DNA-Based HLA Typing Methods

American Foundation for Donation and Transplantation Histocompatibility Specialist Course

2021

Gerald P. Morris MD PhD
University of California San Diego
Conflict of Interest statement:

• Member, Scientific Advisory Board, Transplant-related NGS, Thermo-Fisher/One Lambda

• Research support and scientific meeting travel reimbursement, Thermo-Fisher/One Lambda

• Research support and scientific meeting travel reimbursement, CareDx
Learning Objectives

• Review methodology of common DNA-based HLA typing methods, sequence-specific primer PCR (SSP), sequence-specific oligonucleotide probes (SSO), and sequence-based analysis

• Compare benefits and limitations of DNA-based HLA typing approaches

• Discuss approaches to identify specific HLA alleles
HLA structure influences biologic function

- HLA molecules function to present antigens to T cells

- HLA binds linear peptide fragments in the cleft formed by the 2 alpha helices and the beta-pleated sheet

- T cell antigen receptors (TCR) align atop the alpha helices

Morris GP. 2012. *Nat Immunol*
HLA has genetic variation encoding “functional” differences

- Majority of polymorphisms in HLA genes are in the peptide binding groove and helices
- Variation in HLA causes variation in peptide binding and antigen presentation
HLA polymorphisms influence biologic function

- Amino acid sequences on the beta-pleated sheet influence what antigens can be bound by the HLA molecule and presented to T cells.

- Amino acid sequences on the alpha helices also influence antigen binding and presentation and influence how the HLA molecule interacts with TCRs.

Polymorphic amino acids on alpha helices

Polymorphic amino acids in peptide binding groove
HLA typing by serology

Jean Dausset

• HLA defined in 1952 as primary barrier to allogeneic transplantation
  • Studied serum reactivity of multiparous women
  • Identification of alloantibodies that recognized HLA-A2

• Through the 1980s, HLA typing depended on the identification of mono-specific sera and using that to characterize HLA type
Characterizing the serologic reactivity of cells was a reasonable approach for solid organ transplantation.

- Alloantibodies can cause hyperacute rejection and antibody-mediated rejection.
Typing serologic specificities defines HLA variation on the “outside” of the molecule

• HLA typing in SOT primarily to facilitate serologic compatibility
• Relatively limited number of recognized serologic specificities
Recognized serologic antigens represent only a fraction of HLA genetic diversity

Protein-coding HLA allele variants

IMGT/HLA database, January 2018

• ~18,000 defined protein-coding variants vs. ~150 characterized serologic specificities
• Serologic characterization of HLA using defined reagents may not reflect all possible serologic reactivity

• T cell reactivity against HLA is driven by differences in peptide presentation determined by amino acid differences in the beta-sheet
Challenges in clinical HLA typing

• Large amount of genetic diversity
  • Requirement to resolve ambiguity
  • Need for robust methodology
DNA-based HLA typing

- Majority of polymorphisms are in predictable areas
  - Class I- exons 2, 3, and 4
  - Class II- β chain exon 2
  - Enables focused DNA analysis for polymorphisms

Figure 5-18 Immunobiology, 7ed. (© Garland Science 2008)
HLA nomenclature defines polymorphism as expected to affect biologic function.
HLA nomenclature defines polymorphism as expected to affect biologic function.

Amino acids that define HLA-A2 serologic reactivity.

Amino acids that differentiate HLA-A*02 alleles - not serologically recognizable, but can stimulate T cell responses.
DNA-based HLA typing methods

**Primer-Specific PCR**
- Multiple amplifications with primers specific for polymorphic sequences

**Oligonucleotide Probes**
- Amplification with primers to conserved regions
- Hybridization with probes specific for polymorphic sequences
- High-throughput multiplexed flow cytometry

**DNA Sequence Analysis**
- Amplification with primers to conserved regions
- Sequencing reaction
- Sequence analysis to resolve phasing, identify alleles
Low-resolution HLA genotyping

- In most cases, the first-field typing describes the genotype encoding the serologic antigen.
- In some cases, second-field typing is required to refine the serologic reactivity to the split-antigen level.

