Received: 16 August 2021 | Revised: 16 November 2021 | Accepted: 2 December 2021

DOI: 10.1111/tan.14505

ORIGINAL ARTICLE

HLA Immune Response Genetics WILEY

HLA-EPI: A new EPIsode in exploring donor/recipient epitopic compatibilities

Estelle Geffard[1] | **Léo Boussamet**[1] | **Alexandre Walencik**[1,2] | **Florent Delbos**[2] | **Sophie Limou**[1,3] | **Pierre-Antoine Gourraud**[1] | **Nicolas Vince**[1]

[1]Centre de Recherche en Transplantation et Immunologie, Université de Nantes, CHU Nantes, Inserm, Nantes, France

[2]Laboratoire d'Histocompatibilité et d'Immunogénétique, EFS Centre - Pays de la Loire, Nantes, France

[3]Département Informatique et Mathématiques, Ecole Centrale de Nantes, Nantes, France

Correspondence
Pierre-Antoine Gourraud and Nicolas Vince, CRTI UMR1064 – ITUN, CHU Nantes Hôtel Dieu, 30 bld Jean Monnet, 44093 Nantes Cedex 01, France.
Email: pierre-antoine.gourraud@univ-nantes.fr and nicolas.vince@univ-nantes.fr

Funding information
Conseil Régional des Pays de la Loire; H2020 Marie Skłodowska-Curie Actions, Grant/Award Number: 846520; Institut National de la Santéet de la Recherche Médicale

The HLA system plays a pivotal role both in transplantation and immunology. While classical HLA genotypes matching is made at the allelic level, recent progresses were developed to explore antibody–antigen recognition by studying epitopes. Donor to recipient matching at the epitopic level is becoming a trending topic in the transplantation research field because anti-HLA antibodies are epitope-specific rather than allele-specific. Indeed, different *HLA* alleles often share common epitopes. We present the HLA-Epi tool (hla.univ-nantes.fr) to study an *HLA* genotype at the epitope level. Using the international HLA epitope registry (Epregistry.com.br) as a reference, we developed HLA-Epi to easily determine epitopic and allelic compatibility levels between several *HLA* genotypes. The epitope database covers the most common *HLA* alleles ($N = 2976$ *HLA* alleles), representing more than 99% of the total observed frequency of *HLA* alleles. The freely accessible web tool HLA-Epi calculates an epitopic mismatch load between different sets of potential recipient-donor pairs at different resolution levels. We have characterized the epitopic mismatches distribution in a cohort of more than 10,000 kidney transplanted pairs from European ancestry, which showed low number of epitopic mismatches: 56.9 incompatibilities on average. HLA-Epi allows the exploration of epitope pairing matching to better understand epitopes contribution to immune responses regulation, particularly during transplantation. This free and ready-to-use bioinformatics tool not only addresses limitations of other related tools, but also offers a cost-efficient and reproducible strategy to analyze HLA epitopes as an alternative to *HLA* allele compatibility. In the future, this could improve sensitization prevention for allograft allocation decisions and reduce the risk of alloreactivity.

KEYWORDS
epitopic compatibility, HLA, HLA epitope, HLA eplet, HSCT, solid organ transplantation

Estelle Geffard and Léo Boussamet contributed equally to this study.

1 | INTRODUCTION

With a constantly increasing world population, average life expectancy and growing prevalence of chronic diseases (kidney, lung, etc.), more and more individuals will need organ transplantation during their lifetime. In 2018, around 150,000 transplantations were performed worldwide, representing 17 solid organ transplants per hour.[1] Currently, most graft allocation systems are partly based on generic *HLA* donor-recipient compatibilities, mainly for hyperimmunized patients.

HLA corresponds to the human major histocompatibility complex (MHC). It is a major component of self-recognition by the immune system and determines the tissue compatibility in organ transplantation. *HLA* genes show a high level of polymorphisms with more than 30,000 described alleles[2,3] listed in the IPD-IMGT/HLA Database (www.ebi.ac.uk/ipd/imgt/hla/about/statistics/). Currently, compatibility levels are evaluated from *HLA* genotypes by computing the number of shared *HLA* antigens between donor and recipient. Coupled to the use of immunomodulatory compounds, HLA allelic matching increases the chances of successful transplant and enables longer graft and overall survival.[4,5] However, this *HLA* allele-based allocation system presents several limitations. Among them, individuals with rare *HLA* genotypes or already immunized against one or several *HLA* alleles have drastically decreased chances to benefit from a compatible donor based on allelic compatibility, hence have less chances to get a good quality graft. Moreover, only pre-existing anti-HLA antibodies are taken into account for graft allocation but not the risk of developing new ones after transplantation (de novo donor-specific antibodies or DSA). As transplantations are limited by the number of available organs, it is of great importance to minimize the risks of graft rejection using all possible means. As an example, each year around 600 patients die in France because of lack of available compatible organs.[6] Improving the organ allocation system could therefore improve therapeutic options for rare *HLA* genotypes patients and prevent hyperimmunization in patients undergoing multiple grafts over time.

