



Published in final edited form as:

Hum Immunol. 2019 July ; 80(7): 449–460. doi:10.1016/j.humimm.2019.03.001.

Next-generation HLA typing of 382 International Histocompatibility Working Group reference B-Lymphoblastoid cell lines: report from the 17th International HLA and Immunogenetics Workshop

Lisa E. Creary^{1,23,*}, Sandra G. Guerra², Winnie Chong², Colin J. Brown³, Thomas R. Turner^{4,5}, James Robinson^{4,5}, Will P. Bultitude^{4,5}, Neema P. Mayor^{4,5}, Steven G.E. Marsh^{4,5}, Katsuyuki Saito⁶, Kevin Lam⁶, Jamie L. Duke⁷, Timothy L. Mosbrugger⁷, Deborah Ferriola⁷, Dimitrios Monos^{7,8}, Amanda Willis⁹, Medhat Askar⁹, Gottfried Fischer¹⁰, Chee Loong Saw¹¹, Jiannis Ragoussis¹², Martin Petrek¹³, Carles Serra-Pagés¹⁴, Manel Juan¹⁴, Catherine Stavropoulos-Giokas¹⁵, Amalia Dinou¹⁵, Reem Ameen¹⁶, Salem Al Shemmari¹⁶, Eric Spierings¹⁷, Ketevan Gendzekhadze¹⁸, Gerald P. Morris¹⁹, Qiuhe Zhang²⁰, Zahra Kashi²¹, Susan Hsu²², Sridevi Gangavarapu²³, Kalyan C. Mallempati²³, Fumiko Yamamoto²³, Kazutoyo Osoegawa²³, Tamara Vayntrub²³, Chia-Jung Chang²⁴, John A. Hansen²⁵, and Marcelo A. Fernández-Vi a^{1,23}

¹Department of Pathology, Stanford University School of Medicine, Palo Alto, CA, USA

²Histocompatibility and Immunogenetics Service Development Laboratory, NHS Blood and Transplant, London, UK ³Department of Histocompatibility and Immunogenetics, NHS Blood and Transplant, London, UK ⁴Anthony Nolan Research Institute, Royal Free Hospital, London, UK

⁵UCL Cancer Institute, Royal Free Campus, London, UK ⁶Molecular Biology Research Department, One Lambda, Thermo Fisher Scientific, Canoga Park, CA, USA ⁷Immunogenetics Laboratory, The Children's Hospital of Philadelphia, Philadelphia, PA, USA ⁸Department of Pathology and Lab Medicine, University of Pennsylvania, Philadelphia, PA, USA ⁹Department of Pathology and Laboratory Medicine, Baylor University Medical Center, Dallas, USA ¹⁰Department for Blood Group Serology and Transfusion Medicine, Medical University of Vienna, Vienna, Austria ¹¹HLA Laboratory, Division of Haematology, McGill University Health Centre, Montreal, Canada ¹²Department of Human Genetics, McGill University & McGill University and Genome Québec Innovation Centre, Montreal, Canada ¹³Department of Pathological Physiology and Immunogenomics, IMTM, Faculty of Medicine and Dentistry, Palacky University, Olomouc, Czech Republic ¹⁴Immunology Department, Hospital Clinic de Barcelona, University of Barcelona, IDIBAPS, Barcelona, Spain ¹⁵Biomedical Research Foundation Academy of Athens, Hellenic

* **Corresponding author:** Lisa E. Creary, Ph.D., Department of Pathology, Stanford University School of Medicine, 3155 Porter Drive, Palo Alto, CA 94304, USA, Tel: +1 650 498 1248, Fax: +1 650 724 0294, lcreary@stanford.edu.

Author contributions

LEC was involved in the conception of the study, analysed the data and drafted the manuscript. MFV was involved in the conception of the study. JH provided DNA samples to participants. TV helped to coordinate distribution of the samples to participants. CJH assisted in data analysis. JR, TRT, NPM, and WPB assisted in data analysis and performed NGS genotyping assays. The remaining authors performed NGS genotyping assays.

Disclosure of interests

The authors declare no competing financial or other interests.

Cord Blood Bank, Athens, Greece ¹⁶Health Sciences Center, Kuwait University, Kuwait
¹⁷Laboratory of Translational Immunology, UMC Utrecht, Utrecht, Netherlands ¹⁸HLA Laboratory,
City of Hope, Duarte, CA, USA ¹⁹Department of Pathology, University of California San Diego, La
Jolla, CA, USA ²⁰Department of Pathology and Laboratory Medicine, UCLA Immunogenetics
Center, Los Angeles, CA, USA ²¹HLA department, Kashi Clinical Laboratories, Inc., Portland, OR,
USA ²²HLA Laboratory, American Red Cross, Philadelphia, PA, USA ²³Histocompatibility,
Immunogenetics and Disease Profiling Laboratory, Stanford Blood Center, Palo Alto CA, USA
²⁴Stanford Genome Technology Center, Palo Alto, CA, USA ²⁵Fred Hutchinson Cancer Research
Center, Seattle, WA, USA

Abstract

Extended molecular characterization of HLA genes in the IHWG reference B-lymphoblastoid cell lines (B-LCLs) was one of the major goals for the 17th International HLA and Immunogenetics Workshop (IHIW). Although reference B-LCLs have been examined extensively in previous workshops complete high-resolution typing was not completed for all the classical class I and class II HLA genes. To address this, we conducted a single-blind study where select panels of B-LCL genomic DNA samples were distributed to multiple laboratories for HLA genotyping by next-generation sequencing methods. Identical cell panels comprised of 24 and 346 samples were distributed and typed by at least four laboratories in order to derive accurate consensus HLA genotypes. Overall concordance rates calculated at both 2- and 4-field allele-level resolutions ranged from 90.4% to 100%. Concordance for the class I genes ranged from 91.7 to 100%, whereas concordance for class II genes was variable; the lowest observed at *HLA-DRB3* (84.2%). At the maximum allele-resolution 78 B-LCLs were defined as homozygous for all 11 loci. We identified 11 novel exon polymorphisms in the entire cell panel. A comparison of the B-LCLs NGS HLA genotypes with the HLA genotypes catalogued in the IPD-IMGT/HLA Database Cell Repository, revealed an overall allele match at 68.4%. Typing discrepancies between the two datasets were mostly due to the lower-resolution historical typing methods resulting in incomplete HLA genotypes for some samples listed in the IPD-IMGT/HLA Database Cell Repository. Our approach of multiple-laboratory NGS HLA typing of the B-LCLs has provided accurate genotyping data. The data generated by the tremendous collaborative efforts of the 17th IHIW participants is useful for updating the current cell and sequence databases and will be a valuable resource for future studies.

Keywords

Human leukocyte antigen; Next-generation sequencing; B-Lymphoblastoid cell lines; International HLA and Immunogenetics Workshop; multiple-laboratory testing

1. Introduction

The curation of Epstein-Barr virus (EBV) transformed B-lymphoblastoid cell lines (B-LCLs) for testing factors of the HLA system was initially conceptualized and conducted during the 10th International Histocompatibility Workshop and Conference (IHCW) held in

1987. At the completion of the 10th IHWC, 107 cell lines were extensively characterized by biochemical, serological and molecular techniques, providing a fundamental reference resource for future quality control studies (1,2,3,4,5,6). During the 11th IHWC (1991) the 107 B-LCLs were further genotyped using a PCR sequence-specific oligonucleotide probe (PCR-SSO) method for several class II HLA genes (*HLA-DRA*, *HLA-DRB1*, *HLA-DRB3*, *HLA-DRB4*, *HLA-DRB5*, *HLA-DRB6*, *HLA-DQA1*, *HLA-DQB1*, *HLA-DPA1*, *HLA-DPBI*) (7). The 12th IHWC (1996) saw the expansion of the reference cell panel to include cells with novel HLA sequences: 97 unique cell lines from the 4th Asian-Oceania Histocompatibility workshop were added to the inventory as well as 162 new cell lines gathered from several laboratories located worldwide (8). As a resource for the 13th IHWC, the IHWG Cell and DNA bank was established in 1999 to house and maintain an extensive inventory of the expanded reference B-LCL panel (9). Today, the IHWG cell repository contains 754 B-LCLs obtained from various sources including HLA homozygous and heterozygous donors, Hematopoietic stem cell transplantation (HSCT) patients, selected CEPH (Centre d'Etude du Polymorphisme Humain) families, and individuals from diverse ethnic groups (10). Despite the availability of numerous B-LCLs in the IHWG Cell and DNA bank, only a small fraction of all cells lines have been typed comprehensively for all or the majority of classical class I and class II HLA genes (11,12,13). This is due in part to most laboratories previously typing only the most polymorphic exons that constitute the antigen-recognition site for a few HLA genes commonly *HLA-A*, *HLA-B*, and *HLA-DRB1*. In addition traditional molecular methods used for HLA typing, such as PCR-rSSOP (14), PCR sequence-specific primer (PCR-SSP) (15) and PCR Sanger sequencing based typing (PCR-SSBT) (16,17,18) required too much effort to resolve allele and genotype ambiguities. These ambiguities are often generated due to lack of gene coverage as well as limitations inherent to the technique such as the inability to set phase by heterozygous sequencing (19).

The advent of next-generation sequencing (NGS) and the commercialization of NGS HLA kits around the time of the 16th International HLA and Immunogenetics Workshop (IHIW) (20) provided an ideal opportunity to re-type the reference cell panel at full-gene coverage. The inherent attributes of NGS, namely clonal-based amplification and massive parallel sequencing, mitigates the majority of the HLA phase ambiguities encountered with PCR-SSBT. In addition, NGS combined with long-range PCR, encompassing all exons and introns or key regions representing the antigen binding sites and some introns, ensures wide genomic coverage thereby eliminating allele ambiguities. Furthermore NGS of extended amplicons allows a thorough interrogation of intronic and untranslated regions that may contain pertinent and as yet undefined polymorphisms involved in expression and structural variation. The patch-work pattern of polymorphism across the HLA genes often clustered within exons required more advanced molecular methods to define and discriminate between the numerous alleles collated in the IPD-IMGT/HLA Database (21), as well as to characterize novel alleles that are often detected in the clinical HLA laboratory. Due to the numerous benefits of NGS, namely high-accuracy, high-throughput, low ambiguity and low cost, the 17th IHIW held in Asilomar, California, September 2017 focused on the application of NGS for Histocompatibility testing. Enhancing the IHWG reference cell line repository by extending the molecular characterization of HLA genes for the B-LCLs was one of the major goals of the 17th IHIW. We endeavored to do this by conducting a worldwide

collaborative study of participating groups who were capable of performing NGS typing for HLA. In this report we present the findings of multi-laboratory NGS typing of HLA genes in a select panel of B-LCLs. We also describe novel haplotype associations observed at the 4-field allele level of resolution as well as novel exon and intronic variants characterized by examination of sequence data generated by the participating laboratories.

