt curve analysis, peak temperatures measured up to second decimal place were not enough to distinguish different alleles as confirmed by subsequent sequencing and were much more dependent on the amplicon length. Melt-curve profiles were analyzed using Rotor-Gene Q Series Software 2.1.0. Melt curve peak temperature is a common parameter which may differ for every amplicon and may be considered in analyzing melt curve profiles. (Dwight et al., 2011; Bruzzone et al., 2013)

Sequencing Results Analysis. BLAST was performed using the IMGT/HLA database. Pairwise sequence alignments were performed using EMBOSS-Needle version in 2016 (https://www.ebi.ac.uk/Tools/psa/emboss_needle/nucleotide.html), (Li et al., 2015) and multiple sequence alignments using Clustal Omega version in 2016 (https://www.ebi.ac.uk/Tools/msa/clustalo/) (Li et al., 2015).

3. RESULTS AND DISCUSSION

3.1 Selection of Supertypes

The dbMHC database provides the most data on allele frequencies from a representative Filipino population. The only study on Filipino alleles was from Erlich et al., 2007, which provided a total sample size of 94. Only the top three alleles in terms of frequency for Class I and Class II HLA were chosen for the study due to the limitations in the budget. The alleles and their corresponding supertypes chosen were: A*02 (A2), A*11 (A3), A*24 (A24), B*15 (B27), C*07 (C1), DRB1*15 (DR1) and DQB1*05 (DQ1). **Table 1** shows the selected Class I and II alleles, their frequencies, and their corresponding supertypes.

Table 1. Allele Frequencies for the Filipino Population (Erlich et al., 2007).

Allele	Frequency (N=94)	Supertype
A*02	0.106	A2
A*11	0.266	A3
A*24	0.383	A24
B*15	0.229	B27
C*07	0.362	C1
DRB1*15	0.452	DR1
DQB1*05	0.511	DQ1

3.2 Primer Design

After primer assessment and alignment in the IMGT/HLA database, only widely spaced single mismatches between closely related alleles were found. False priming is unlikely even with the presence of 3'-end mismatches (Lefever et al., 2013). Studies have shown that mismatch specificity is only present during the first few cycles (Lefever et al., 2013; Liu, et al. 2012). A single mismatch at the 3'-end has little or no impact on yield, and there must be at least 4 mismatches to completely block the reaction (Lefever et al., 2013; Liu, et al. 2012).

The study used high resolution melting curve analysis to resolve this problem. The specificity in identity matching will depend on the melt-curve profile and not on whether there was an amplification or not.

Table 2. Designed Primers and Their Amplicon Lengths

Allele	Primer	Sequence (5' → 3')	Amplicon (bp)
A*02	F	TTCTTCACATCCGTGTCCCG	199
	R	GAGTCTGTGAGTGGGCCTTC	
A*24	F	TTTCTCCACATCCGTGTCCC	196
	R	CTGTGAGTGGGCCTTCACTT	
B*15	F	CCCAGTTCGTGAGGTTCGAC	178
	R	GCCTCGCTCTGGTTGTAGTA	
C*07	F	GGTCTCACACCTTCCAGAGG	169
	R	AACTTGCGCTGGGTGATCTG	
DRB1*15	F	TGGCAGCCTAAGAGGAAGTG	187
	R	CCTGCTCCAGGATGTCCTTC	
DQB1*05	F	TGTGCTACTTCACCAACGGG	220
	R	TACGCCACCTCGTAGTTGTG	

3.3 Real-Time PCR Assay and Melt-Curve Analysis

Amplification was confirmed using gel electrophoresis, which showed clear, distinct, single bands, as shown in Figure 2. As can be seen, the amplicon lengths are consistent with the expected lengths of ~200 bp, no multiple bands are present, and all samples had amplicons.