For the past few years, HLA epitopic matching has become a trending topic in the field of solid organ and hematopoietic stem-cell transplantation (HSCT), and potential clinical implementations are underway.[7,8] Indeed, despite their high level of polymorphisms, *HLA* alleles often share some epitopes, indicative of their respective levels of similarity. Epitope matching could therefore improve graft allocation. For example, in kidney transplantation, epitope load anticipation allowed to identify patients at risk for allosensitization and to adjust the immunosuppressive drug target levels.[8,9] In HSCT, a lifesaving therapy for several blood cancers, HLA epitope mismatches were associated with delayed engraftment in the case of HvG (host vs. graft), while increased HLA eplet mismatches give rise to a protective effect on the risk of relapse in the case of GvH (graft vs. host) in haplo-identical transplantation.[10]

Antibodies recognize specific surface regions of the antigens called structural epitopes, themselves containing functional epitopes.[11] Functional epitopes carry essential amino acid residues, that is, substitution of one of these residues would significantly decrease antibody affinity to the epitope.[12] These functional epitopes, composed of 1–5 amino acid residues, are eplets, mainly located within α1-α2 and α1-β1 domains of HLA class I and HLA class II molecules, respectively. Their highly polymorphic nature defines, at least in part, the different *HLA* alleles.[13] An eplet can either constitute linear or three-dimensional epitopes, and can be located in cryptic or exposed areas of the HLA proteins (Figure 1A).[14] Moreover, some eplets are shared by several *HLA* alleles, from the same or different gene, while some are unique to one *HLA* allele. As such, HLA eplet mismatch load between donor and recipient can be determined to refine the allocation system. It might provide better-suited transplants, requiring less immunosuppressive treatment in transplanted patients[8] and could unravel compatible grafts for highly immunized or rare *HLA* genotypes patients, overall aiming to increase patients' survival chances.

Several informatics programs have become available to study epitopic matching between donor and recipient.[15] HLAMatchmaker[16] is used to identify three-dimensional eplets on the HLA molecular surface toward which antibodies can be directed. This Excel-formatted program for HLA antibody analysis and eplet-based matching defines an eplet according to differences within repeating triplets of linear amino acid sequence. The OneLambda Incorporated (OLI) Fusion MatchMaker software is based on HLAMatchmaker. It uses the same database and the same calculation process but also includes additional eplet data, and offers a user-friendly eplet calculation tool. HLA-EMMA[17] analyzes HLA class I and class II compatibilities between donor and recipient on amino acid level focusing on the antibody-accessible amino acid mismatches, with both single or batch analysis options. The PIRCHE algorithm was established to identify peptides presented by antigen-presenting cells (APC) to recipient CD4 T-cells after digestion of HLA protein residues.[18,19] In this case, alloantigen recognition is qualified as indirect since the APCs belong to the recipient. In contrast, HLAMatchmaker focuses on direct alloantigen recognition. In HSCT, PIRCHE algorithm compares HLA protein fragments of the recipient potentially detected by donor CD8+ (presented by shared HLA class I) and CD4+ (presented by shared HLA class II) T cells and