2. Materials and Methods

2.1. Selection of IHWG reference B-lymphoblastoid cell lines

The 17th IHWG NGS HLA typing reference cell line committee selected B-LCLs, typed at low or intermediate allele-level resolution in previous workshops, for further HLA characterization. Cells were obtained from the IHWG Cell and DNA bank housed at the Fred Hutchinson Cancer Research Center, Seattle, WA (<http://www.ihwg.org>). Selection was primarily based on the following criteria:

1. Common and Well-Documented (CWD) HLA alleles; according to the CWD version 2.0. catalogue (22)
2. Rare (non-CWD) HLA alleles
3. Homozygosity at one or more HLA loci (consanguineous cell lines identical by descent)
4. Heterozygosity at one or more HLA loci
5. Representation of donor cells from diverse ethnic groups

In total we selected 419 B-LCLs from a total of 754 available IHWG cells for a single-blind study. The cells were derived from individuals from a number of ethnic backgrounds including but not limited to; European, Jewish, African American, Australian Aboriginal, Chinese, Japanese, and Native American, ensuring that a wide diversity of HLA alleles were adequately represented in the study. The 419 B-LCLs were compiled into five panels; four small panels consisting of 24 cells (denoted 24U1, 24U2, 24U3, 24U4) and an extended panel (denoted ExtU) comprised of 23 samples from the 24U4 panel plus 323 unique cell samples. Various reference panels were compiled to widen the options of the number of cell lines that would be available for typing by the participating laboratories.

2.2. Selection of participating laboratories

To ensure suitability of the participating laboratory to perform accurate HLA typing by NGS, laboratories were requested to complete NGS HLA typing of quality control (QC) DNA samples derived from B-LCL reference samples included in the NGS of full length genes pilot study and/or the quality control proficiency testing study (Osoegawa *et al.* unpublished data). Laboratories that were at least 95% concordant at 2-field resolution for at least 3 HLA loci in the QC panel were deemed to be successful to perform highly accurate NGS HLA typing. For this study, recruitment began in late 2016 and by September 2017 a total of 18 accredited histocompatibility testing laboratories from 18 institutions agreed to participate in a collaborative single-blind study.

The reference cell line 17th IHIW committee at Stanford University contacted the IHWG cell bank at the Fred Hutchinson Cancer Research Center to distribute genomic DNA derived from the 24U and ExtU panels to participating laboratories. Purified DNA material, each unit labeled with a unique 17th IHIW database code, was provided blinded to participating groups at a concentration of 100 ng/μl per sample, total amount 1.5μg. Participating groups could choose how many reference panels they wanted to type but it was imperative that at least 4 groups typed a single panel in order to maximize accuracy and reproducibility of the HLA genotypes generated. All DNA samples were typed by each laboratory using their current preferred NGS method. The participating laboratories and principal investigators are listed in Table 1.

2.3. HLA typing

Reference cell-line panels were typed for HLA loci using the established NGS methods of the participating laboratories. Most participating laboratories typed using commercial NGS HLA reagents and alignment software: TruSight HLA Assign (Illumina Inc. CA, USA); Holotype HLA & HLA Twin (Omixon, Budapest, Hungary); MIA FORA NGS (Immucor, Inc., Norcross, GA, USA); NGS Engine (GenDx, Utrecht, Netherlands); TypeStream (One Lambda/Thermo Fisher Scientific Inc., CA, USA); Genetics Management System (Cisco Genetics Inc., Seattle, WA). One laboratory used in-house bioinformatics software. Six different types of NGS machines were used: Illumina MiSeq (n = 13), Illumina MiniSeq (n = 1), Illumina NextSeq (n = 1), Ion Torrent™ Personal Genome Machine system (n = 1), Ion S5™ system (n = 1), Ion S5™ XL system (n = 3), and Pacific Biosciences RS II system (n = 1). Some NGS reagents and software were kindly provided by the vendors at no cost as contribution to the 17th IHIW. The number of HLA loci typed at full-gene level for the majority of class I loci and full-gene or wide coverage (encompassing key polymorphic exons and some intronic regions) for class II genes ranged from 3 to 11 across the groups. All operations from PCR amplification through generation of HLA data were performed at the individual participant sites. Details of the NGS instrument and software utilized by participating laboratories as well as HLA genomic regions amplified and sequenced are described in Table 2.

2.4. Submission of HLA data into the 17th IHIW Database

Laboratories submitted HLA data electronically by uploading Histoimmunogenetics Markup Language (HML) or Extensible Markup Language (XML) encoded files containing paired genotype HLA allele lists as well as DNA sequences to the 17th IHIW database <https://ihiws17.stanford.edu>. In order to standardize data input such as HLA allele names, the IPD-IMGT/HLA Database version 3.25.0 (released July 2016) was used as the only reference source for the 17th IHIW database. On those rare occasions where the data submitted by the participating group was analysed using a different IPD-IMGT/HLA Database version (21), the 17th IHIW database automatically converted HLA allele names to correspond to the most similar allele present in version 3.25.0. To further standardize input data and facilitate downstream analyses HLA genotype data was automatically converted to genotype list (GL) string format (23,24) in the workshop database. Additional information regarding the output HML and XML files generated by the different NGS vendor systems, and upload into the 17th IHIW database has recently been reported by Jung C *et al.* (25).

2.5. Data checking by knowledge of HLA haplotypes

Submitted genotype data was interrogated for the well-known *HLA-DRB3/4/5* and *HLA-DRB1* haplotypes (26). General patterns of *HLA-DRB3/4/5~HLA-DRB1* linkage exist for specific allele families: *HLA-DRB3* alleles associates with *HLA-DRB1*03, 11, 12, 13, 14*; *HLA-DRB4* with *HLA-DRB1*04, 07, 09*; *HLA-DRB5* with *HLA-DRB1*15, 16*; and absence of *HLA-DRB3/4/5* alleles for *HLA-DRB1*01, 08, and 10* alleles. These associations have been well-characterized particularly in specific ethnic groups such as individuals of European ancestry. However, exceptions do exist particularly in non-European populations such as *HLA-DRB5*01~HLA-DRB1*01* and *HLA-DRB5*absent~HLA-DRB1*15* haplotypes in African descent groups. Data that did not conform to the general pattern of *HLA-DRB3/4/5~HLA-DRB1* linkage were further confirmed by checking the ethnic background of the cell line as well as confirming the alleles with the genotypes submitted by the other participating groups.

2.6. Concordance Calculation

Concordance calculations and data processing were performed using scripts written using R (v3.3.2) programming (<http://cran.r-project.org>). Due to incomplete and low-resolution reference HLA genotypes, concordance rates were calculated by comparing NGS HLA genotypes from each individual laboratory with consensus genotypes generated from at least four laboratories. Overall HLA concordance was scored as the number of shared class I and II alleles. Concordance of test NGS HLA genotypes with consensus reference NGS HLA typing was calculated at 2- and 4-field allele resolution using the following four classifications:

- i. Match of unambiguous reference allele + unambiguous NGS allele (referred to as Identical)
- ii. Ambiguous reference alleles + ambiguous NGS allele, matched at one allele (denoted Con-2)
- iii. Ambiguous reference alleles + unambiguous NGS alleles, matched at one allele (denoted Con-3)
- iv. Unambiguous reference allele + ambiguous NGS alleles, matched at one allele (denoted Con-4)

No match between the reference consensus alleles and NGS test alleles was scored as discordant.

3. Results

Overall 17 groups located worldwide participated in the reference panel single-blind study. One group had previously typed the majority of the cell lines and contributed the HLA genotype data to this study to help derive consensus genotypes. The IHWG and 17th IHIW identification numbers of these cell-lines were known to this group therefore they were not evaluated as part of the single-blind study. To maintain confidentiality the 17 groups participating in the single-blind study were assigned arbitrary numbers from 1 to 17 for the following sections of this report. Individual laboratories shown in Table 1 were randomized

so that there is no direct correlation between the arbitrary numbers assigned to each group and the order of the groups listed in Table 1. Genotype and sequence consensus data for 11 HLA loci (*HLA-A*, *HLA-C*, *HLA-B*, *HLA-DRB1*, *HLA-DRB3*, *HLA-DRB4*, *HLA-DRB5*, *HLA-DQA1*, *HLA-DQB1*, *HLA-DPA1*, and *HLA-DPB1*) were generated from IHWG B-LCLs samples using a variety of NGS methods. Forty-three samples included in the 346 ExtU panel had discrepant genotype calls when compared to the genotypes in the IPD-IMGT/HLA Database Cell Repository and we deduced this was due to sample mix-ups. Of the 43 discrepant samples, 6 samples could be corrected as they were previously typed by NGS methods by several testing laboratories in both the quality control pilot and proficiency testing studies. The 37 samples that were a result of preventable sampling handling errors, which could be traced back to the center distributing the cell lines, were removed prior to performing further analyses. Therefore 309 cells remained with genotypes we could be confident of in the ExtU panel. However of the 309 cells, 23 samples were also tested in the 24U4 panel, resulting in 382 unique cell lines (96 cells from the 24U and 286 cells from the ExtU panel) with valid HLA genotype data tested in the entire study.

3.1. Evaluation of Concordance

3.1.1. 24U panel study—Sixteen laboratories contributed HLA data to the 24U study but the number of panels typed by each laboratory varied; panel 24U1 was typed by 10 groups, panel 24U2 typed by 8, panel 24U3 by 7, and panel 24U4 by 6 laboratories. The number of HLA loci typed for each of the 24U panels ranged from 3 (all class I genes) to 11, however each panel was typed by at least 4 groups for all 11 loci. Overall concordance across all HLA loci was examined at 2- and 4-field allele level resolution for each laboratory that contributed data to the 24U study; the results are shown in Table 3a. The average concordance rates between genotypes generated by test laboratories and consensus genotypes were 99.0% (90.4 – 100.0%) at the 2-field allele resolution and 97.8% (89.6 – 99.9%) at the 4-field allele resolution.