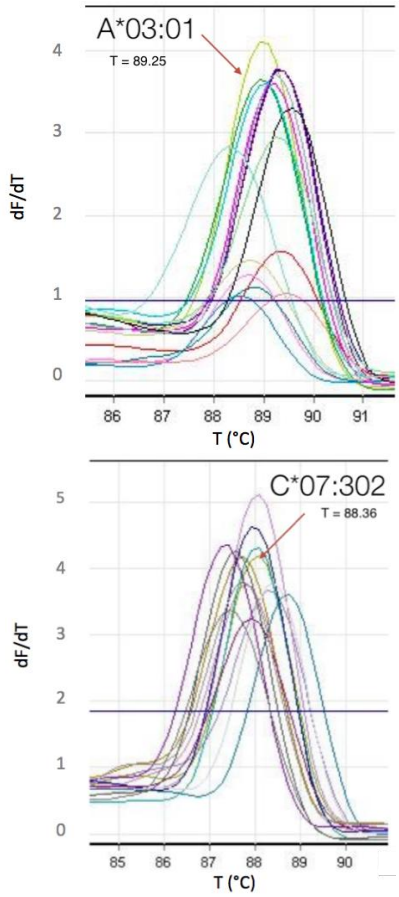


Figure 1. Melting Curves of A*03:01 and C*07:302 showing identity upon superimposition with other amplicons in a similar allele group. Top: melting curve of A*03:01 (arrow) superimposed with other amplicons for primer A*02; Bottom: melting curve of C*07:302 (arrow) superimposed with other amplicons for primer C*07. All other melting curves have inconclusive sequencing results (see Appendix B for more details).

Previous studies have shown that melting curves have different shapes for different specific sequences (Zhou, et al. 2004; Reed and Wittwer, 2004). It is theorized that DNA may be assuming an intermediate state during the melting process, as regions of the amplicon that are more stable do not melt immediately (e.g. G-C rich regions) (Huguet et al., 2010).

All melt curves obtained showed unique profiles as seen in **Appendix A**, suggesting that the HLA alleles of the 17 patient samples were all unique. Also, because of mis-priming due to suboptimal PCR conditions, different PCR products tend to have heterozygous sequences, which can have significant effects on the shape of their melt curve. This is consistent with the single bands seen in the agarose gel electrophoresis.

3.3 Sequencing Results

The actual HLA alleles of the samples were not previously identified, their PCR products were gel extracted and sent for sequencing. Selected samples

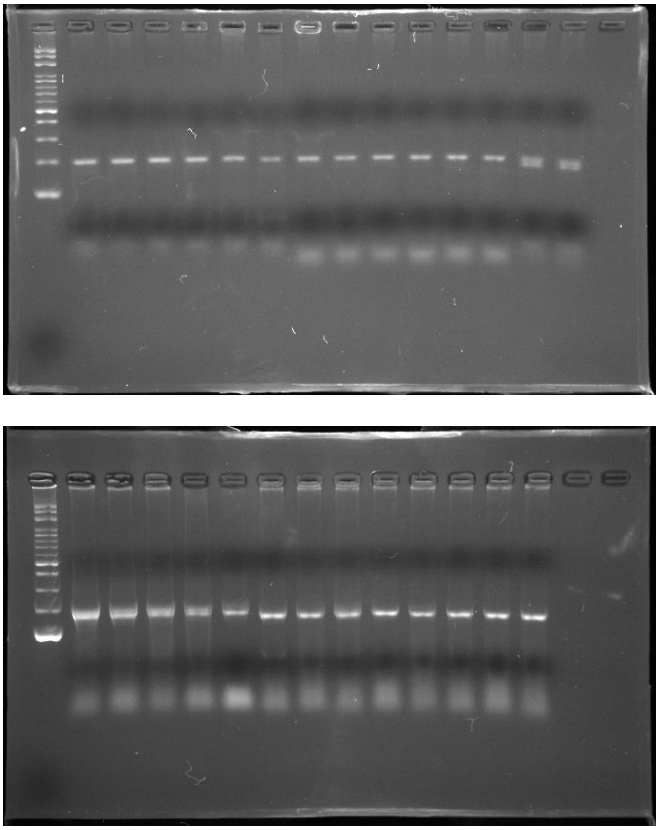


Figure 2. Sample agarose gel electrophoresis for sample codes AA (top) and AB (bottom). Leftmost – 100 bp ladder. (top after ladder: AA: 2, 3, 5, 10, 11, 15, 18, 19, 21, 22, 27, 28, AB1,2; bottom after ladder: AB: 3, 10, 13, 16, 17, 21, 22, 24, 26, 27, 28, 31, 32)