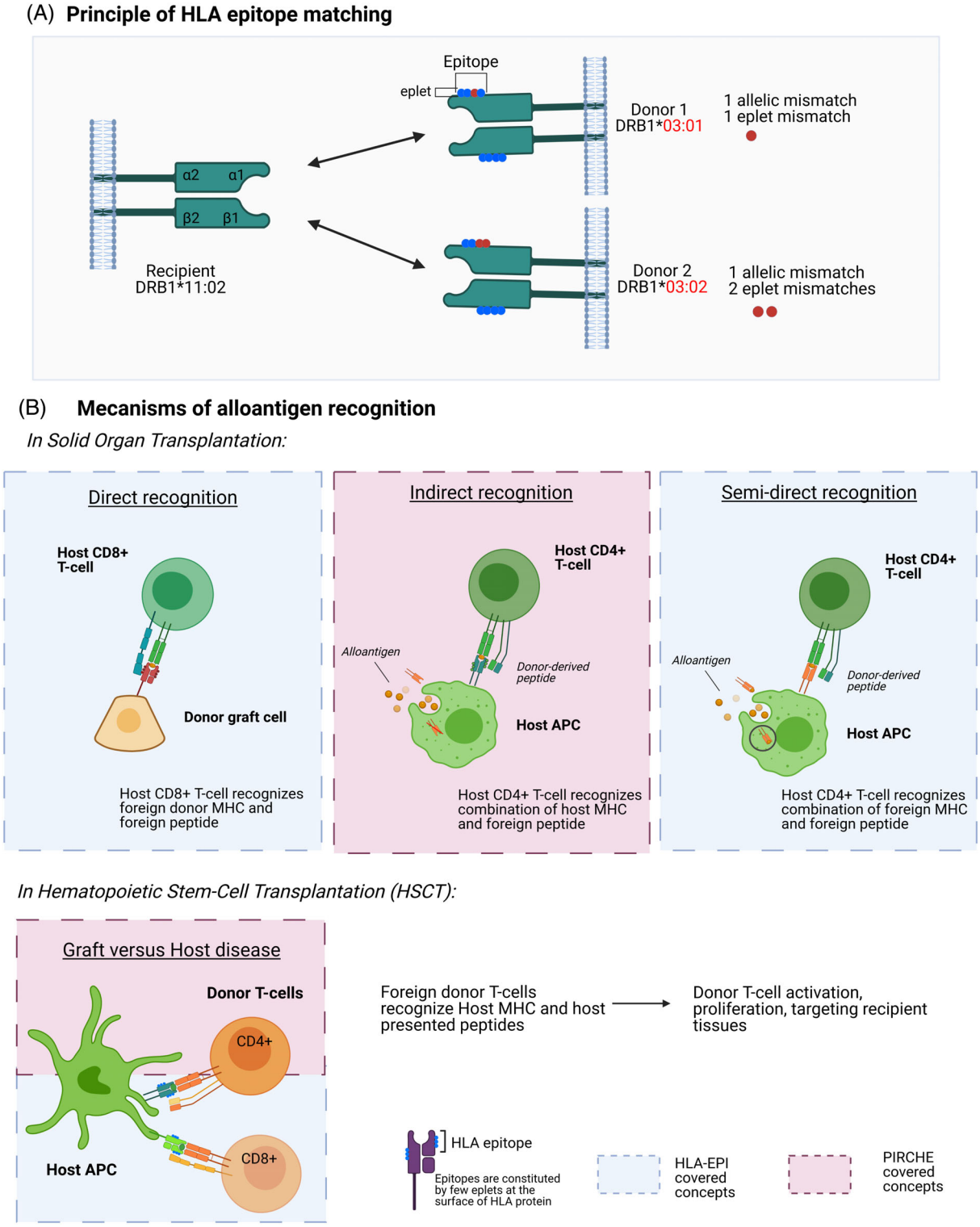

FIGURE 1 (A) Principle of epitope matching. Example of two different potential grafts: *HLA-DRB1*11:02* recipient may receive an organ from two distinct donors with different HLA-DRB1 alleles: 03:01 or 03:02. There is a single allelic mismatch in each case but *HLA-DRB1*03:01* (donor 1) shows higher compatibility in terms of eplet mismatch load compared with *HLA-DRB1*03:02* (donor 2). In blue: Matched eplets between the different individuals, in red: Mismatched eplets, in green: HLA Class II proteins. (B) Principles of alloantigen recognition covered by the HLA-Epi tool: HLA-Epi covers the concepts of direct and semi-direct recognitions while the PIRCHE algorithm focuses on indirect recognition for solid organ transplantation. In the case of HSCT, both HLA-Epi and PIRCHE algorithms cover the Graft versus host disease. (Created with BioRender.com)

considered as mismatched. In solid organ transplantation, PIRCHE considers peptide mismatches from the donor presented by recipient's HLA class II molecules to CD4+ T cells.[19]

Here, we propose a new tool to study HLA compatibility based on eplets direct recognition using our new Easy-HLA web suite tool: HLA-Epi. In the present work, we aimed at creating a user-friendly and freely available web tool to improve the study of donor/recipient tissue compatibility levels with an up-to-date public database. Our tool allows the complete study of a genotype regardless of its resolution (low resolution or one field, e.g., *A*01*, high resolution, or two fields, e.g., *A*01:01*). Figure 1B summaries alloantigen recognition mechanisms covered by the HLA-Epi and PIRCHE algorithms while highlighting their differences.