For each HLA gene, concordance rates calculated at 4-field resolution ranged from: 91.7% to 100% for all class I genes; 89.6 – 100% for *HLA-DRB1*; 84.2 – 100% *HLA-DRB3*; 86.4 – 100% *HLA-DRB4*; 91.7 – 100%; *HLA-DQA1*; 83.3 – 100% *HLA-DQB1*; 87.5 – 100% *HLA-DPA1*; 80 – 100% for *HLA-DPB1*. All *HLA-DRB5* alleles typed by the groups were perfectly matched with the consensus *HLA-DRB5* alleles. As expected concordance rates for each HLA gene across all participating groups were improved although not dramatically when examined at 2-field resolution; concordance was highest for class I loci (91.7 to 100%) and lower for class II loci (88.9 – 100%). HLA gene concordance rates are shown in Tables 3b. Discrepant allele types between the test laboratory and the consensus genotypes could be ascribed to a number of reasons, such as sample mix-ups by the testing site; there was a single sample swap between H00005B3 (IHIW9041, J0528239) and H00005E4 (IHW09106, MANIKA) by laboratory #15 which were clearly evident from the very different HLA allele groups of these cell lines. Disparities were also due to failure of testing laboratories to call homozygous calls correctly, incorrect heterozygous genotypes, and limitations of the NGS HLA software to correctly assign alleles automatically at 4-field allele resolution.

3.1.2. Extended panel study—Five laboratories (1, 2, 16, 17, and 18) typed the samples included in the extended panel. Each group typed 11 HLA loci at full-gene coverage for class I genes as well as some class II genes and wide-coverage for the remaining class II genes. The average overall concordance rates computed at 2-field were 99.8% (99.1 – 100%) and at 4-field 98.0% (96.7 – 99.9%). The results are summarized in Table 4a.

For the class I genes concordance ranged from 98.9 – 100% and 92.9 – 100% at 2- and 4-field allele resolution levels respectively (Table 4b). At the 4-field level the following concordance ranges were observed for the class II genes; *HLA-DRB1* 99.4 – 100%, *HLA-DRB3* 93.6 – 100%, *HLA-DRB4* 97.6 – 100%, *HLA-DRB5* 87.1 – 100%, *HLA-DQA1* 95.4 – 100%, *HLA-DQB1* 95.9 – 100%, *HLA-DPA1* 95.8 – 100%, and *HLA-DPB1* 86.1 – 100%. At the 2-field level concordance rates for all class II genes fared better than the 4-field concordance rates. The lowest observation was at *HLA-DRB5* (95.8%) for group 18, which was due to an over-call of *HLA-DRB5*01:01:01* in 3 samples.

3.2. Characteristics of 383 B-LCLs characterized by NGS

A comprehensive description of the characteristics for the cell lines included in the 24U and extended panel studies such as cell names, ethnic group, consanguineous parents, and consensus genotypes generated by NGS from this study are shown in Supplementary Table 1. In total 470 alleles were identified of which 11 alleles were non-expressed; *HLA-A*01:04N*, *HLA-A*02:53N*, *HLA-A*23:08N*, *HLA-A*24:09N*, *HLA-A*68:11N*, *HLA-A*68:18N*, *HLA-C*04:09*, *HLA-B*39:25N*, *HLA-B*44:23N*, *HLA-DRB4*01:03:01:02N*, and *HLA-DRB5*01:08N*. All alleles characterized at *HLA-A*, *HLA-C*, *HLA-B*, *HLA-DPA1*, and *HLA-DRB5* loci were non-ambiguous. For the remaining loci, both non-ambiguous and ambiguous 4-field alleles were detected which were mostly due to indiscernible short tandem repeat (STR) enriched regions located within introns of some class II genes. For example alleles *HLA-DRB1*15:01:01:01*, *HLA-DRB1*15:01:01:02*, and *HLA-DRB1*15:01:01:03* are indistinguishable due to GT and GA dinucleotide repeats located at ~5700–5790 bp within intron 2. Phase ambiguities, due to low polymorphic variation across intron 2 of the *HLA-DPB1* gene, accounted for 27 % of the total (n = 59) *HLA-DPB1* alleles detected.

We detected a genotype ambiguity in three cell lines (IHW09267 LEO23, IHW09289 LB, and IHW09374 FH2) at the *HLA-DQB1* locus due to incomplete reference sequences; the *HLA-DQB1*06:03:01+HLA-DQB1*06:04:01* genotype is indistinguishable from *HLA-DQB1*06:39+HLA-DQB1*06:41*. This ambiguity occurs because the genomic sequence is only available for *HLA-DQB1*06:03:01+HLA-DQB1*06:04:01*. Also various combinations of alleles from the two possible genotypes share identical sequence across exons 2 and 3. For instance the following alleles have identical sequences across exon 2; *HLA-DQB1*06:03:01+HLA-DQB1*06:41* and *HLA-DQB1*06:39+HLA-DQB1*06:04:01*. Whereas *HLA-DQB1*06:03:01+HLA-DQB1*06:39* and *HLA-DQB1*06:41+HLA-DQB1*06:04:01* share identical sequence across exon 3. However, we can deduce the plausible genotypes by examining the extended haplotypes. We note that cells LEO23, LB, and FH2 all bear the *HLA-DRB1*13:01:01~HLA-DRB3*03:01:01~HLA-DQB1*06:03:01/HLA-DQB1*06:39~HLA-DQA1*01:03:01* haplotype. In comparison, the IPD-IMGT/HLA

Database cell repository shows *HLA-DQB1*06:39* reported in cell LUMC-DQB64 bearing *HLA-DRB1*15:01* but lacking *HLA-DRB1*13:01* or *HLA-DRB1*13:02*. Therefore we conclude that cells LB, LEO23, and FH2 carry the common genotype *HLA-DQB1*06:03:01+HLA-DQB1*06:04:01*. A summary of the unique alleles found in the cell panel is shown in Supplementary Table 2.

3.3. Homozygous genotypes

In the entire cohort of 382 B-LCLs, 51 cells were reported as consanguineous by descent, 274 samples were not consanguineous and for 57 samples the status of consanguinity was unknown. Of the total 51 consanguineous B-LCLs, 41 cells were homozygous for all alleles defined at maximum allele resolution at all 11 HLA loci. In addition, homozygous genotypes were also identified in other non-consanguineous cell lines at some or all 11 loci; 37 cells were homozygous at all 11 loci. The total number of homozygous genotypes detected at class I genes were 150 at *HLA-A*, 143 for *HLA-C*, and 141 for *HLA-B*. At all class II genes examined homozygous genotypes ranged from 80 (*HLA-DRB5*) to 197 (*HLA-DRB3*). The results are summarized in Supplementary Table 3.

3.4. Detection of novel variants

NGS data was analysed using the multiple sequence alignment tool known as hlaPoly a program developed specifically for the 17th IHIW to identify novel nucleotide polymorphisms for a given HLA consensus sequence (25). Thirty-three samples in the entire B-LCL panel contained novel alleles, which are HLA alleles that are not present in version 3.25.0 of the IPD-IMGT/HLA Database. Eleven exon variants were detected in 18 samples. In cell line C1R (IHW09208), which was included in the extended panel study, 2 exon variants were detected; an exon 3 deletion of nucleotide G at codon 179 causing a frameshift mutation in allele *HLA-A*02:01:01:01* and a transversion mutation that changes the translation initiation codon of *HLA-B*35:03:01* from ATG to TTG at codon -24 in exon 1. Examination of the consensus sequence data revealed that 4 of the 5 groups included in the extended panel study detected these exon variants; for 1 group the sample failed typing. The *HLA-A*, *-B* negative mutant C1R cell line is well-described and is widely used in functional studies of class I MHC genes. C1R was derived from a normal EBV-transformed B cell line, Hmy2, by three rounds of γ -irradiation followed by immunoselection with monoclonal anti-HLA class I antibodies and complement. Mutagenesis resulted in one entire HLA haplotype being deleted in the C1R cell line (27,28).

Novel variants were also characterized at other class I alleles. *HLA-A*33:03:01* in cell line V.E.C (IHW09314) harbors a point mutation within codon -20 exon 1 resulting in an alanine to valine amino acid change. Also in two samples, HIN, JP (IHW09163) and HIN, S (IHW09165) a point mutation located within exon 4 of *HLA-C*14:02:01:01* was detected.

The majority of exon variants were detected in the *HLA-DPA1*02:02* alleles. We characterized three substitution variants (one missense and two silent) in exon 3 of *HLA-DPA1*02:02:01* for 8 samples, and for all samples the new nomenclature for this allele, according to IPD-IMGT/HLA Database version 3.30.0 is *HLA-DPA1*02:07:01:01*. Other exonic variants of *HLA-DPA1*02:02* were detected in cell FH9 IHW09383; *HLA-*

*DPA1*02:02:02* exon 3 variant, codon 114, Pro to Leu, the most similar allele is an exon 3 variant of *HLA-DPA1*02:02:01*. Also cell DUG150 IHW09397 harbored three exon 3 variants and one missense exon 4 variant in *HLA-DPA1*02:02:01*; the closest allele is an exon 4 variant of *HLA-DPA1*02:07:01:01*. A novel *HLA-DRB1*14:08* exon 3 variant (missense mutation) was detected in sample WON, I (IHW09194) derived from an Australian Aboriginal donor. All complete sequences containing novel exon variants were submitted to GenBank (<https://www.ncbi.nlm.nih.gov/genbank>). The submitted sequences have the prefix MK in the GenBank accession numbers listed in Table 5.

We detected variants located within introns 1 and 2 of the *HLA-DQB1*03:01:01* allele that appeared to be a hybrid of *HLA-DQB1*03:01:01:01* recombined with *HLA-DQB1*03:01:01:03*. Interestingly these *HLA-DQB1*03:01:01:01/HLA-DQB1*03:01:01:03* 'hybrid' alleles were found in cell lines of Asian descent. Three different hybrid alleles were detected and were associated with two distinct *HLA-DRB1* alleles; *HLA-DRB1*08:03:02* and *HLA-DRB1*12:02:01*:

*HLA-DRB1*12:02:01~HLA-DQB1*03:01:01:01/HLA-DQB1*03:01:01:03_Hybridx1~HLA-B*15/B*27/B*35/B*46*

*HLA-DRB1*12:02:01~HLA-DQB1*03:01:01:01/HLA-DQB1*03:01:01:03_Hybridx2~HLA-B*15/B*55*

*HLA-DRB1*08:03:02~HLA-DQB1*03:01:01:01/HLA-DQB1*03:01:01:03_Hybridx3~HLA-B*51:01:01:01*

Recently these novel *HLA-DQB1*03:01:01:01/HLA-DQB1*03:01:01:03* hybrid alleles have been submitted to the IPD-IMGT/HLA Database by other groups and have been named; *HLA-DQB1*03:01:01:07* (Hybridx1), *HLA-DQB1*03:01:01:12* (Hybridx2), *HLA-DQB1*03:01:01:17* (Hybridx3). Details of all novel alleles characterized in this study are summarized in Table 5.