with different melting peak temperatures were sent for sequencing. After performing BLAST using the IMGT/HLA database, HLA allele types to the medium-resolution level were not identified for most samples, despite clear non-overlapping chromatograms. There have been inconsistencies with the sequences for some samples (<75% pairwise sequence alignment identity between forward and reverse sequences, and different allele hits), which could reflect poor sample quality. (See **Appendix B**) Many other samples were only identifiable at the low-resolution level, i.e. only up to the HLA allele group (e.g. HLA-A*02) vs medium resolution which also identifies the specific HLA protein (e.g. HLA-A*02:101). Although % identity for the top hit results from available sequences in the BLAST database were generally high (>90%), specific HLA proteins either could not be identified or there were multiple hits within the same allele group, rendering the results to be inconclusive. This may be due to some alleles having almost perfectly similar sequences in the amplified region, considering the amplicon lengths were all less than 200 bp, which shows some weakness in the primer design. Heterozygous sequences were also seen in some chromatograms. Although multiple sequence alignment showed conserved regions, the primers were not specific enough to be used for sequence-based typing,

and therefore the samples will need to be typed using another method as a standard. On the other hand, there were two alleles identified at the medium-resolution level using this method, A*03:01 and C*07:302. **Figure 1** shows their specific melt curve profile.

With further validation through replication and resequencing using other positive controls, while keeping other factors such as DNA template concentration and purity constant, this method can be used as a tool for possible HLA typing of allogeneic cell therapy specimens.

4. CONCLUSION AND RECOMMENDATIONS

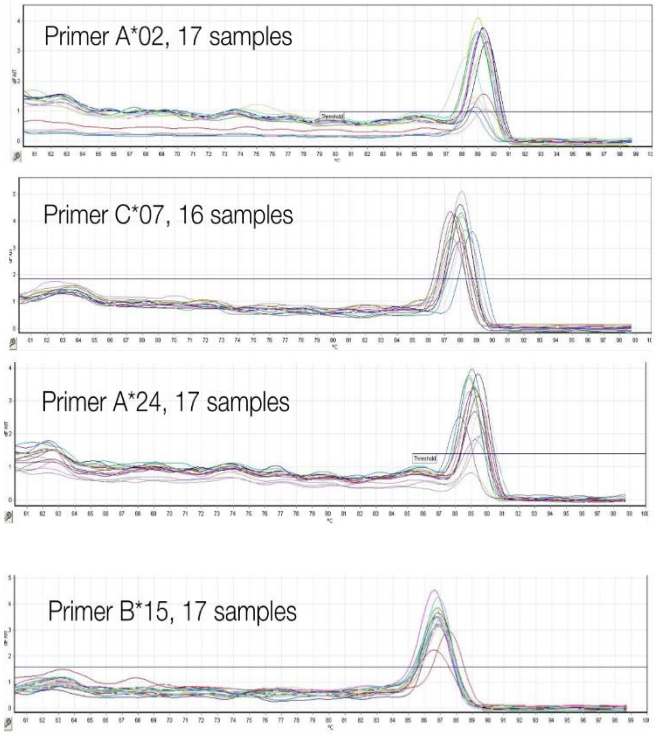
The study explored the use of real-time PCR and melting curve analysis in describing some common HLA Class I alleles in Filipinos. A real-time PCR assay for six selected HLA alleles was optimized. However, only 2 out of 17 samples were identified at the HLA allele and protein level, namely, HLA A*03:01 and C*07:302, which were confirmed by sequencing. This may be due to highly similar sequences in the amplified region and heterozygous amplification. More specific primers may be designed to address this problem. Melt-curve peak temperature alone is not enough to differentiate HLA alleles so the actual profiles may need to be compared. The findings in this study will need further validation through replication and sequencing to establish the melting curve profiles for HLA identification use.

