LINEAR ARRAY GENOTYPING SYSTEM FOR HLA

1 PURPOSE

The purpose of this Standard Operating Procedure (SOP) is to outline procedures for genotyping nPOD donor cases based on human leukocyte antigen (HLA).

2 SCOPE

This SOP will be applied to all nPOD donor cases submitted for Hi Res HLA genotyping.

3 RESPONSIBILITIES

3.1 Managers and supervisors are responsible for making sure that technicians are properly trained and equipment and facility are maintained in good working order.

3.2 Laboratory personnel are responsible for reading and understanding this SOP and related documents and to perform these tasks in accordance with the SOPs. They are responsible for following clinical laboratory and tissue banking best practices.

4 EQUIPMENT and MATERIALS

The equipment and materials required are determined as per laboratory specifications, whichever is suitable for the specific task or procedure.

5 SAFETY

5.1 Use universal safety precautions when handling human samples and personal protective equipment (e.g., face mask with shield, gloves, lab coat or apron).

6 PROCEDURE

Background

The linear array genotyping system is a sequence-specific oligonucleotide probe (SSOP)-based genotyping method. In this method, a series of unlabelled oligonucleotide probes, corresponding to sequence motifs in a given HLA gene locus, is immobilized as a series of parallel lines on a backed nylon membrane. The nylon membrane with immobilized probes will be referred to as a "strip" in this document. PCR amplification is used to amplify the relevant polymorphic exon(s) of the locus, and the PCR product is hybridized to the probe array. The resultant data resemble a UPC, or bar code, and the series of positive probe signals on the strip is used to determine the genotype of the sample at that locus. The procedure for generation of data for all of the assays is similar and will be outlined in detail below.

Assays

The following set of six assays will be utilized for the nPOD genotyping effort:

1) HLA-A
2) HLA-B
3) HLA-C
4) DQA1/DQB1
5) DPA1/DPB1
6) DRB1 HiRes

All reagents for the assays are provided with the kits. Procedures are described in more detail below. Briefly, the procedure begins with a PCR amplification for which a "Master Mix" and $MgCl_2$ solution are provided. The user need only provide water and template DNA. All assays have been validated on a Perkin Elmer 9600 thermal cycler. The use of this machine, if available, is recommended. The PCR product is subsequently denatured and hybridized to the linear array (strip) in the Bee Blot machine. Denaturation solution and all reagents for the hybridization/color development process are provided with the kits. After the color development, strips are placed in the StripScan tray (provided) for image scanning. Initial data analysis is performed in StripScan, then the data are transferred to SCORE software for generation of "authorized calls" and generation of data output files. SCORE genotyping software contains "typing kits" for all of the assays. The probes are listed by number, and each has a designation (e.g., "EEFARF") that is usually based on the one-letter amino-acid code for the region encoded by the sequence motif. Once the data are transferred to SCORE, the user selects the "authorized" call and truncates all allele calls to conform to the 4-digit standard. The linear arrays are hybridization-based, and not every PCR product/probe hybridization will be equally robust. Some probes in some of the assays may tend to be produce faint signals, while some may give a very dark signal with a tendency to cross-hybridize (i.e., give a false positive result). As the user becomes familiar with a particular assay, he/she will recognize patterns and become adept at interpretation of less-than-optimal signals. The Noble laboratory has extensive experience with this HLA genotyping system, which is utilized both for projects within the lab and for samples from the Type 1 Diabetes Genetics Consortium.

HLA Class I

The HLA Class I assays involve co-amplification of exons 2 and 3 of the Class I locus, followed by a single hybridization to a strip which contains probes corresponding to sequence motifs in each exon. A positive control, or "ALL" probe, is included for each exon. The HLA-C assay consists of 36 probes, with the exon 2 ALL probe at position 17 and the exon 3 ALL probe at position 36. The HLA-B assay consists of 81 probes, with the exon 2 ALL probe at position 41 and the exon 3 ALL probe at position 81. The HLA-A assay consists of 57 probes, with the exon 2 ALL probe at position 21 and the exon 3 ALL probe at position 57.

HLA Class II

The assays for the DQ- and DP-encoding genes are similar. These assays involve a co-amplification of the DQA1 and DQB1, or of the DPA1 and DPB1, loci followed by a single hybridization. All probe-binding data are processed together: the SCORE typing kits sorts out the DQA1 (or DPA1) from the DQB1 (or DPB1) probes and allows calling of the genotypes separately.