3.5. Patterns of association

A large majority of the cells tested were derived from Caucasoid donors (42.1%, n = 161), the second largest represented group were from Asia (17.5%, n = 67), followed by African ancestry (5.8%, n = 22), Native Americans (3.1%, n = 12), Hispanic (2.4%, n = 9), Jewish (2.4%, n = 9) and Australian Aboriginal (1.6%, n = 6). Ethnic groups from North Africa, India and the Pacific Islands collectively represented 1.6% of the cell panel and for 23.6% the ancestry was unknown. Although Caucasoid cell donors represented the majority of the cell panel we had sufficient numbers of samples in the other broad ethnic categories to observe distinct haplotypes, due to linkage disequilibrium (LD) across the MHC, at maximum resolution. In the Caucasoid group the following haplotypes were frequently observed at the *HLA-C* and *HLA-B* loci; *HLA-C*07:02:01:03~HLA-B*07:02:01*, *HLA-C*07:01:01:01~HLA-B*08:01:01:01*; *HLA-C*05:01:01:01~HLA-B*18:01:01:01*; *HLA-C*05:01:01:02~HLA-B*44:02:01:01*

At the *HLA-DPA1* locus, if *HLA-DPA1*01:03:01:05* was detected it was exclusively associated with *HLA-DPB1*04:02:01:01* or *HLA-DPB1*04:02:01:02*, regardless of the

ethnic group. Whereas *HLA-DPA1*01:03:01:01* associates with *HLA-DPB1*02:01:02*, *HLA-DPA1*01:03:01:02* with *HLA-DPB1*04:01:01:01* and *HLA-DPB1*04:01:01:02* across all ethnic groups. In non-European groups we observed the following haplotypes *HLA-DPA1*02:02:02~HLA-DPB1*01:01:01* (African and Asian) *HLA-DPA1*02:02:02~HLA-DPB1*05:01:01* (Asian and Australian Aboriginal).

Distinct extended haplotypes encompassing *HLA-DQA1*, *HLA-DQB1* and *HLA-DRB1*^{3/4/5} loci were apparent in specific ethnic groups for example; *HLA-DRB5*01:02~HLA-DRB1*15:02:01:01/HLA-DRB1*15:02:01:02~HLA-DQB1*06:01:01~HLA-DQA1*01:03:01:01* (Asian, Caucasoid, Iranian) and *HLA-DRB1*08:03:02~HLA-DQB1*06:01:01~HLA-DQA1*01:03:01:03/HLA-DQA1*01:03:01:04* (Asian and Australian Aboriginal). These haplotypes illustrate how NGS extends and improves the accuracy of HLA allelic and haplotype diversity in populations. Further examples of haplotypes observed in the cell panel are shown in Table 6.

3.6. Evaluation of HLA NGS genotypes with genotypes listed in the IPD-IMGT/HLA Database Cell Repository

We compared the consensus NGS HLA genotype data to the historical genotype data collated in the IPD-IMGT/HLA Database Cell Repository which were characterized mostly by low and medium HLA resolution molecular methods. Overall 68.4% (n = 4630) of the NGS HLA alleles matched with the corresponding alleles listed in the IPD-IMGT/HLA Database Cell Repository. Of this total, 51.0% (n = 3453) NGS HLA alleles matched at an increased resolution, 12.2% (n = 828) were at the same resolution, and for 2.5% (n = 170) the resolution was decreased. The latter result could be attributed to ambiguous class II alleles that could not be distinguished by NGS methods due to low complexity intronic regions.

In general alleles found to be ambiguous by NGS were not listed as ambiguous alleles in the IPD-IMGT/HLA Database Cell Repository. An example is shown for cell line 26/27 (IHW09251), for which the IPD-IMGT/HLA Database cell report is *HLA-DPB1*02:01:02+HLA-DPB1*04:02*. By NGS, two genotypes are equally possible; *HLA-DPB1*02:01:02+HLA-DPB1*04:02:01:02* and *HLA-DPB1*105:01 HLA-DPB1*416:01*. Alleles *HLA-DPB1*02:01:02* and *HLA-DPB1*416:01* share identical sequence across exon 2, whilst *HLA-DPB1*04:02:01:02* and *HLA-DPB1*105:01* have identical sequence over exon 3. Such discrepancies are likely to be the result of the date of the original submission of the cell information to the IPD-IMGT/HLA Database and naming of the alleles at that time; a cell could have been submitted with an unambiguous type before the discovery of a novel allele that created an ambiguity.

For the remaining proportion of matched alleles zygosity discrepancies were observed at various levels of resolution; increased resolution at 2.4% (n = 162) and 0.3% (n = 17) were matched at the same resolution. Zygosity discrepancies denote alleles that differ in the number of copies at a given locus. It was observed that the IPD-IMGT/HLA cell reports would sometimes list two identical alleles for a given locus but by NGS the corresponding alleles were defined as either heterozygous or a single allele. Such instances were observed at the *HLA-DQB1* gene for cell 2012 (IHW09347): NGS genotype *HLA-*

*DQB1*05:03:01:01/HLA-DQB1*05:03:01:02 +HLA-DQB1*06:03:01* compared to the IPD-IMGT/HLA Database genotype *HLA-DQB1*06:03:01*.

Discordant alleles (6.2%, n = 421) between the two datasets were due to the greater accuracy and resolution of the NGS methods compared to older lower resolution molecular and serological typing methods. In addition, due to the dynamic nature of the HLA system, the number of HLA alleles reported has increased over the years, for example in 2002 during the 13th IHIW *HLA-DRB1*14:54:01* had not been reported because exon 3 of the *HLA-DRB1* gene was not routinely sequenced, for this reason *HLA-DRB1*14:01* was reported as unambiguous. In addition, some alleles have undergone name changes, for instance allele *HLA-C*02:10:01:01* characterized by NGS is listed as *HLA-C*02:02* or *HLA-C*02:02:04* in the IPD-IMGT/HLA Database Cell Repository. These examples illustrate the importance of storing and re-examining primary sequence data rather than solely relying on reported genotype calls. A comparison of the various categories of matched and discrepant alleles between NGS data and old typing results are illustrated in Figure 1.

4. Discussion

In this study we have successfully broadened the molecular characterization of reference IHWG B-LCLs through high-resolution typing of classical class I and class II HLA genes using various NGS methods. In this endeavor, we have also resolved inconsistencies found when comparing NGS genotypes to historical typing data generated using lower-resolution typing techniques and added new allelic information as the majority of cell lines collated in both the IPD-IMGT/HLA and the IHWG Cell databases lacked genotype information for some of the HLA genes. Our multi-center study, which is the largest comprehensive analyses of reference cells carried out to-date, provides a reference source for alleles and common haplotypes defined at maximum allele resolution in diverse ethnic groups. The reference B-LCLs are an important resource for the Immunogenetics community, an observation further underscored by the many studies reported over the years that have utilized these cells for different purposes such as quality control and optimization of novel typing methods for HLA (11,12), and KIR (13), defining other immunogenetic markers (29,30,31) as well as MHC extended haplotype studies (32,33). Recently, Turner *et al.* applied a ‘Third generation’ NGS system, Pacific Biosciences Single Molecule Real-Time (SMRT) DNA sequencing, to sequence 126 IHW reference cells for class I loci (full genomic sequencing) and partial sequencing for *HLA-DRB1*, *HLA-DQB1*, and *HLA-DPB1* loci (12). There was some overlap of the cells typed in this study and those characterized by the aforementioned study. Our results have confirmed and extended the HLA genotypes generated by the Turner study.

In this multi-center study, through the application of NGS combined with long range PCR, we have successfully characterized eleven HLA loci with minimal ambiguity. The HLA ambiguities we observed, due to low complexity regions and *HLA-DPB1* phase ambiguities could not be resolved and appeared to be independent of the NGS methods and platforms used. We also noted differences in the level of resolution of particular class I and II alleles generated by some laboratories due to differences in genomic coverage. For instance in the extended cell line panel study, 4 of the 5 groups could not discern *HLA-B*35:01:01:01* from *HLA-B*35:01:01:02*. The difference between these two alleles according to the IPD-

IMGT/HLA Database v3.25.0, is a single nucleotide difference located within the 3'UTR at position 2936 bp; *HLA-B*35:01:01:01* carries a G nucleotide whilst *HLA-B*35:01:01:01* bears an A nucleotide. Further evidence for the correct *HLA-B*35:01* 4-field allele is provided by the distinct associations observed in larger population studies of HLA diversity. For instance in a large cohort of European ancestry individuals the following *HLA-C*04~HLA-B*35:01:01:02* associations were observed; *HLA-C*04:01:01:01~HLA-B*35:01:01:02* (frequency = 0.034), *HLA-C*04:01:01:05~HLA-B*35:01:01:02* (frequency = 0.015), *HLA-C*04:01:01:06~HLA-B*35:01:01:02* (frequency = 0.007). In comparison *HLA-B*35:01:01:01* associates with *HLA-C*04:01:01:01* and is far less common with an estimated haplotype frequency of 0.0005 (Creary *et al.* unpublished data).

In addition to inadequate genomic coverage, inconsistent alleles were reported due to poor critique of NGS HLA genotypes and lack of interrogation of the associated sequence data. It is important to emphasize that the accuracy of NGS for HLA typing, especially for detecting alleles characterized at 4-field, depends upon both the NGS protocol (reagents and sequencing platform) and manual review of the NGS data. An important observation from this study is that some of the inconsistent alleles reported could have been prevented if care was taken to manually review automatic calls generated by the software. These findings highlight how errors may be avoided in future studies examining NGS HLA data.

Our results illustrate how NGS of extended HLA amplicons allow for accurate characterization of genomic regions that were difficult to characterize using traditional methods or not sequenced at all. For instance in the case of non-expressed HLA alleles, which are often due to indels causing a frameshift mutation resulting in a premature stop codon or single point mutations causing a premature stop codon, NGS identified these variants accurately and efficiently. The identification of null alleles is important for HSCT as nonidentification or misidentification could lead to poor engraftment and graft-versus-host disease. Furthermore, these variants are sometimes located outside of the antigen-recognition site as observed in alleles *HLA-A*68:11N* (due to a deletion of the G nucleotide in exon 1 codon -9), *HLA-C*04:09N* (deletion of A nucleotide in exon 7 codon 341), and *HLA-DRB4*01:03:01:02N* (G>A substitution position 9656 intron 1) suggesting that all genomic regions should be sequenced to ensure detection of all null alleles.