Polymorphic amino acids that define HLA-A2

HLA-A2
Low-resolution HLA genotyping – serologic split antigens

First-field represents serologic antigen
A*02 = low-resolution genotyping
A*02:101 = high-resolution genotyping
A2 = serologic antigen equivalent

Second-field defines serologic split antigens
DQB1*03 = low-resolution genotyping, but might not be adequate
DQB1*03:01 = might be low-resolution or high-resolution depending on method used
DQ7 = serologic antigen equivalent
HLA typing by Sequence-Specific Primer PCR (SSP)

- The presence or absence of defining DNA sequences can be determined by PCR.
- Each primer has specific DNA sequence that is the compliment to the target sequence. If PCR with that SSP generates an amplicon, the positive reaction indicates the presence of the target sequence.
- Specificity results from stringent PCR primer annealing temperature.
- Genotyping determined by PCR reaction pattern:
  - Typically by running agarose gel and examining banding pattern.
  - Current software performs pattern analysis.
Real-time PCR is a modified version of SSP

- Traditional SSP utilizes endpoint PCR—typically ~30 cycles of PCR to maximize production of amplicon to enable visualization by gel electrophoresis.
- Real-time PCR measures amplicon generation during the exponential phase:
  - Intercalation of fluorescent dye binding double-stranded DNA amplicons.
  - Generation of fluorescent reporters via probe cleavage during PCR.
- More easily multiplexed.
- Melt curve analysis provides QC for target sequence.
**HLA typing by Sequence-Specific Oligonucleotide probe (SSO)**

- SSO targets allele-defining DNA sequences using labeled oligonucleotide probes
- Basically a multiplexed Southern Blot
- Target DNA (either genomic or amplicon) is hybridized to nitrocellulose-bound SSOs and target sequences detected by visualization
- Specificity results from stringent oligonucleotide annealing temperature
- Genotyping is determined by hybridization pattern
HLA typing by flow cytometry multiplex SSO

- Blot-style hybridization is technically challenging and low-throughput
- Fluorophore-barcoded microbead flow cytometry (Luminex) enables multiplexing of SSO reaction in high-throughput format
- Each barcoded bead has a specific SSO probe
- Typing is determined by detection of double stranded DNA on bead surface
- Current software performs pattern analysis
Ambiguity of SSP/SSO HLA typing

- Low-resolution typing methods are very good at determining low-resolution genotypes
- Not very good at estimating high-resolution genotypes
“Intermediate” resolution typing methods provide “intermediate” data

- Review of 295 deceased donor typings (2017-2018) performed using RT-PCR and confirmed by NGS
- Actual high-resolution type frequently not the “lowest number” in result allele string
- 0.7% (not including DPB1) of NGS-confirmed alleles not in string of likely results by RT-PCR
- Take home message – even when the typing report provides as 2-field typing result, it’s only “high-resolution” if tested by a high-resolution method (ie. NGS)
Comparison of SSP/SSO HLA typing

<table>
<thead>
<tr>
<th>Method</th>
<th>Resolution</th>
<th>Throughput</th>
<th>TAT</th>
<th>Scalable</th>
</tr>
</thead>
<tbody>
<tr>
<td>SSP</td>
<td>Low/~high (depends on SSP primer set and the number of reactions)</td>
<td>Low</td>
<td>Fast</td>
<td>No</td>
</tr>
<tr>
<td>RT-PCR</td>
<td>Low</td>
<td>Moderate</td>
<td>Faster</td>
<td>Sort of</td>
</tr>
<tr>
<td>SSO</td>
<td>Low</td>
<td>High</td>
<td>Not quite as fast</td>
<td>Yes</td>
</tr>
</tbody>
</table>

- All methods have different cost/benefit ratio
- All useful for low-resolution HLA genotyping depending on clinical need and workflow
High-resolution HLA genotyping

- **HLA G groups** – groups of alleles with the same nucleotide sequences for the antigen recognition domain (ARD)
- **HLA P groups** – groups of alleles predicted to have the same ARD amino acid sequence

**Polymorphic amino acids that define HLA-A*02**

Polymorphic amino acids that differentiate HLA-A*02 alleles
Technical advances in HLA typing improved HSCT outcomes

• 1960s- Serologic HLA typing
  • Only identifies limited polymorphisms (first field)

• 1980s- Molecular analysis
  • Sequence-specific oligonucleotide probes (SSO)
  • Sequence-specific primer PCR (SSP)
  • Capable of allele-specific (2-field) genotyping, but requires extensive multiplexing, and may miss differences outside of targeted areas