The DRB1HiRes assay is a single PCR followed by a single hybridization and is performed in the same manner as all of the other assays.

Ambiguities

Ambiguities in genotyping results are a feature of almost any HLA genotyping system. The only way to determine a genotype with absolutely no ambiguity is to use allele-specific amplification followed by DNA sequencing to identify each of the two alleles in a genotype, a process that is extremely expensive and time-consuming and is not feasible for a large-scale project. Useful HLA genotyping data can be obtained with more rapid and less expensive methods, such as the SSOP-based linear array system that will be utilized for the nPOD project, as long as the user understands how to handle the intrinsic ambiguities in the system. The two categories of ambiguities are discussed below.

Allele Ambiguities

Allele ambiguities arise from polymorphisms whose sequence motifs are not included in the probes in the assay. Two alleles that differ only by an untested polymorphism will not be distinguishable with the assay. For example, the two alleles DPB1*0501 and DPB1*3801 differ only at the region encoding amino acid 17. Previous versions of the DPB1 SSOP genotyping system did not include a probe specific for this polymorphism; thus, a probe-binding pattern consistent with DPB1*0501 was also consistent with DPB1*3801. Because DPB1*3801 is very rare and DPB1*0501 is more frequent (f = ~.08 in Caucasians), the assumption could be made with reasonable confidence that an allele consistent with DPB1*0501 and DPB1*3801 was, in fact, DPB1*0501. Although this particular ambiguity has been resolved in the current version of the DPB1 genotyping assay, similar ambiguities are present in other assays, particularly the Class I assays, where thousands of alleles have been identified to date.

Genotype Ambiguities

With all HLA typing methods, simultaneous typing of the two chromosomes in an individual can result in genotype ambiguities. As a simplified example, consider an allele with two polymorphic regions with two sequence motifs at each region. Probes for motifs in region one are termed A and B, and probes for motifs in region 2 are called C and D. If a sample from a heterozygous individual hybridizes to both probes A and B at region 1 and to probes C and D at region 2, the two possible genotypes that can be inferred are AC,BD or AD,BC. The assay data do not allow distinction of the phase of the probe-binding sequences. Methods to distinguish among genotype possibilities include the following: 1) If only one of the possible genotypes includes known alleles, then that genotype can be reasonably inferred as the correct one. However, if both genotypes include two known alleles, the distinction is more difficult to make. 2) If family-based genotyping data are available on a sample from a parent or child of the individual whose sample is being tested, the correct genotype can usually be inferred. For individual samples, such as those found in case-control studies, this is not possible. 3) When family data are unavailable, the ambiguity can be resolved by amplification of the sample with a primer that is specific for one of the polymorphic regions (allele-specific amplification) to allow only one of the two sequences to be amplified. Subsequent genotyping of the PCR product will determine which other probe-binding regions are on the same chromosome as the allele-specific primer. For the simplified example above, if an allele-specific primer complementary to the A sequence in region one were used for amplification, and the subsequent genotyping indicated the presence of the D sequence at region 2, the genotype of the sample could be unambiguously determined to be AD,BC. The DRB1 linear array assay is designed to incorporate allele-specific genotyping, but the other genotyping assays to be utilized for the T1DGC are not. Thus, in a situation where multiple genotype possibilities exist for a T1DGC sample genotype, the user will be required to make a best estimation to create the authorized genotype call.

HLA Typing Protocol for T1DGC Samples

DNA will be supplied to the HLA genotyping laboratories in screw-capped tubes with bar-coded labels. A total of 2 ug of DNA at a concentration of 20 ng per ul (total volume = 100 ul) will be shipped to each laboratory. Samples will be spun briefly and 50 ul of the sample will be transferred to a 96-well plate. Batches of 11 samples will be placed in wells 1 through 11 of a 96-well plate. These samples will then be transferred to a 96-well PCR amplification tray containing reagents for each of the 6 assays in 6 separate rows. Because the amplification and hybridization conditions are standardized for all assays, generation of primary data can be completed for all 6 loci simultaneously. Subsequently, data analysis will be performed separately for each locus.