REFERENCES

1. Allard M, Oger R, Benlalam H, Florenceau L, Echasserieau K, Bernardeau K, Labarrière N, Lang F, & Gervois N. 2014. Soluble HLA-I/peptide monomers mediate antigen specific CD8 T cell activation through passive peptide exchange with cell-bound HLA-I molecules. J Immunol. 192(11):5090–7. doi:10.4049/jimmunol.1303226
2. Bruzzone CM, Tawadros PS, et al. (2013) Enhanced Primer Selection and Synthetic Amplicon Templates Optimize High-Resolution Melting Analysis of Single-Nucleotide Polymorphisms in a Large Population.Genet Test Mol Biomarkers, 17(9):675–680.
3. Cao Y, Huschtscha L, Nouwens A, Pickett H, Neumann A, Chang A, Toouli C , Bryan T, Reddel R. 2008. Amplification of hTERT in human mammary epithelial cells with limiting hTR expression levels. Cancer Res. 68(9):3115-23.
4. Choi S, Levine J, & Ferrara, J. 2009. Management of Graft-versus-Host Disease. Immunol Allergy Clin N Am 30 (2010) 75–101.doi:10.1016/j.iac.2009.10.001 immunology.theclinics.com
5. Choo SY. 2007. The HLA System: Genetics, Immunology, Clinical Testing, and Clinical Implications. Yonsei Medical Journal. 48(1):11-23. doi:10.3349/ymj.2007.48.1.11.
6. Crux NB and Elahi S. 2017. Human Leukocyte Antigen (HLA) and Immune Regulation: How Do Classical and Non-Classical HLA Alleles Modulate Immune Response to Human Immunodeficiency Virus and Hepatitis C Virus Infections? Front Immunol. 2017 Jul 18;8:832. doi: 10.3389/fimmu.2017.00832.
7. Doytchinova, I, Guan, P, & Flower, D. 2004. Identifiying Human MHC Supertypes Using Bioinformatic Methods. Journal of Immunology.172:4314-4323.
8. Dwight Z, Palais R, Wittwer CT. (2011) uMELT: prediction of high-resolution melting curves and dynamic melting profiles of PCR products in a rich web application. Bioinformatics. 27(7):1019–1020.
9. Erlich HA, Alejandrino M, Pozzili P, Panelo A, & Bugawan TL. 2007. Filipino from Luzon Island, Philippines. Anthropology/human genetic diversity population reports. In: J.A. Hansen, ed. Immunobiology of the Human MHC: Proceedings of the 13th International Histocompatibility Workshop and Conference. I:620-621.
10. Gersuk, V & Nepom, G. 2006. A real-time PCR approach for rapid high resolution subtyping of HLA-DRB1*04. J Immunol Methods. 317(1-2):64–70.
11. Huguet, J. M., Bizarro, C. V., Forns, N., Smith, S. B., Bustamante, C., & Ritort, F. (2010). Single-molecule derivation of salt dependent base-pair free energies in DNA. *Proceedings of the National Academy of Sciences of the United States of America*, *107*(35), 15431–15436. https://doi.org/10.1073/pnas.1001454107
12. Leddon SA, Sant AJ. 2010. Generation of MHC class II-peptide ligands for CD4 T-cell allorecognition of MHC class II molecules. Curr Opin Organ Transplant. 15(4):505–11. doi:10.1097/MOT.0b013e32833bfc5c
13. Lefever S, Pattyn F, Hellemans J, & Vandesompele J. 2013. Single-Nucleotide Polymorphisms and Other Mismatches Reduce Performance of Quantitative PCR Assays. Clinical Chemistry. 59,10:1470–1480.
14. Li, W., Cowley, A., Uludag, M., Gur, T., McWilliam, H., Squizzato, S., Park, Y. M., Buso, N., & Lopez, R. (2015). The EMBL-EBI bioinformatics web and programmatic tools framework. *Nucleic acids research*, *43*(W1), W580–W584. https://doi.org/10.1093/nar/gkv279
15. Liu J, Huang S, Sun M, Liu S, Liu Y, Wang W, Zhang X, Wang H, & Hua W. 2012. An improved allele-specific PCR primer design method for SNP marker analysis and its application. Plant Methods. 8:34.
16. Lyubchenko YL, Frank-Kamenetskii MD, Vologodskii AV, Lazurkin YS, & Gause GG Jr. 1976. Fine Structure of DNA Melting Curves. Biopolymers. 15:1019-1036.
17. Paunic V, Gragert L, Madbouly A, Freeman J, Maiers M. 2012. Measuring Ambiguity in HLA

Typing Methods. PLoS ONE 7(8): e43585. doi:10.1371/journal.pone.0043585
18. Perng CL, Chang LF, Chien WC, Lee TD, & Chang JB. 2012. Effectiveness and limitations of resolving HLA class I and class II by heterozygous ambiguity resolving primers (HARPs) — a modified technique of sequence-based typing (SBT). Clinical Biochemistry. 45:1471–1478.
19. Reed G and Wittwer C. 2004. Sensitivity and Specificity of Single-Nucleotide Polymorphism Scanning by High-Resolution Melting Analysis. Clinical Chemistry 50:10: 1748–1754.
20. Smith CJ, Osborn AM. Advantages and limitations of quantitative PCR (Q-PCR)-based approaches in microbial ecology. FEMS Microbiol Ecol. 2009 Jan;67(1):6-20. doi: 10.1111/j.1574-6941.2008.00629.x. PMID: 19120456.
21. Ye et al. 2012. Primer-BLAST: A tool to design target-specific primers for polymerase chain reaction. BMC Bioinformatics 2012 13:134.
22. Zhou L, Vandersteen J, Wang L, Fuller T, Taylor M, Palais B, & Wittwer CT. 2004. High-resolution DNA melting curve analysis to establish HLA genotypic identity. Tissue Antigens. 64:156–164.