Extended coverage and sequencing of intronic and untranslated regions has revealed silent substitutions that have distinct haplotype associations indicating higher levels of linkage disequilibrium across the HLA regions than previously estimated using lower resolution typing data. Increased haplotype diversity was observed in the HLA class II region. For example, there are three 4-field subtypes of *HLA-DQA1*01:03* that associate with different *HLA-DRB1* alleles in different ethnic groups; *HLA-DQA1*01:03:01:01~HLA-DQB1*06:01:01~HLA-DRB1*15:02:01:01/02~HLA-DRB5*01:02* (Asian, Gypsy, Middle Eastern), *HLA-DQA1*01:03:01:02/06~HLA-DQB1*06:03:01~HLA-DRB1*13:01:01:01/02~HLA-DRB3*01:01:02:01/02* (European, Asian, African, Australian Aboriginal), and *HLA-DQA1*01:03:01:03/04~HLA-DQB1*06:01:01~HLA-DRB1*08:03:02* (Australian Aboriginal, Japanese). In addition we observed diversity at the 4-field level of the *HLA-DQA1*03:03* alleles; *HLA-DQA1*03:03:01:01~HLA-DQB1*03:01:01:01~HLA-DRB1*04:01:01:01/02* (European, African American, Australian

Aboriginal, Hispanic), *HLA-DQA1*03:03:01:02~HLA-DQB1*04:02:01~HLA-DRB1*04:05:01/HLA-DRB1*04:10:01* (Asian), *HLA-DQA1*03:03:01:03~HLA-DQB1*04:01:01~HLA-DRB1*04:05:01* (Asian, Hispanic). These haplotypes illustrate how variants in introns could be useful markers of population differences.

In conclusion the B-LCL panel highly defined by NGS is a valuable resource; with the expansion of sequence coverage in both flanking and untested regions novel polymorphisms may lead to name changes. The 18th IHIW will focus on studies evaluating the untested regions as well as other loci in the HLA region. The improvement of algorithms used in NGS HLA assignment software, paired with longer sequence reads and expanded reference databases will achieve faster non-ambiguous sequence assignments. The data generated from this study will be used to update the IPD-IMGT/HLA and IHWG cell repositories and will provide a valuable resource for developing assays and characterizing haplotypes.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgements

We gratefully acknowledge the Fred Hutchinson Cancer Research center for providing the IHWG DNA samples, in particular Mr. Emil Madraimov and Ms. Angela Bryce for distributing samples to participating laboratories. We are also immensely grateful to the many vendors that provided reagents and software free of charge to some of the investigators. We extend thanks to the Stanford Blood center for their financial and general administrative support for the 17th IHIW. LEC and MFV (Stanford University) were supported by grant U19NS095774 from the U.S. National Institutes of Health (NIH). MP (Palacky University) was supported by RVO: 61989592. SAS (Kuwait University) was supported by grant 2012-130-204 from the Kuwait Foundation for Advancement of Science (KFAS). IR (McGill University) was supported by the Genome Canada Science Technology Innovation Center, Compute Canada Resource Allocation Project (WST-164-AB) and Genome Innovation Node (244819).

Abbreviations:

HLA	Human leukocyte antigen
B-LCL	B-lymphoblastoid cell lines
NGS	Next-generation sequencing
IHIW	International HLA and Immunogenetics Workshop
IHWG	International Histocompatibility Working Group

References

1. Yang SY, Milford E, Hammerling U, Dupont B. Description of the reference panel of B-lymphoblastoid cell lines for factors of the HLA system: The B-cell line panel designated for the Tenth International Histocompatibility Workshop In Dupont B, editor. Immunobiology of HLA, Volume I, Histocompatibility Testing 1987. New York: Springer-Verlag; 1989 p. 11–19.
2. Milford EL, Kennedy LJ, Yang SY, Dupont B, Lalouel JM, Yunis EJ. Serologic Characterization of the Reference Panel of B-Lymphoblastoid Cell Lines for Factors of the HLA system In Dupont B, editor. Immunobiology of HLA Volume I, Histocompatibility Testing 1987. New York: Springer-Verlag; 1989 p. 19–38.

3. Mickleson E, Reinsmoen N, Robbins FM, Hartzman R, Odum N, Svejgaard A, et al. HLA-Dw and HLA-DP Typing of the Reference Panel of B-Lymphoblastoid Cell Lines In Dupont B, editor. Immunobiology of HLA Volume I, Histocompatibility Testing 1987. New York: Springer-Verlag; 1989 p. 38–42.
4. Yang SY. Assignment of HLA-A and HLA-B Antigens for the Reference Panel of B-Lymphoblastoid Cell Lines Determined by One-Dimensional Isoelectric Focusing (1D-IEF) Gel Electrophoresis In Dupont B, editor. Immunobiology of HLA Volume I, Histocompatibility Testing 1987. New York: Springer-Verlag; 1989 p. 43–44.
5. Knowles RW. Assignment of HLA-Class II alpha and beta Chain 2-D Gel Patterns for the Workshop Reference Panel of B-Lymphoblastoid Cell Lines In Dupont B, editor. Immunobiology of HLA Volume I, Histocompatibility Testing 1987. New York: Springer-Verlag; 1989 p. 44–46.
6. Flomenberg N Assignment of T-Cell_Defined (TCD) HLA Class II Specificities for the Reference Panel of B-Lymphoblastoid Cell Lines In Dupont B, editor. Immunobiology of HLA Volume I, Histocompatibility Testing 1987. New York: Springer-Verlag; 1989 p. 46–48.
7. Kimura A, Dong RP, Harada H, Sasazuki T. DNA typing of HLA class II genes in B-lymphoblastoid cell lines homozygous for HLA. In Tsuji K, Aizawa M, Sasazuki T, editors. HLA 1991, Proceedings of the Eleventh International Histocompatibility Workshop and Conference held in Yokohama, Japan, 6–13 November 1991 New York: Oxford University Press; 1992 p. 419–425.
8. Marsh SGE, Packer R, Heyes JM, Bolton B, Faucher R, Charron D, et al. The 12th International Histocompatibility Workshop cell lines panel In Charron D, editor. Genetic diversity of HLA Functional and Medical Implication. Paris: EDK; 1997 p. 26–27.
9. Mickleson E, Rosielle J, Pei J, Hansen JA. Quality control and reference cells for DNA typing: The International Human MHC Cell and Gene Bank. In Hansen JA, editor. Immunobiology of the Human MHC, Proceedings of the 13th International Histocompatibility Workshop and Conference, Volume I Denmark: PJ Schmidt; 2006 p. 417–420.
10. Hansen JA. Fred Hutchinson Cancer Research Center. [Online].; 2018 [cited 2018 June 29]. Available from: <https://www.fredhutch.org/en/labs/clinical/projects/iHW/reference-panels.html>.
11. Mayor NP, Robinson J, McWhinnie AJM, Ranade S, Eng K, Midwinter W, et al. HLA Typing for the Next Generation. PLOS ONE. 2015 5; 10(e0127153).
12. Turner TR, Hayhurst JD, Hayward DR, Bultitude WP, Barker DJ, Robinson J, et al. Single molecule real-time DNA sequencing of HLA genes at ultra-high resolution from 126 International HLA and Immunogenetics Workshop cell lines. HLA. 2018 2; 91(2): p. 88–101. [PubMed: 29171935]
13. Norman PJ, Hollenbach JA, Gorgani-Nemat N, Marin WM, Norberg SJ, Ashouri E, et al. Defining KIR and HLA Class I Genotypes at Highest Resolution via High-Throughput Sequencing. The American journal of Human Genetics. 2016 8; 99(2): p. 375–91. [PubMed: 27486779]
14. Cao K, Chopek M, Fernandez-Vina MA. High and intermediate resolution DNA typing systems for class I HLA-A, B, C genes by hybridization with sequence-specific oligonucleotide probes (SSOP). Reviews in Immunogenetics. 1999; 1(2): p. 177–208. [PubMed: 11253946]
15. Bunce M, O'Neill CM, Barnado MC, Krausa P, Browning MJ, Morris PJ, et al. Photocopying: comprehensive DNA typing for HLA-A, B, C, DRB1, DRB3, DRB4, DRB5 & DQB1 by PCR with 144 primer mixes utilizing sequence-specific primers (PCR-SSP). Tissue Antigens. 1995 11; 46(5): p. 355–67. [PubMed: 8838344]
16. Cereb N, Maye P, Lee Y, Kong S, 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 1; 45(1): p. 1–11. [PubMed: 7725305]
17. Kotsch K, Wehling J, Kohler S, Blasczyk R. Sequencing of HLA class I genes based on the conserved diversity of the non-coding regions: sequence-based typing of the HLA-A gene. Tissue Antigens. 1997 8; 50(2): p. 178–91. [PubMed: 9271828]
18. Kotsch K, Wehling J, Blasczyk R. Sequencing of HLA class II genes based on the conserved diversity of the non-coding regions: sequencing based typing of HLA-DRB genes. Tissue Antigens. 1999 5; 53(5): p. 486–97. [PubMed: 10372544]
19. Erlich H HLA DNA typing: past, present, and future. Tissue Antigens. 2012 7; 80(1): p. 1–11. [PubMed: 22651253]