1. DNA Amplification

Premixes are provided at 2X concentration for each test. Premix contains buffer, primers, and enzyme. Template DNA and $MgCl_2$ (provided) must be added.

a. Amplification Reaction

30 ul PCR premix for desired assay
3 ul genomic DNA template at 20 ng/ul
15 ul 6mM $MgCl_2$
12 ul H_2O
60ul total volume

Combine sufficient reagents for 100 reactions, excluding DNA, mix by inversion, and distribute 57 ul of the mixture to each tube in the PCR tray. Transfer 3 ul of template DNA from the master plate to each tube in the PCR setup plate with a multichannel pipettor to avoid sample mixups. Be sure to have both the master plate and the PCR plate in the correct orientation, with position A1 in the upper left corner. Seal the tubes and cycle as specified below.

b. Cycling Parameters

35 cycles of:
15 seconds at 95°C
45 seconds at 60°C
15 seconds at 72°C

Follow the cycling with 5 minutes at 72°C. Afterwards, set the hold cycle for 15°C indefinitely.

In a Perkin-Elmer 9600 thermal cycler, link the "cycling" and the two "hold" programs together into a method so that they are run sequentially.

2. Gel Electrophoresis (optional)

Gel electrophoresis of PCR products is not routinely performed; however, this procedure can be utilized, in the case of less than optimal results, to check the efficiency of the amplification. The PCR products are in the 200-600bp size range. When electrophoresis is warranted, 5 ul of the 60 ul PCR reaction may be electrophoresed on an agarose gel to confirm the presence, size, and quantity of PCR product. A mixture

of 3% NuSeive agarose with 1% regular agarose works well for resolving bands in the desired size range. Invitrogen makes a buffer-free system ("e-gels") that works well for visualizing these products.

3. Hybridization/Detection

a. Materials

The following materials are used:

- DNA denaturing solution
- SSPE concentrate
- SDS concentrate
- Streptavidin-HRP conjugate
- Substrate A
- Substrate B
- Locus-specific Linear Array Strips
- Distilled water
- Waterproof marking pen
  - Erasable ink pens work well, as do Extra Fine Point Sharpie markers. DO NOT USE VWR brand markers because the ink will run and may obscure some of the probe-binding data.

b. Preparation of Solutions

Hybridization (hyb), wash, and citrate solutions should be prepared in advance and may be stored for up to 3 months at room temperature.

1X wash solution (1 liter)
50 ml 20X SSPE   Final: 1X
25 ml 20% SDS   0.5%
925 ml distilled water

4X hyb solution (1 liter)
200 ml 20X SSPE   Final: 4X
25 ml 20% SDS   0.5%
775 ml distilled water

Note: Always add H$_2$O first, then SSPE and SDS. Do not mix concentrated SSPE and SDS because an insoluble precipitate will form.

Citrate solution (600 ml)
30 ml Citrate concentrate
570 ml distilled water

Enzyme conjugate and Working substrate solutions should be freshly prepared just prior to starting the hybridization procedure

Enzyme conjugate
1 part Streptavidin-HRP
33.125 parts 1X wash solution

Working substrate
1 part Substrate B
4 parts Substrate A

c. Denaturation of PCR product

To denature the PCR product, add 35ul of denaturing solution to 35ul PCR product. This is most easily done by pipetting 35ul of denaturant into all wells of a 96-well plate (any round-bottom or V-bottom plate should work) and then adding 35ul of PCR product to the appropriate wells with a multichannel pipettor and gently pipetting up and down a few times. Let the mixture of PCR product and denaturant sit at room temperature for approximately 15 minutes before using to allow the DNA time to denature. Denaturing the DNA just prior to setting up the Bee Blot machine (see below) works well.

d. Hybridization/Color Development

Label strips for the appropriate locus at the right-hand end with a waterproof marker. Avoid writing on the probe lines, and DO NOT WRITE on the left of the reference ink line, or StripScan will not be able to find the ink line and scan your strip. Label the strip with the plate number and location, e.g., 5-F11 will designate the sample that can be found in your 5th plate of samples in position F11. This labeling system is more legible and less error-prone than trying to write the 7-digit barcode number on the strip.

The strips are 4 mm wide, and the StripScan trays have been manufactured to fit them exactly. DO NOT CUT the strips in half horizontally in an attempt to conserve reagents. This will adversely affect the ability of the StripScan software to assist you with data analysis.