APPENDIX A. RAW MELTING CURVE PROFILES



APPENDIX B: SEQUENCING SUMMARY DATA

Primer	Sample (DNA#)	Length	Melt Peak Temp	% Identity	Top Hit
A*02	AA1	173	89.25	100	Axx
	(3)	171		100	A03:01
	AA2	171	89.6	100	A02x
	(1)	168		100	A02x
	AA3	145	88.9	100	A24:29
	(11)	165		100	Axx
	AA4	147	89.4	100	A24x
	(16)	167		100	A03:01
	AA17	190	89.25	95	A02x
	(17)	174		94	A02x
	Y1	150	89.6	100	A24x
	(2)	169		100	A03x
	Y2	145	89.65	100	A24x
	(8)	170		100	Axx
A*24	AA10	183	89.1	93	A02x
	(10)	173		97	A02:01
	Y3	144	89.6	100	A24x
	(2)	164		97	A68:142
	Y4	144	89.85	99	A24x
	(8)	164		99	A24x
B*15	AB3	100	87.2	97	B27:109
	(1)	156		93	B37:06
	AB13	68	87.85	96	Cxx
	(6)	151		98	Gxx
	AB14	63	87.35	97	Bxx
	(3)	145		92	Bxx
	Y5	100	87.1	xxx	xxx
	(2)	153		95	B37:52
	Y6	100	87.15	xxx	xxx
	(8)	163		96	Bxx
C*07	AB21	144	88.1	100	Bxx
	(11)	100		xxx	xxx
	AB22	145	88.36	100	C07x
	(10)	144		100	C07:302
	AB24	100	89.11	98	Axx
	(7)	141		99	A03:95
	AB31	143	87.85	100	Cxx
	(19)	141		97	C04x
	AB32	100	87.95	96	Bxx
	(18)	100		xxx	xxx
	Y7	143	87.75	100	B40x
	(2)	145		95	C04x
	Y8	141	87.6	100	C08x
	(8)	143		98	C08x

Note:
Top Hit: Sequence similarity top hit from BLAST – IMGT/HLA database (top: forward, bottom: reverse)
x: specific HLA protein could not be identified
xx: HLA group could not be identified
xxx: no significant matches found
% identity: percent extent to which the sample sequence and the top hit sequence have the same alignment residues (top: forward, bottom: reverse)