HLA-Epi has the ambition to become a ready-to-use tool for the graft allocation system, and is currently oriented for research applications. HLA-Epi is based on the international reference eplet registry database: epregistry (www.epregistery.com.br).[14]

2 | MATERIALS AND METHODS

2.1 | HLA-Epi, a scalable tool

Epitopic compatibilities are defined by mismatches between donor and recipient *HLA* genotypes; a mismatch being defined by an eplet brought by the donor and unknown to the recipient. HLA-Epi allows to study epitopic mismatch between one *HLA* genotype compared with either one or more *HLA* genotypes (up to 30 can currently be added). Depending on downstream applications, one might need to compare a recipient to a single or several potential donors (e.g., in the case of HSCT or living donor kidney transplant), whereas others would search to compare a donor to one or several potential recipients (e.g., in solid organ transplantation). To adapt for these needs, HLA-Epi was made scalable with the possibility to switch donors and recipients. HLA-Epi allows the study of *HLA-A*, *-B*, *-C*, *-DRB1*, *-DQB1*, *-DQA1*, and *-DPB1* genes, and in addition, high and low-resolution *HLA* genotypes can also be evaluated. Our tool offers two main research modes: 1) via a web interface where genotypes are written manually or 2) via a multiple search option where several pairs of genotypes are submitted within a file.

2.2 | HLA-Epi development

2.2.1 | Web development

HLA-Epi is integrated into the Easy-HLA website (hla.univ-nantes.fr).[20] We developed this new tool using HTML5, PHP7.4 (www.php.net), Javascript1.8.5 languages and implemented a PostgreSQL database (www.postgresql.org). This eplet database is queried through a PostgreSQL database management system and will gradually be updated as new eplets are identified.

Easy-HLA website and related database are hosted on Nantes Université secure server. All reference data are anonymized and coming from public database. Database access is restricted to authorized person only. Submitted data are deleted immediately after analysis completion and the output data files are safely conserved on our server for 1 week. Source code and used data to generate our results are available on the following git repository https://gitlab.univ-nantes.fr/crtiteam5/easy-hla).

2.2.2 | Eplets database

The HLA-Epi tool is based on a public database including 560 distinct eplets correspondences carried by 2976 *HLA* alleles coming from EpRegistry3.0, the International Registry of HLA Epitopes (http://www.epregistry.com.br).[21] Among them, 154 have been antibody checked. EpRegistry represents a rich resource to study histocompatibility at the epitope level and antibody responses against HLA mismatches in transplant patients. Overall, although 10% of the total described *HLA* alleles are represented in the eplet database, it represents more than 99% of the *HLA* alleles observed frequency. This database contains epitopes carried by proteins coded by seven *HLA* genes *HLA-A*, *-B*, *-C*, *-DRB1*, *-DQB1*, *-DQA1*, and *-DPB1*. Depending on the user's goal, different epitopes properties such as exposed, cryptic, antibody checked and unchecked eplets can be considered. Exposed eplets correspond to the accessible ones to compounds present in the media such as antibodies. These eplets consequently have a greater immunogenicity risk compared with cryptic eplets. Eplets are annotated with the position of its first constituting amino acids, followed by the amino acids names symbol (e.g., 71QS). In the present work, we characterized the epitope EpRegistery database, through the use of our developed web tool.

2.2.3 | Eplet mismatch load calculation

The eplet mismatch computations are performed for each class I (*HLA-A, -B, -C*) and class II (*HLA-DRB1, -DQB1, -DPB1*) *HLA* alleles given as inputs, counting each eplet carried by the donor but unknown to the recipient as one mismatch. In the HvG way (default way for solid organ transplantation), the mismatched eplets are given by the donor and not carried by the recipient: mismatched eplets present at the cell surface of the graft might be

immunogenic, and activate recipient's immune response. On the opposite, in the GvH way, mismatched eplet belong to the recipient but are unknown to the donor: in this HSCT context, immune cells from the graft may react against recipient tissues. Consequently, results are dependent of the way of calculation: HvG vs. GvH. We display class I and class II mismatch numbers ($MMCL_I$ and $MMCL_{II}$), as well as the overall mismatch load (MMload) as the sum of $MMCL_I$ and $MMCL_{II}$ annotated "score" in the tool. MMload indicates global compatibility levels between donor and recipient. The main calculations are summarized in the following equations:

$$MMCL_I = \sum (epletsI_d \notin epletsI_r)$$

$$MMCL_{II} = \sum (epletsII_d \notin epletsII_r)$$

$$MM_{load} = MMCL_I + MMCL_{II}$$

Equation: Calculation equations for the main output. $MMCL_I$: number of class I eplet mismatches, $MMCL_{II}$: number of class II eplet mismatches, MMload: mismatch load, $epletsI_d$: eplets harbored by the class I HLA alleles belonging to the donor, $epletsI_r$: eplets harbored by the class I HLA alleles belonging to the recipient.