20. Middleton D, Marsh SGE. 16th International HLA and Immunogenetics Workshop (IHIW) Introduction. *International Journal of Immunogenetics*. 2013 2; 40(1): p. 1. [PubMed: 23280276]
21. Robinson J, Halliwell JA, Hayhurst JD, Flicek P, Parham P, Marsh SGE. The IPD and IMGT/HLA database: allele variant databases. *Nucleic Acids Research*. 2015 1; 43(D423-D431).
22. Mack SJ, Cano P, Hollenbach JA, He J, Hurley CK, Middleton D, et al. Common and Well-Documented HLA Alleles: 2012 Update to the CWD Catalogue. *Tissue Antigens*. 2013 4; 81(4): p. 194–203. [PubMed: 23510415]
23. Milius RP, Mack SJ, Hollenbach JA, Pollack J, Heuer M, Gragert L, et al. Genotype List String: a grammar for describing HLA and KIR genotyping results in a text string. *Tissue Antigens*. 2013 8; 82(2): p. 106–12. [PubMed: 23849068]
24. Milius RP, Heuer M, George M, Pollack J, Hollenbach JA, Mack SJ, et al. The GL Service: Web Service to exchange GL String encoded HLA & KIR genotypes with complete and accurate allele and genotype ambiguity. *Human Immunology*. 2016 3; 77(3): p. 249–56. [PubMed: 26621609]
25. Chang CJ, Osoegawa K, Milius RP, Maiers M, Xiao W, Fernandez-Vina MA, et al. Collection and storage of HLA NGS genotyping data for the 17th International HLA and Immunogenetics Workshop. *Human Immunology*. 2018 2; 79(2): p. 77–86. [PubMed: 29247682]
26. Andersson G Evolution of the human HLA-DR region. *Frontiers in Bioscience*. 1998 7; 3: p. d739–745. [PubMed: 9675159]
27. Storkus WJ, Howell DN, Salter RD, Dawson JR, Cresswell P. NK susceptibility varies inversely with target cell class I HLA antigen expression. *The Journal of Immunology*. 1987 3; 138(6): p. 1657–59. [PubMed: 3819393]
28. Zemmour J, Little AM, Schendel DJ, Parham P. The HLA-A,B “negative” mutant cell line C1R expresses a novel HLA-B35 allele, which also has a point mutation in the translation initiation codon. *The Journal of Immunology*. 1992 3; 148(6): p. 1941–48. [PubMed: 1541831]
29. Dunn DS, Naruse T, Inoko H, Kulski JK. The Association between HLA-A alleles and young Alu dimorphisms near the HLA-J, -H, and -F genes in workshop cell lines and Japanese and Australian populations. *Journal of Molecular Evolution*. 2002 12; 55(6): p. 718–26. [PubMed: 12486530]
30. Ugolotti E, Vanni I, Raso A, Benzi F, Malnati M, Biassoni R. Human leukocyte antigen-B (-Bw6/-Bw4 I80, T80) and human leukocyte antigen-C (-C1/-C2) subgrouping using pyrosequencing analysis. *Human Immunology*. 2011 5; 72: p. 859–68. [PubMed: 21664941]
31. Cox ST, Madrigal JA, Saudemont A. Diversity and characterization of polymorphic 5’ promoter haplotypes of MICA and MICB genes. *Tissue Antigens*. 2014 9; 84: p. 293–303. [PubMed: 24962621]
32. Dorak MT, Shao W, Machulla HKG, Lobashevsky ES, Tang J, Park MH, et al. Conserved extended haplotypes of the major histocompatibility complex: further characterization. *Genes and Immunity*. 2006 6; 7: p. 450–67. [PubMed: 16791278]
33. Norman PJ, Norberg S, Nemat-Gorgani N, Royce T, Hollenbach JA, Shults Won M, et al. Very long haplotype tracts characterized at high resolution from HLA homozygous cell lines. *Immunogenetics*. 2015 9; 67(9): p. 479–85. [PubMed: 26198775]

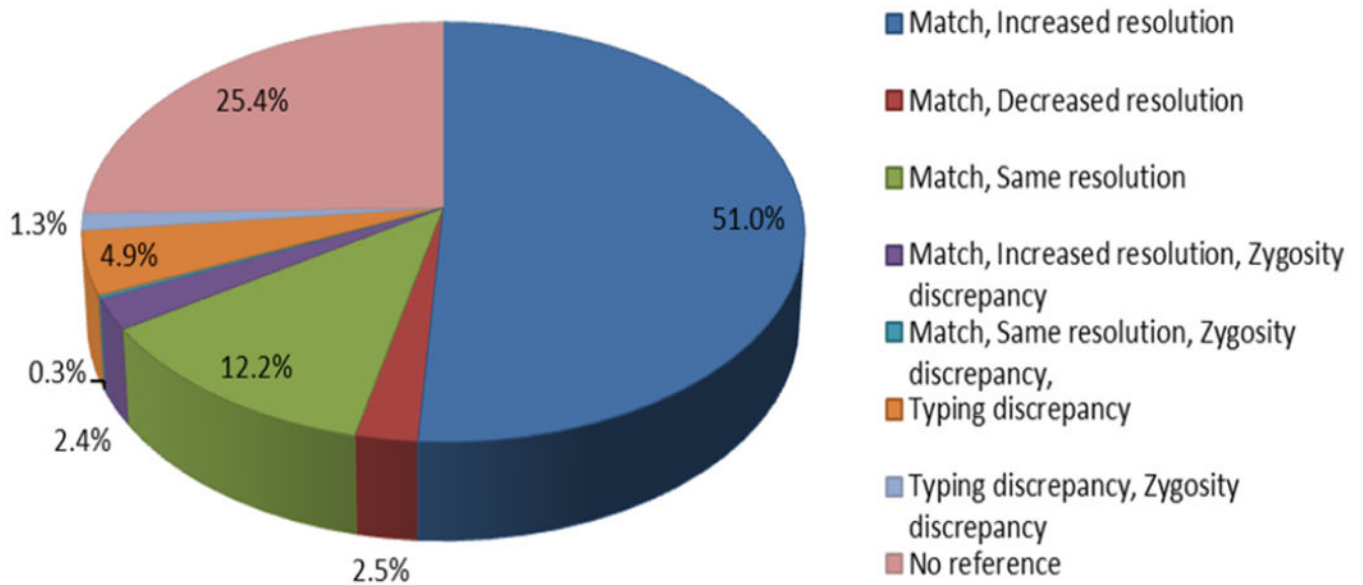


Figure 1.

Comparison of 382 B-LCL HLA alleles characterized by NGS with alleles catalogued in the IPD-IMGT/HLA Database Cell Repository. Various categories of NGS HLA alleles matched with corresponding alleles in the IPD-IMGT/HLA Database Cell Repository are shown; NGS HLA alleles matched with increased allelic resolution (dark blue slice), NGS HLA alleles matched with decreased allelic resolution (red slice), NGS HLA alleles matched with the same allelic resolution (green slice). Zygosity discrepancies describe alleles that differ in the number of copies at a given locus. NGS HLA alleles matched with increased allelic resolution but a zygosity discrepancy was detected (purple slice). NGS HLA alleles matched at the same allelic resolution but a zygosity discrepancy was detected (turquoise slice). Typing discrepancies (orange slice) denotes HLA alleles mismatched between the two datasets. The proportion of typing discrepancies detected in the presence of zygosity discrepancies is illustrated in the light-blue slice. NGS HLA alleles that do not have HLA allele data listed in the IPD-IMGT/HLA Database Cell Repository are denoted as no reference (pink slice).

Table 1.

Principal Investigators and associated laboratories included in the study

Labcode	PI(s) names	Laboratory
autfis	Gottfried Fischer	Medical University of Vienna, Austria
cansaw	Chee Loong Saw, Jiannis Ragoussis	McGill University Health Center and McGill University, Canada
czepet	Martin Petrek	IMTM, Palacky University, Czech Republic
espcs	Carles Serra-Pages, Manel Juan	Hospital Clinic de Barcelona, Spain
gbrcn	Colin J Brown	NHS Blood and Transplant, London, UK
gbrsm	Steven GE Marsh	Anthony Nolan Research Institute, London, UK
grcsta	Catherine Stavropoulos-Giokas, Amalia Dinou	Hellenic Cord Blood Bank, Athens, Greece
kwfame	Reem Ameen	Health Sciences Center, Kuwait University, Kuwait
nldspi	Eric Spierings	UMC Utrecht, Netherlands
ucagen	Ketevan Gendzekhadze,	City of Hope, CA, USA
ucamor	Gerald P Morris	UCSD Immunogenetics and Transplantation Laboratory, CA, USA
ucazha	Qiheng Zhang	UCLA Immunogenetics Center, CA, USA
uorkas	Zahra Kashi	Kashi Clinical Laboratories, OR, USA
upabsu	Susan Hsu	American Red Cross, PA, USA
ussta1	Marcelo A Fernández-Vi a	Stanford University, CA, USA
utxask	Medhat Askar	Baylor University Medical Center, TX, USA
ucasai	Katsuyuki Saito	One Lambda, Thermo Fisher, CA, USA
upamon	Dimitrios Monos	The Children's Hospital of Philadelphia, PA, USA

Table 2.

NGS systems and HLA Loci typed by participating laboratories

Labcode	Software	Hardware	HLA loci typed	Coverage class I/class II
autfis	TypeStream™ Visual, One Lambda-Thermo Fisher Scientific and NGSengine, GenDx, Inc	Ion Torrent™ PGM system	A, C, B, DRB1/3/4/5, DQA1, DQB1, DPA1, DPB1	Full-gene/wide-coverage
cansaw	HLA Twin, Omixon	MiSeq, Illumina	A, C, B, DRB1/3/4/5, DQA1, DQB1, DPA1, DPB1	Full-gene/full-gene (DQA1, DQB1, DPA1) and wide-coverage
czepet	HLA Twin, Omixon	MiSeq, Illumina	A, C, B, DRB1/3/4/5, DQA1, DQB1, DPA1	Full-gene/full-gene (DQA1, DQB1 and DPA1 full-gene) and wide-coverage
especs	NGSengine, GenDx	MiSeq, Illumina	A, C, B, DRB1, DQB1	Full-gene/wide-coverage
gbren	NGSengine, GenDx TypeStream™ Visual, One Lambda-Thermo Fisher Scientific	MiSeq, Illumina Ion S5™ system	A, C, B, DRB1, DPB1 A, C, B, DRB1/3/4/5, DQA1, DQB1, DPA1, DPB1	Full-gene/full-gene Full-gene/full-gene (DQA1, DPA1), and wide-coverage
gbrsm	SMRT Analysis, PacBio and In-house tools	PacBio RS II system	A, C, B, DRB1, DQB1, DPB1	Full-gene/wide-coverage
gresta	HLA Twin, Omixon	MiSeq, Illumina	A, C, B, DRB1/3/4/5, DQA1, DQB1, DPB1	Full-gene/wide-coverage
kwame	HLA Twin, Omixon	MiSeq, Illumina	A, C, B, DRB1/3/4/5, DQA1, DQB1, DPA1, DPB1	Full-gene/full-gene (DQA1, DQB1 and DPA1 full-gene) and wide-coverage
nldspi	NGSengine, GenDx	MiSeq, Illumina	A, C, B, DRB1, DQA1, DQB1, DPA1, DPB1	Full-gene/wide-coverage
ucamor	TruSight HLA Assign, Illumina	MiSeq, Illumina	A, C, B, DRB1/3/4/5, DQA1, DQB1, DPA1, DPB1	Full-gene (A, C), B (ex1-intron 6)/full-gene (DQB1, DQA1, DPA1), wide-coverage
ucazha	TypeStream™ Visual, One Lambda-Thermo Fisher Scientific	Ion S5™ XL system	A, C, B, DRB1/3/4/5, DQA1, DQB1, DPA1, DPB1	Full-gene/full-gene (DQA1, DPA1), and wide-coverage
uorkas	TypeStream™ Visual, One Lambda-Thermo Fisher Scientific	Ion S5™ XL system	A, C, B	Full-gene/NT
upahsu	TypeStream™ Visual, One Lambda-Thermo Fisher Scientific	MiSeq, Illumina	A, C, B, DRB1/3/4/5, DQA1, DQB1, DPA1, DPB1	Full-gene/full-gene (DQA1, DPA1), and wide-coverage
ussta1-1	MIA FORA FLEX, Immucor	MiniSeq and NextSeq, Illumina	A, C, B, DRB1/3/4/5, DQA1, DQB1, DPA1, DPB1	Full-gene/wide-coverage
ussta1-2	TruSight HLA Assign, Illumina	MiSeq, Illumina	A, C, B, DRB1/3/4/5, DQA1, DQB1, DPA1, DPB1	Full-gene (A, C), B (ex1-intron 6)/full-gene (DQB1, DQA1, DPA1), wide-coverage
utxask	TruSight HLA Assign, Illumina HLA Twin, Omixon	MiSeq, Illumina MiSeq, Illumina	A, C, B, DRB1/3/4/5, DQA1, DQB1, DPA1, DPB1 A, C, B, DRB1/3/4/5, DQA1, DQB1, DPA1, DPB1	Full-gene/wide-coverage Full-gene/full-gene (DQA1, DQB1 and DPA1 full-gene) and wide-coverage
ucasai	TypeStream™ Visual, One Lambda-Thermo Fisher Scientific	Ion S5™ XL system	A, C, B, DRB1/3/4/5, DQA1, DQB1, DPA1, DPB1	Full-gene/full-gene (DQA1, DPA1), and wide-coverage
upamon	HLA Twin, Omixon	MiSeq, Illumina	A, C, B, DRB1/3/4/5, DQA1, DQB1, DPA1, DPB1	Full-gene/full-gene (DQA1, DQB1 and DPA1 full-gene) and wide-coverage