Handle the strips carefully, and use gloves when handling. As with any hybridization to Nylon membrane, touching the strips with bare hands may leave residues on them that can interfere with hybridization. Place the strips into the Bee Blot tray. The Bee Blot tray is coated with Teflon: do not use metal forceps to place strips in or remove strips from the trays, or you might damage the trays. At the appropriate point in the Bee Blot Program, which can be found in the Bee Blot manual from Bee Robotics, pipet the denatured PCR product into the wells of the Bee Blot tray. The tray has a capacity of 48 assays and is configured with appropriate spacing for multi-channel pipetting. Thus, 48 samples, or one-half of a PCR tray, can be added to the Bee Blot tray with only four pipetting operations, using a 12-channel pipettor. Be careful to align the tips properly when pipetting. Misaligning the tips will cause the samples to be hybridized to the wrong strips. Misaligning the tips, then correcting the error before dispensing, can still interfere with hybridization results. Enough residual denatured PCR product is on the outside of the tip that touching it to the hybridization solution in the wrong well can lead to background bands for robust probes.

Run the machine according to the manufacturer's directions.

4. Detection

When the hybridization and color development process is finished, carefully remove the strips from the Bee Blot tray, and place them in the ScanPad. Do not use metal forceps to handle the strips. Plastic

forceps may be used. Alternatively, a plastic pipet tip may be used to push up one end of the strip from the tray so that it may be gently grasped with a gloved hand or metal filter forceps.

Carefully transfer strips to the ScanPad and cover with an acetate sheet. Blot any excess liquid surrounding the strips to avoid a splotchy image. Make sure that the acetate sheet (transparency), used to cover the strips, is clean. Scan the image according to the StripScan protocol.

5. Genotyping with SCORE Software

The StripScan and SCORE software provided by the T1DGC will create a database with your sample and the probe-binding data. You must proofread the data prior to analysis, incorporating your knowledge of the probe behavior patterns for the assay. Remember that the SCORE software will create a "virtual strip" for you that is based directly on the pixel values imported from StripScan. These virtual strip images are useful in that placing the cursor over a probe position will reveal the probe number and pixel value. However, these images are not the actual data, but rather a virtual reconstruction of them. Looking at the original data, whether on the strips themselves, the .jpeg file from the original scan, or a photographic image of the original strips taken shortly after they are run, is of great utility. Taking a photograph or digital image of all of the strips from a given run immediately after the scanning procedure is finished is an excellent way to preserve an easy-to-see record of the primary data for later reference. Looking at all of the strips in the context of the others in the same run, as opposed to just looking at the virtual image of a given strip, can be quite beneficial in determining the appropriate call for weak or cross-hybridizing probes. Creating a digital or photographic image of the strips is strongly recommended but not required.

Genotype analysis of the scanned strip images begins in the SrtipScan program, where the user selects among possible genotype patterns, based on the probe reactivity. Data are then exported into the SCORE software program, the probe binding pattern is analyzed by the software, and the user selects an "authorized call" from among the possible genotype calls. This is described below.

After importing the StripScan data, navigate to the sample and locus of interest and click "Results" then "Analysis." One or more condensed genotypes (e.g., DRB1*03, DRB1*04) will appear in the window on the left. Click once on the arrow to the left of the condensed genotype. If you accidentally click twice, the condensed genotype will appear in the boxes for the authorized calls. These can be edited later (see below). Clicking once will cause all of the specific genotype possibilities associated with the condensed genotype to appear in the box just to the right of the condensed genotypes.

Tolerance

If the probe-binding pattern is incorrect, a dialog box will appear stating that no genotype call can be made without giving the assay a "tolerance." Applying a tolerance of one means that the software will attempt to call a genotype by checking each probe position and changing the call for one of the probes in the assay. If a genotype is called with a tolerance of one, look at the genotype possibilities in the specific genotypes box. An icon that looks like a bomb will appear to the left of each genotype. Click on the bomb to see which probe call was altered to allow that genotype call.

IMPORTANT: Go back and look at the original strip image to see if the probe indicated by the "bomb" might have been called incorrectly. If so, click the SSO/SSP button, click the "Edit Results" box, and modify the data for that probe. If not, keep looking for the error until you can make a decision as to what

the correct probe-binding sequence should be, and use that sequence to generate the genotype. DO NOT authorize any genotype calls that have a tolerance greater than zero.