• 2000s- DNA sequencing based typing (SBT)
  • Not limited to focusing on “known” DNA sequence variation- can evaluate all sequences
HLA genotyping ambiguity is decreased by DNA Sequence-Based genotyping (SBT)

<table>
<thead>
<tr>
<th></th>
<th>AFA</th>
<th>API</th>
<th>CAU</th>
<th>HIS</th>
</tr>
</thead>
<tbody>
<tr>
<td>SERO</td>
<td>37.55%</td>
<td>37.24%</td>
<td>19.57%</td>
<td>43.34%</td>
</tr>
<tr>
<td>SSO</td>
<td>2.36%</td>
<td>5.30%</td>
<td>1.31%</td>
<td>3.24%</td>
</tr>
<tr>
<td>SBT</td>
<td>&lt; 0.01%</td>
<td>&lt; 0.01%</td>
<td>&lt; 0.01%</td>
<td>&lt; 0.01%</td>
</tr>
</tbody>
</table>

- Technical improvements have reduced HLA typing ambiguity
- Standard requires 2-field typing for HSCT
  - *In silico* “de-resolution” of SBT-determined 2-field genotyping estimated the degree of ambiguity in SSO and serological typing
  - Noteworthy that the degree of ambiguity for targeted analyses disproportionately affects non-Caucasian populations

Adapted from Paunic V. 2012. *PLOS One.*
Sanger sequencing-based typing (SBT)

<table>
<thead>
<tr>
<th>ddA</th>
<th>ddT</th>
<th>ddC</th>
<th>ddG</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ACTAGTCACGTAC</td>
<td>ddA</td>
<td>ddT</td>
<td>ddC</td>
</tr>
</tbody>
</table>

- Sanger DNA sequencing utilizes incorporation of dideoxy- (dd) terminator nucleotides during PCR amplification.
- Sort amplicons by size, know terminal nucleotide base.
- Electrophoretic gel-based analysis has limited throughput.
- Capillary electrophoresis, combined with multiplexed fluorescently-labeled dd nucleotides improved throughput and facilitated clinical application.
Technical limitations of SBT limit HLA genotyping

- SBT limited to ~800-1000 bp
  - HLA genes range from ~4 - >14 kb
  - May require multiple amplicons for analysis

- Extended homopolymers can be difficult to resolve

- Unable to define cis/trans relationship between multiple polymorphic nucleotides
  - Polymorphic nucleotides appear as positive for 2 nucleotides at a given position
  - n polymorphisms generate $2^n$ possible genotypes

Sanger data: G T A/T G C T/G A

Possible genotypes: GTAGCTA
GTTGCTA
GTAGCGA
GTTGCGA
SBT ambiguity

- Complimentary polymorphisms between alleles requires stranded sequencing
- Polymorphisms outside sequenced exons
- Incomplete reference allele databases

IMGT Sequence Alignment 3.15.0
SBT ambiguity impairs clinical utility

2-field HLA genotyping ambiguity for 95 patient samples typed by SBT

- Standard requires 4-digit typing for HSCT
- NMDP allele codes were developed to “soften” the ambiguity
  - ie. HLA-A*02:CVEG = HLA-A*02:01/02:24/02:101
Technical advances in HLA typing improving HSCT outcomes- next?

• 1960s- Serologic HLA typing
  • Only identifies limited polymorphisms (first field)

• 1980s- Molecular analysis
  • Sequence-specific oligonucleotide probes (SSO)
  • Sequence-specific primer PCR (SSP)

• 2000s- DNA sequencing based typing (SBT)

• Late 2010s- Next-generation DNA sequencing (NGS)
Summary

• Multiple molecular biology methods can be used to perform HLA genotyping

• Different DNA-based HLA typing methods have different benefits and limitations
  • SSP – simple, can use at high- or low-resolution, low-throughput
  • SSO – simple, limited resolution, high throughput
  • SBT – more complicated, true high-resolution, intermediate throughput

• The level of HLA genotyping and the method used should fit the clinical need
  • Solid organ transplant – genotype to level to predict serologic split antigens
  • Deceased donor HLA typing – need to type rapidly
  • HSCT - 2-field high-resolution is required
  • Disease association/Pharmacogenomics – often high-resolution
  • Integrated use of multiple methods can provide excellent results