Overall concordance rates of 24U B-LCL HLA-A genotypes characterized by NGS compared to consensus NGS genotypes

Table 3a.

Group	2-Field Concordance (%)					4-Field Concordance (%)				
	Identical	Con-2	Con-3	Con-4	Total	Identical	Con-2	Con-3	Con-4	Total
01	97.97	0.00	1.79	0.00	99.77	75.53	2.96	12.63	5.22	96.34
02	96.59	1.31	1.31	0.26	99.48	84.34	7.61	1.68	5.66	99.29
03	97.90	1.87	0.00	0.23	100.0	69.86	10.75	5.61	11.45	97.66
04	96.00	1.18	0.00	2.82	100.0	75.53	11.76	2.12	9.65	99.06
05	97.66	1.41	0.00	0.94	100.0	78.69	7.26	6.32	7.03	99.30
06	97.96	1.53	0.00	0.51	100.0	82.14	4.08	6.12	7.14	99.49
07	95.08	1.17	0.00	3.28	99.53	76.58	8.67	5.39	7.26	97.89
08	95.08	0.52	0.00	1.55	97.15	70.91	7.27	9.35	8.05	95.58
09	98.61	0.00	0.00	1.39	100.0	95.83	0.00	0.00	3.47	99.31
10	97.42	1.41	0.00	1.17	100.0	79.34	5.63	10.56	3.52	99.06
11	95.75	1.39	0.65	1.11	98.89	76.16	10.72	3.60	7.95	98.43
12	96.77	1.11	0.17	0.68	98.72	75.38	7.13	7.47	8.32	98.30
13	98.59	0.70	0.70	0.00	100.0	78.17	10.33	5.87	5.16	99.53
14	97.25	0.86	1.10	0.31	99.53	64.82	11.08	5.23	14.20	95.32
15	85.00	0.00	0.83	4.58	90.42	69.58	7.08	0.83	12.08	89.58
16	98.24	1.76	0.00	0.00	100.0	82.46	16.18	0.00	1.30	99.94

Abbreviations: Identical denotes unambiguous reference allele and unambiguous NGS allele concordance rates; Con-2 denotes ambiguous reference allele and ambiguous NGS allele matched at one allele; Con-3 denotes ambiguous reference alleles and unambiguous NGS alleles matched at one allele; Con-4 denotes unambiguous reference allele and ambiguous NGS allele matched at one allele.

Table 3b. Concordance rates of 24U B-LCL HLA genes characterized by NGS compared to consensus NGS genotypes

Group	Allelic resolution	HLA gene										
		A	B	C	DRB1	DRB3	DRB4	DRB5	DQA1	DQB1	DPA1	DPB1
01	2-Field	100.0	100.0	100.0	100.0	100.0	96.4	100.0	99.3	100.0	100.0	100.0
	4-Field	93.3	100.0	97.2	99.3	89.8	96.4	100.0	91.7	98.6	87.5	100.0
02	2-Field	100.0	100.0	100.0	97.9	NT	NT	NT	NT	99.5	NT	99.5
	4-Field	99.5	100.0	100.0	97.3	NT	NT	NT	NT	99.4	NT	99.5
03	2-Field	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0
	4-Field	95.8	95.8	100.0	100.0	84.2	100.0	100.0	95.8	100.0	97.9	100.0
04	2-Field	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0
	4-Field	97.9	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	93.8	100.0
05	2-Field	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	97.9	100.0
	4-Field	100.0	100.0	100.0	100.0	95.0	100.0	100.0	97.9	100.0	97.9	100.0
06	2-Field	100.0	100.0	100.0	100.0	NT	NT	NT	100.0	100.0	NT	100.0
	4-Field	100.0	100.0	100.0	100.0	NT	NT	NT	96.4	100.0	NT	100.0
07	2-Field	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	95.8	100.0
	4-Field	100.0	100.0	100.0	100.0	85.0	100.0	100.0	97.9	100.0	93.8	95.8
08	2-Field	100.0	100.0	100.0	95.8	90.0	95.0	100.0	100.0	97.7	93.8	80.0
	4-Field	100.0	100.0	100.0	89.6	84.2	95.0	100.0	97.9	95.5	93.8	80.0
09	2-Field	100.0	100.0	100.0	NT	NT	NT	NT	NT	NT	NT	NT
	4-Field	100.0	97.9	100.0	NT	NT	NT	NT	NT	NT	NT	NT
10	2-Field	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0
	4-Field	100.0	100.0	100.0	100.0	90.0	100.0	100.0	100.0	100.0	100.0	95.8
11	2-Field	100.0	100.0	100.0	96.6	NT	NT	NT	99.3	99.3	97.1	98.4
	4-Field	100.0	100.0	100.0	96.6	NT	NT	NT	99.3	98.5	94.3	98.4

Group	Allelic resolution	HLA gene											
		A	B	C	DRB1	DRB3	DRB4	DRB5	DQA1	DQB1	DPA1	DPB1	
12	2-Field	99.3	98.5	98.6	100.0	96.2	88.9	100.0	100.0	100.0	98.6	100.0	99.2
	4-Field	99.3	98.6	98.6	99.2	96.2	86.4	100.0	99.2	98.6	98.5	99.2	99.2
13	2-Field	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0
	4-Field	100.0	100.0	100.0	100.0	90.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0
14	2-Field	100.0	100.0	100.0	99.3	98.3	98.1	100.0	99.3	98.6	100.0	100.0	100.0
	4-Field	97.2	96.5	95.8	99.3	93.1	98.1	100.0	92.4	95.1	89.6	95.8	95.8
15	2-Field	91.7	91.7	91.7	91.7	NT	NT	NT	NT	85.4	NT	NT	NT
	4-Field	91.7	91.7	91.7	89.6	NT	NT	NT	NT	83.3	NT	NT	NT
16	2-Field	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0
	4-Field	100.0	100.0	100.0	100.0	100.0	100.0	100.0	99.5	100.0	100.0	100.0	100.0

Abbreviations: NT, HLA gene not typed by the participating group

Overall concordance rates of the extended B-LCL panel HLA genotypes characterized by NGS compared to consensus NGS genotypes

Table 4a.

Group	2-Field Concordance (%)				4-Field Concordance (%)					
	Identical	Con-2	Con-3	Con-4	Total	Identical	Con-2	Con-3	Con-4	Total
01	97.47	0.84	0.55	0.25	99.12	72.43	6.58	10.20	7.50	96.71
02	96.61	2.81	0.02	0.51	99.96	78.35	5.47	11.84	1.62	97.28
16	97.04	2.60	0.36	0.00	100.0	82.37	17.55	0.04	0.00	99.96
17	97.04	2.62	0.26	0.06	99.98	75.14	6.13	11.12	6.92	99.30
18	97.17	1.86	0.00	0.68	99.71	77.62	5.71	11.57	1.97	96.87

Abbreviations: Identical denotes unambiguous reference allele and unambiguous NGS allele concordance rates; Con-2 denotes ambiguous reference allele and ambiguous NGS allele matched at one allele; Con-3 denotes ambiguous reference alleles and unambiguous NGS alleles matched at one allele; Con-4 denotes unambiguous reference allele and ambiguous NGS allele matched at one allele.

Concordance rates between B-LCL HLA genes in the extended panel characterized by NGS compared to consensus NGS genotypes

Table 4b.

Group	Allelic resolution	HLA gene										
		A	B	C	DRB1	DRB3	DRB4	DRB5	DQA1	DQB1	DPA1	DPB1
01	2-Field	100.0	99.6	98.9	99.6	100.0	97.6	96.6	99.2	97.0	99.6	99.2
	4-Field	99.7	92.9	97.4	99.6	93.6	97.6	87.1	95.4	95.9	95.8	99.2
02	2-Field	100.0	100.0	100.0	99.6	100.0	100.0	100.0	100.0	100.0	100.0	100.0
	4-Field	100.0	95.3	100.0	99.6	96.7	100.0	100.0	96.9	99.8	99.4	86.1
16	2-Field	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0
	4-Field	99.8	100.0	100.0	100.0	99.6	100.0	100.0	100.0	100.0	100.0	100.0
17	2-Field	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	99.8	100.0	100.0
	4-Field	100.0	99.8	100.0	100.0	96.3	100.0	100.0	97.2	99.2	99.5	99.7
18	2-Field	99.8	100.0	100.0	99.8	99.5	100.0	95.8	99.6	99.2	100.0	99.8
	4-Field	98.6	94.8	100.0	99.4	95.7	100.0	95.7	96.6	98.6	99.6	87.1

Table 5.