APPENDIX C: SEQUENCES

AA1
(F)GGNNNGGNANCGCTTCTCGCCGTGGGCTACGTGGACGACACGCAGTTCGT
GCGGTTCGACAGCGACGCCGCGAGCCAGAGGATGGAGCCGCGGGCGCCGTG
GATAGAGCAGGAGGGGCCGGAGTATTGGGACCGGAACACACGGAATGTGAAG
GCCCACTCACAGACTACCAC
(R)GTTTCTTCACATCCGTGTCCCGGCCCGGCCGCGGGGAGCCCCGCTTTATC
GCCGTGGGCTACGTGGACGACACGCAGTTCGTGCGGTTCGACAGCGACGCC
GCGAGCCAGAGGATGGAGCCGCGGGCGCCGTGGATAGAGCAGGAGGGGCC
GGAGTATGACCGAANCCNGN
AA2
(F)GNNNGGNGGGAACGCTTCTCGCAGTGGGCTACGTGGACGACACGCAGTTC
GTGCGGTTCGACAGCGACGCCGCGAGCCAGAGGATGGAGCCGCGGGCGCCG
TGGATAGAGCAGGAGGGTCCGGAGTATTGGGACGGGGAGACACGGAAAGTG
AAGGCCCACTCACAGACT
(R)TTTCTTCACATCCGTGTCCCGGCCCGGCCGCGGGGAGCCCCGCTTCATCG
CAGTGGGCTACGTGGACGACACGCAGTTCGTGCGGTTCGACAGCGACGCCGC
GAGCCAGAGGATGGAGCCGCGGGCGCCGTGGATAGAGCAGGAGGGTCCGGA
GTATGGACGGNNANG
AA3
(F)GNCCGGNACNCTTCTCGCCGTGGGCTACGTGGACGACACGCAGTTCGTGC
GGTTCGACAGCGACGCCGCGAGCCAGAGGATGGAGCCGCGGGCGCCGTGGA
TAGAGCAGGAGGGGCCGGAGTATTGGGACGAGGAGACACGGAAA
(R)TTTCTTCACATCCGTGTCCCGGCCCGGCCGCGGGGAGCCCCGCTTCATCG
CAGTGGGCTACGTGGACGACACGCAGTTCGTGCGGTTCGACAGCGACGCCGC
GAGCCAGAGGATGGAGCCGCGGGCGCCGTGGATAGAGCAGGAGGGTCCGGA
GTNTGNCGNNNN
AA4
(F)GGCCGGGNNCCGCTTCTCGCCGTGGGCTACGTGGACGACACGCAGTTCGT
GCGGTTCGACAGCGACGCCGCGAGCCAGAGGATGGAGCCGCGGGCGCCGTG
GATAGAGCAGGAGGGGCCGGAGTATTGGGACGAGGAGACAGGGAAA
(R)TTTCTTCACATCCGTGTCCCGGCCCGGCCGCGGGGAGCCCCGCTTTATCG
CCGTGGGCTACGTGGACGACACGCAGTTCGTGCGGTTCGACAGCGACGCCG
CGAGCCAGAGGATGGAGCCGCGGGCGCCGTGGATAGAGCAGGAGGGGCCG
GAGTNTGACGAGANNN
AA17
(F)NCNNTNNGNNCCNGCTTCNNCGCAGTGGGCTACGTGGACGACACGCAGCT
CCGTGCGCGATCGACAGCGACCCCGCGAGCCAGAGCATGGAGCCGCGGGCC
CGTGGATAGAGCAGGAGGGTCCGGAGTATTGGGACGGGGAGACACGGAAAG
TGAAGGCCCACTCACAGACTACGGGACACGGAAGTGAA
(R)GAGTCTGTGAGTGGGCTTTCTTTACATCCGTGTCCGCCNGNCGCGGGGAG
CCCGCTTCATCGCAGTGGGCTACGTGGACGACACGCAGTTCGTGCGGTTCGA
CAGGCGCCGCGAGCCAGAGGAGGGACGCGGGCGCCGTGGATAGAGCAGGA
GGGTCCGGAGTNTGACAGANGN
Y1
(F)CNGNGNCNGGGAACGCTTCTCGCCGTGGGCTACGTGGACGACACGCAGTT
CGTGCGGTTCGACAGCGACGCCGCGAGCCAGAGGATGGAGCCGCGGGCGCC
GTGGATAGAGCAGGAGGGGCCGGAGTATTGGGACGAGGAGACAGGGAAA
(R)TTTTCTTCACATCCGTGTCCCGGCCCGGCCGCGGGGAGCCCCGCTTCATC
GCCGTGGGCTACGTGGACGACACGCAGTTCGTGCGGTTCGACAGCGACGCC
GCGAGCCAGAGGATGGAGCCGCGGGCGCCGTGGATAGAGCAGGAGGGGCC
GGAGTNTGNCGNGANNAGN
Y2
(F)GGGGGNGGNACCGCTTCTCGCCGTGGGCTACGTGGACGACACGCAGTTC
GTGCGGTTCGACAGCGACGCCGCGAGCCAGAGGATGGAGCCGCGGGCGCCG
TGGATAGAGCAGGAGGGGCCGGAGTATTGGGACGAGGAGACAGGG
(R)TTTCTTCACATCCGTGTCCCGGCCCGGCCGCGGGGAGCCCCGCTTCATCG
CCGTGGGCTACGTGGACGACACGCAGTTCGTGCGGTTCGACAGCGACGCCG
CGAGCCAGAGGATGGAGCCGCGGGCGCCGTGGATAGAGCAGGAGGGGCCG
GAGTATGACGAGGANNNGT
AA10
(F)CNNNNGNACCCANNCTCATCGCAGTGGCTACGCGGACGACACGCAGTCGT
GCCCATCGACAGCGACCCGCGAGCCAGAGGACGGAGCCGCGGGCCCGTGCA
TAGAGCAGGAGGGTCCGGAGTATTGGGACGGGGAGACACGGAAAGTGAAGG
CCCACTCACAGAGGACACGGATGTGAGAAAA
(R)TCTGTGAGTGTGCTTACTTTCTCTACACCGTTCCCAGCCGGGGCGGGGACC
CCGCTTCATCGCAGTGGGCTACGTGGACGACACGCAGTTCGTGCGGTTCGAC
ACGACGCCGCGAGCCAGAGGATGGGACCCGGGCGCCGTGGATAGAGCAGGA
GGGTCCGGAGTCNNACNNN
Y3
(F)GNNNNGNNACCGCTTCTCGCCGTGGGCTACGTGGACGACACGCAGTTCGT
GCGGTTCGACAGCGACGCCGCGAGCCAGAGGATGGAGCCGCGGGCGCCGTG
GATAGAGCAGGAGGGGCCGGAGTATTGGGACGAGGAGACAGGG
(R)TTTGATTTACAATAACTTAAGTCCGTCGAAAAGAGCTTCCTTCATCGCCGTG