The tolerance feature is a wonderful tool built into SCORE that is extremely valuable to aid in finding errors introduced due to anomalies in probe behavior. It should never be used as a crutch to allow a call from incorrect primary data. **ONLY THE USER CAN DECIDE WHETHER OR NOT A PROBE SHOULD BE POSITIVE. THE SOFTWARE CANNOT MAKE THIS DECISION!!!!!** The tolerance feature of SCORE is only capable of suggesting where to look for potential errors in the probe-binding pattern. It is not capable of looking at your primary data and making a decision about the intensity of a probe signal. In some cases, the probes suggested as possibly miscalled by the tolerance feature of SCORE are not, in fact, miscalled. For example, the addition or deletion of a single probe signal, as suggested by the tolerance feature in SCORE, may allow a genotype call to be made; however, the suggested probe is robust and no question exists about its call. In that case, the user must look further to find the source of the error, which may result from multiple miscalled probes. The importance of careful examination of the primary data cannot be overstated here.

Authorizing Genotypes

When you are satisfied that the probe-binding pattern is correct, and the SCORE software is able to give resulting genotype calls without the use of the tolerance tool, then you may authorize a genotype call. Because so many alleles (including silent mutations, intron mutations, null and low expression alleles) exist for many of the loci, these assays will usually produce multiple specific genotype possibilities, any of which is consistent with the probe-binding pattern. The responsibility of the HLA genotyper is to select an "authorized" genotype call that the statisticians will use for their analyses. All primary probe-binding data are retained in the database for future use if needed.

The genotyping lab will use their knowledge and experience with HLA to select the authorized genotype. This decision is based on the likelihood of finding the alleles in the genotype as well as, for families, having consistency of the allele call within the family, i.e., if an allele is called for a proband, that allele should also be present in one of the parents. In general, alleles with high frequency were discovered first and have the low numbers at the third and fourth position. SCORE lists the genotype call possibilities in numeric order; therefore, the first genotype on the list is frequently the most likely and, therefore, becomes the authorized call. However, exceptions to the general rule occur, so the top genotype should not be automatically selected without looking at the other possibilities.

The T1DGC authorized calls **MUST** be in 4-digit nomenclature, e.g., DPB1*0301, rather than DPB1*030101. The first four digits of any allele designation are sufficient to determine the primary amino acid sequence. Additional numbers represent silent or intronic polymorphisms and do not affect protein structure. For example, DPB1*030101 and DPB1*030102 encode the same protein, which should have the same effect on disease susceptibility. In many cases, reducing the genotype possibilities to their four-digit forms significantly reduces the number of possible genotypes. For example, the following three genotypes:

A*020101, A*680101
A*020102, A*680101
A*020104, A*680101

reduce to A*0201, A*6801 in four-digit nomenclature, reducing the number of possible genotypes from 3 to 1.

To authorize calls, double click on a specific genotype that contains the four-digit genotype of choice. In the above example, any of the choices could be used. A dialog box will ask you if you want to authorize the results. Click yes. The genotype will appear in the "authorized interpretation" section of the screen. Manually edit the first allele in the authorized genotype to create the 4-digit allele, i.e., remove the last two digits. When you go to the next allele to edit, the dialog box will again appear asking if you want to authorize these results. Click yes, then proceed with editing the second allele. Click on the "close" button at the bottom right of the screen. You will again be prompted to authorize the results. Click yes. The results screen will close and you will now see the sample screen, where your authorized interpretation will be present in the appropriate boxes. All authorized calls must contain two alleles. For homozygous samples, manually enter the allele call into the second box for the authorized call. This should be done with "Copy" and "Paste" functions to avoid typographical errors. DO NOT LEAVE ANY AUTHORIZED CALL BOXES EMPTY. Also note that the "authorized call" fields are simple text fields. If the probe-binding pattern and resulting analysis are modified to create a different genotype, the user MUST manually re-create the authorized call.

Once all authorized calls are created, results may be generated as an xml file or as an Excel spreadsheet for transfer to nPOD.

7 REFERENCES

7.1 N/A

8 REVISION HISTORY

Version	Date	Revision

	Name	Signature	Date
Prepared by	Janelle Noble		
Approved by	Mark Atkinson		