Novel alleles identified in the B-LCL panel

Cell ID / Cell Name	Most similar allele	Nucleotide Substitution ^a	Codon / Amino Acid change ^b	New allele name	GenBank Accession number	IPD-IMGT/HLA Database Accession number
IHW09208 / CIR	HLA-A*02:01:01:01	G deletion	179 exon 3 / FMS	N/A result of artificial mutation	N/A	N/A
IHW09163 / HIN, JP IHW09165 / HIN, S	HLA-B*35:03:01	$\underline{\Delta}TG > \underline{\Delta}TG$	-24 exon 1 / Met > Leu	N/A result of artificial mutation	N/A	N/A
IHW09314 / V.E.C.	HLA-A*33:03:01	$\underline{GCG} > \underline{CTG}$	-20 exon 1 / Ala > Val	HLA-A*33:172	MK182707	HWS10053663
IHW09163 / HIN, JP IHW09165 / HIN, S	HLA-C*14:02:01:01	$\underline{GTG} > \underline{\Delta}TG$	248 exon 4 / Val > Met	HLA-C*14:102	MK182704	HWS10053937
IHW09194 / WON, I	HLA-DRB1*14:08	$\underline{ATG} > \underline{ACG}$	160 exon 3 / Met > Thr	HLA-DRB1*14:211	MK205123	HWS10053955
IHW09077 / T7527 IHW09121 / CHI013 IHW09182 / DCH002 IHW09183 / DCH007 IHW09184 / DCH008 IHW09198 / MIN IHW09199 / DJU, S IHW09205 / PETCH IHW09227 / JAP-NF IHW09441 / FH53	HLA-DQB1*03:01:01:01/HLA-DQB1*03:01:01:03	Hybridx1 ^c	Introns 1 and 2	HLA-DQB1*03:01:01:07	LC217997	HLA17167
IHW09333 / KI IHW09253 / THAI742	HLA-DQB1*03:01:01:01/HLA-DQB1*03:01:01:03	Hybridx2 ^d	Intron 1	HLA-DQB1*03:01:01:12	MG571442	HLA17367
IHW09070 / LUY IHW09163 / HIN, JP IHW09165 / HIN, S	HLA-DQB1*03:01:01:01/HLA-DQB1*03:01:01:03	Hybridx3 ^e	Introns 1 and 2	HLA-DQB1*03:01:01:17	MFI56573	HLA17462
IHW09182 / DCH002	HLA-DQB1*05:01:01:03	$\underline{TCC} > \underline{TCT}$	-6 exon 1 / Ser > Ser	HLA-DQB1*05:01:24:02 (Intron variant)	MG686541	HLA18105
IHW09060 / CB6B IHW09103 / KTI4	HLA-DPA1*02:02:01	$\underline{GTG} > \underline{\Delta}TG$	91 exon 3 / Val > Met	HLA-DPA1*02:07:01:01	KP774801	HLA15619
IHW09214 / KNE IHW09220 / XLI-ND IHW09248 / KHAGNI IHW09298 / HAS-15		$\underline{CCA} > \underline{CCG}$ $\underline{GTA} > \underline{GTG}$	127 exon 3 / Pro > Pro 154 exon 3 / Val > Val			

Cell ID / Cell Name	Most similar allele	Nucleotide Substitution ^a	Codon / Amino Acid change ^b	New allele name	GenBank Accession number	IPD-IMG/HLA Database Accession number
IHW09340 / 1066						
IHW09397 / DUG150	HLA-DPA1*02:02:01	<u>G</u> TG > <u>A</u> TG	91 exon 3/Val > Met	TBA	MK182706	HWS10053939
		<u>C</u> CA > <u>C</u> CG	127 exon 3/Pro > Pro			
		<u>G</u> TA > <u>G</u> TG	154 exon 3/Val > Val			
		<u>C</u> GG > <u>C</u> AG	224 exon 4/ Arg > Gln			
IHW09383 / FH9	HLA-DPA1*02:02:02	<u>C</u> CA > <u>C</u> TA	114 exon 3/Pro > Leu	HLA-DPA1*02:14	MK182705	HWS10053941
IHW09347 / 2012	HLA-DPA1*03:01	<u>G</u> TG > <u>G</u> TC	204 exon 4/Val > Val	HLA-DPA1*03:01:02	KY807146	HLA17562
IHW09275 / GU373	HLA-DPA1*04:01	<u>G</u> CG > <u>A</u> CG	190 exon 4/ Ala > Thr	HLA-DPA1*04:02	KY807145	HLA17563

^aNucleotide of the previously reported allele is listed first, differences are underlined.

^bAmino acid encoded by the previously reported allele is shown first.

^cHybridx1 consists of the following mutations: HLA-DQB1*03:01:01 = Intron-1 G > T substitution at 1490 bp, Intron-2 G > A substitution at 3371 bp; HLA-DQB1*03:01:01:03 = Intron-1 T > C substitution at 742 bp and G > T at 1490 bp (HLA-DQB1*03:01:01:07)

^dHybridx2 consists of the following mutations: HLA-DQB1*03:01:01 = Intron-1 G > T substitution at 1490 bp, Intron-2 G > A substitution at 2979 bp and G > A at 3371 bp; HLA-DQB1*03:01:01:03 = Intron-1 T > C substitution at 742 bp and G > T at 1490 bp; Intron-2 G > A substitution at 2979 bp (HLA-DQB1*03:01:01:12)

^eHybridx3 consists of the following mutations: HLA-DQB1*03:01:01 = Intron-1 G > T substitution at 1490 bp, Intron-2 C > G substitution at 3209 bp and G > A at 3371 bp; HLA-DQB1*03:01:01:03 = Intron-1 T > C substitution at 742 bp and G > T at 490 bp; Intron-2 C > G substitution at 3209 bp (HLA-DQB1*03:01:01:17)

Abbreviations: FMS, Frame shift mutation; N/A, non-applicable; TBA, to be assigned

Table 6.

HLA haplotypes observed in the B-LCL panel

Loci	Haplotype	Ethnic Group
HLA-C-HLA-B	C*07:02:01:03-B*07:02:01	Caucasoid
HLA-C-HLA-B	C*07:01:01:01-B*08:01:01:01	Caucasoid
HLA-C-HLA-B	C*05:01:01:01-B*18:01:01:01	Caucasoid
HLA-C-HLA-B	C*05:01:01:02-B*44:02:01:01	Caucasoid
HLA-C-HLA-B	C*03:04:01:01-B*40:01:02	Caucasoid
HLA-C-HLA-B	C*04:01:01:01-B*35:01:01:02	Caucasoid
HLA-DRB3-HLA-DRB1	DRB3*02:02:01:01~DRB1*03/12	Caucasoid, African American, Australian Aboriginal, Asia
HLA-DRB3-HLA-DRB1	DRB3*02:02:01:02~DRB1*11	Caucasoid, Hispanic, Asia
HLA-DPA1~HLA-DPB1	DPA1*01:03:01:01~DPB1*02:01:02	Caucasoid, African American, Australian Aboriginal, Native American, Asia
HLA-DPA1~HLA-DPB1	DPA1*01:03:01:02~DPB1*04:01:01:01	Caucasoid, African American, Australian Aboriginal, Native American, Asia
HLA-DPA1~HLA-DPB1	DPA1*01:03:01:03~DPB1*03:01:01	Caucasoid, Native American
HLA-DPA1~HLA-DPB1	DPA1*01:03:01:04~DPB1*04:01:01:01	Caucasoid, Asia
HLA-DPA1~HLA-DPB1	DPA1*01:03:01:05~DPB1*04:02:01:01	Caucasoid, African American
HLA-DPA1~HLA-DPB1	DPA1*01:03:01:05~DPB1*04:02:01:02	Caucasoid, Native American, Asia
HLA-DPA1~HLA-DPB1	DPA1*02:02:02~DPB1*05:01:01	Asia, Australia Aboriginal
HLA-B-HLA-DRB4-HLA-DRB1	B*57:01:01~DRB4*01:03:01:02N~DRB1*07:01:01:01/02	Caucasoid
HLA-DRB1~HLA-DQA1~HLA-DQB1	DRB1*03:01:01:01/02~DQA1*05:01:01:01/02~DQB1*02:01:01	Caucasoid
HLA-DRB1~HLA-DQA1~HLA-DQB1	DRB1*04:01:01:01/02~DQA1*03:03:01:01~DQB1*03:01:01:01	Caucasoid, African American, Australia Aboriginal, Hispanic
HLA-DRB1~HLA-DQA1~HLA-DQB1	DRB1*04:05:01~DQA1*03:03:01:03~DQB1*04:01:01	Asia, Hispanic
HLA-DRB1~HLA-DQA1~HLA-DQB1	DRB1*04:05:01/DRB1*04:10:01~DQA1*03:03:01:02~DQB1*04:02:01	Asia
HLA-DRB1~HLA-DQA1~HLA-DQB1	DRB1*07:01:01:01/02~DQA1*02:01:01:01/02~DQB1*02:02:01:01	Caucasoid, African American
HLA-DRB1~HLA-DQA1~HLA-DQB1	DRB1*08:03:02~DQA1*01:03:01:03/04~DQB1*06:01:01	Australia Aboriginal, Japanese
HLA-DRB3-HLA-DRB1~HLA-DQA1~HLA-DQB1	DRB3*01:01:02:01/02~DRB1*13:01:01:01/02~DQA1*01:03:01:02/06~DQB1*06:03:01	Caucasoid, Thailand, African, African American, Australia Aboriginal
HLA-DRB3-HLA-DRB1~HLA-DQA1~HLA-DQB1	DRB3*01:01~DRB1*13:02:01~DQA1*01:02:01:04/06/07~DQB1*06:04:01	Caucasoid, Asia

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Loci	Haplotype	Ethnic Group
HLA-DRB5~HLA-DRB1~HLA-DQA1~HLA-DQB1	DRB5*01:01:01~DRB1*15:01:01:01/02/03~DQA1*01:02:01:01/03/05~DQB1*06:02:01	Caucasoid, Hispanic
HLA-C~HLA-B~HLA-DRB1~HLA-DRB4~HLA-DRB1~HLA-DPA1~HLA-DPB1	C*06:02:01:01~B*13:02:01~DRB4*01:03:01:01/03~DRB1*07:01:01:01/02~DPA1*02:01:01:01~DPB1*17:01	Caucasoid, African American, Asia, Australia Aboriginal
HLA-C~HLA-B~HLA-DRB5~HLA-DRB1~HLA-DQA1~HLA-DQB1	C*12:02:02~B*52:01:01:02~DRB5*01:02~DRB1*15:02:01:01/02~DQA1*01:03:01:01~DQB1*06:01:01	Asia, Caucasoid, Middle Eastern