GGCTACGTGGACGACACGCAGTTCGTGCGGTTCGACAGCGACGCCGCGAGC
CCAGAGGATGGAGCCGCGGGCGCCGTGGATAGAGCAGGAGGGCCGGAGTAN
TGNNNNNNNN
Y4
(F)GNNNNGGNCCCGCTTCTCGCCGTGGGCTACGTGGACGACACGCAGTTCGT
GCGGTTCGACAGCGACGCCGCGAGCCAGAGGATGGAGCCGCGGGCGCCGTG
GATAGAGCAGGAGGGGCCGGAGTATTGGGACGAGGAGACAGGG
(R)TTTTCTCCACATCCGTGTCCCGGCCCGGCCGCGGGGAGCCCCGCTTCATC
GCCGTGGGCTACGTGGACGACACGCAGTTCGTGCGGTTCGACAGCGACGCC
GCGAGCCAGAGGATGGAGCCGCGGGCGCCGTGGATAGAGCAGGAGGGCCG
GAGTANGNNANNN
AB3
(F)GNNTNNTTCGAGGATGGAGCGCGGGCGCCGTGGATAGAGCAGGGAGGGG
CCGGAGTATTGGGACCGAGGAGACACGGATCTGCAAGGACCCCGCACAGAC
(R)TCCCAGTTTGTGAGGTTTGACAGCGACTCGGCGAGTCCGAGGATGGAGCC
GCGGGCGCCGTGGATGGAGCAGGAGGGGCCGGAGTATTGGGACGCGGGAG
ACACAGATCGTGTCAAGGCCCACGGCACAGACTGACACGANNAACNNTNGAT
NTNN
AB13
(F)NNNNGGGTCAGGGAGNAGCGCGGGCGCCGTGGANGGGCAGGAGGGGCC
GGAGTATTGGGACCGGGAGA
(R)TCCCAGTTCGTGAGGTTCGACAGCGACTCGGCGTGTCCGAGGATGGAGCC
GCGGGCGCCGTGGGTGGAGCAGGAGGGGCCGGAGTATTGGGACGCGGGAG
ACACGAGATCGTGCAAGGCCCACGCACAGACTACAGAGAGANCGCGNNNNN
AB14
(F)GNNNNNNNCGAGGAGGAGCGCGGGCGCCGTGGATAGAGCAGGAGGGGC
CGGAGTATTGCGACC
(R)TCCCAGTTCGTGAGGTTCGACAGCGACGCCGCGAGTCCGAGGATGGAGCC
GCGGGCGCCGTGGATGGAGCAGGAGGGGCCGGAGTATTGGACCGGGAGACA
CGGACCTGCAAGCCCACGCACAGACTGACCGAAGAACGCGTTNN
Y5
(F)NNNNNNNNGGTCAAAGAGNGCGCGGGCGCGTGGATAGAACCGGGAGGGG
CCCGGAGTACTGGGACCGGGACAGAAGTGTTGTTAGTTTTATTTCCGATCG
(R)TTCCCAGTTCGTGAGGTTTGACAGCGACTCCGCGAGTCCGAGGATGGAGC
CGCGGGCGCCGTGGGTGGAGCAGGAGGGGCCGGAGTATTGGGACGCGGGA
GACACAGATCGTGCAAGGCCCACGCACAGACTACAGGAAGGAACCGCGNNNC
N
Y6
(F)TGCACNGGTCCGAGAGGAGCGCGGGCGCCGTGGATAGANCAGAGGGGCC
GAGAGTATTGGGACACGAGGAGACACGGAACGCCAAGACAACGCCAACTGA
(R)TCCCAGTTCGTGAGGTTCGACAGCGACGCCGCGAGTCCGAGGATGGAGCC
GCGGGCGCCGTGGATGGAGCAGGAGGGGCCAGAGTATTGGGGACGCGGGG
AGGACACGAGGATCATCTGAAGGCCCACNCACCAGGACTGTCNNAGATGNCC
CGAANAAGNNA
AB21
(F)NNNGGCNCNGNNGNCCGGACGGGCGCCTCCTCCGCGGGCATAACCAGTA
CGCCTACGACGGCAAGGATTACATCGCCCTGAACGAGGACCTGCGCTCCTGG
ACCGCCGCGGACACGGCGGCTCAGATCACCCAGCGCAAGTACC
(R)TGTGACGGCTTCTCGTCCACATCCGCATCGGCTGGCGGCCGGGGGGGCG
GCCGGGGTTCTGGCCCGAACGAGGACCGCGCTCCGANCCNNNTTNNNNNGA
AB22
(F)GGNANGNNGCACTGGGGCGACGGGCGCCTCCTCCGCGGGTATGACCAGT
CCGCCTACGACGGCAAGGATTACATCGCCCTGAACGAGGACCTGCGCTCCTG
GACCGCCGCGGACACCGCGGCTCAGATCACCCAGCGCAAGTTAT
(R)GTGGTCTCACACCTTCCAGAGGATGTCTGGCTGCGACCTGGGGCCCGACG
GGCGCCTCCTCCGCGGGTATGACCAGTCCGCCTACGACGGCAAGGATTACAT
CGCCCTGAACGAGGACCTGCGCTCNTACCGCCCGNGATNCCC
AB24
(F)GNNNTCACTGGGGNCGGACGGGCGCTTCCTCCGCGGGTACCAGCAGGAC
GCCTACGACGGCAAGGATTACATCGCCCTGAACGAGGACCTGCGCTCTTGG
(R)TTGGTCTCACACCTTCCAGAGGATGTATGGCTGCGACGTGGGGTCGGACG
GGCGCTTCCTCCGCGGGTACCGGCAGGACGCCTACGACGGCAAGGATTACAT
CGCCCTGAACGAGGACCTGCGCTCTGNCCNNGGNNTAGN
AB31
(F)GNNNNNCACTGGGGCGGACGGGCGCCTCCTCCGCGGGTATGACCAGTCC
GCCTACGACGGCAAGGATTACATCGCCCTGAACGAGGATCTGCGCTCCTGGA
CCGCCGCGGACACCGCGGCTCAGATCACCCAGCGCAAGTTAA
(R)TTCTCACACATTCCAGAGGATGTTTGGCTGCGACCTGGGGCCGGACGGGC
GCCTCCTCCGCGGGTATGACCAGTTCGCCTACGACGGCAAGGATTACATCGC
CCTGAACGAGGNTCTGCGCTCCTGACCGCCGNGNNNNNG
AB32
(F)CNNGGCCACCGGNNCGGAGGGCGCCTCCTCCGCGGGCATAACCAGGTCG
CCTACGACGGCAAGGATACATCGCCCTGAACAAGGACCTGAGCTCCGGCAC
(R)AATCGTCACCACAAAACCTCCCCCGCGACCGTAAGCAGCCGCCATGACGG
CAAGGATCCTCGCCCTTCACGAGGACCTGGGCTCCTGNAGCCCCNNCCGT
Y7

(F)CGNGNTCACTGGGGCGGACGGGCGCCTCCTCCGCGGGCATAACCAGTTC
GCCTACGACGGCAAGGATTACATCGCCCTGAACGAGGACCTGCGCTCCTGGA
CCGCCGCGGACACGGCGGCTCAGATCACCCAGCGCAAGTTAA
(R)TTGGTCTCACACCTTCTAGAGGATGTATTGGCTGCGACGTGGGGCCGGAC
GGGCGCCTCCTCCGCGGGCATAACCAGTTCGCCTACGACGGCAAGGATTACA
TCGCCCTGANCGAGGATCTGCGCTCCTNGACCGCGGNGNNNNC

Y8

(F)GNNGTNACTGGGGCGACGGGCGCCTCCTCCGCGGGTATAACCAGTTCGCC
TACGACGGCAAGGATTACATCGCCCTGAATGAGGACCTGCGCTCCTGGACCG
CCGCGGACACGGCGGCTCAGATCACCCAGCGCAAGTAAA
(R)TGGTCTCACACCGTCCAGAGGATGTATGGCTGCGACCTGGGGCCCGACGG
GCGCCTCCTCCGCGGGCATAACCAGTTCGCCTACGACGGCAAGGATTACATC
GCCCTGAATGAGGACCTGCGCTCCTGACCGCGCGATCCANN