All calculations are made through specific PHP functions calling SQL requests and retrieving results from the PostgreSQL eplet database. High resolution HLA genotype imputation function is using previously developed HLA-upgrade tool. High resolution genotypes are inferred using the National Marrow Donor Program haplotype frequencies database published in 2013 for uses in clinical transplant and immunological research.[22]

2.3 | Distribution of eplet variability study

In order to characterize the eplet distribution variability in different populations, we ran HLA-Epi on two different datasets: The 1000 Genomes (1KG) public database[23] and a local clinical database of kidney transplanted patients (DIVAT, www.divat.fr). Data from the DIVAT cohort are freely available to academic researchers and were obtained after approval by its scientific board (http://www.divat.fr/access-to-data). All donors were informed of the final use of their blood and signed an informed consent form.

First, we prepared datasets to analyze the number of mismatch distribution between two *HLA* genotypes. (1) We retrieved high-definition HLA genotypes from three large ancestral populations from the 1KG database: European ($n = 332$), East Asian ($n = 335$) and African ($n = 388$).[24] We simulated random sampling genotype pairs with replacement to create 5000 hypothetical donor-recipient pairs within each population (subdivided in 10 groups of 500 pairs). (2) In parallel, 10,667 *HLA* genotypes from renal transplant donor-recipient couples were extracted from the DIVAT database to assess eplets in a real practice donor-recipient matching setting. These pairs were first upgraded through the HLA-upgrade tool[20] to get high-resolution *HLA* genotypes.

Data from simulated pairs, along with real data pairs from DIVAT, were formatted for the HLA-Epi workflow; we used the multiple search mode and selected all eplet types (exposed, cryptic, checked and unchecked). Outputs provided by HLA-Epi were then analyzed and visualized with the R statistical analysis software (https://www.R-project.org/). Analysis of variance (ANOVA) and pairwise *t* test comparisons were performed between the distribution curves in terms of epitopic mismatch load obtained for all the tested populations. In parallel, linear regressions between allelic mismatch counts and epitopic mismatch counts were run to evaluate correlation level for these variables in the DIVAT data.

2.4 | Comparison to other related tools

We validated HLA-Epi by comparing the calculated epitopic HLA class I and class II mismatch scores against two related tools: HLAMatchmaker 3.1, Excel-formatted program for HLA antibody analysis and eplet-based matching based on EpRegistry (as HLA-Epi), and PIRCHE, algorithm established to identify peptides presented by APC to recipient CD4 T-cells after digestion of HLA protein residues. We considered *HLA-A*, *-B*, *-C* alleles for class I *HLA* genes and *HLA-DRB1* and *-DQB1* alleles for class II *HLA* genes from 1000 randomly assigned pairs from 1KG European ancestry. The mismatch scores were calculated in the HvG way. Linear regression analyses were performed to compare similarity between the different scores. For HLA-Epi and HLA_Matchmaker, both class I and class II mismatch scores as well as the only given PIRCHE output (corresponding to PIRCHEII score for HvG) were retrieved and compared HLA-Epi mismatch scores (class I, class II, and class I + class II).

3 | RESULTS

3.1 | Eplet database characterization

The EpRegistry database includes correspondences on 560 distinct eplets carried by 2976 *HLA* alleles. Although this database covers only 10%–15% of all described *HLA*

alleles (30,862 known to date, April 2021, https://www.ebi.ac.uk/ipd/imgt/hla/stats.html), it represents more than 99% of the overall *HLA* alleles observed frequency. Only rare alleles, with frequency below 10e-6, are not represented in the database. Characteristics of the eplet database are summarized in Table 1. "Unique" eplets are exclusively present on one *HLA* gene; in other words, a unique eplet will only appear on some alleles within a gene but not on any other *HLA* gene. On the other hand, some eplets are not only shared between different alleles of the same *HLA* gene but also between different *HLA* genes (Figure 2). Interestingly, inter-gene disparity in eplet composition is higher in Class II *HLA* genes with only a few eplets shared between at least two class II *HLA* genes (Figure 2B).

3.2 | HLA-Epi searching web tool

In manual mode (Figure 3), *HLA* genotypes are directly typed in the tool interface. Several options can be selected corresponding to exposed, cryptic, checked or unchecked eplets. It is possible to compare a genotype against one or several genotypes by adding up to 30 different genotypes. HLA-Epi can provide recipient to one or several donors comparisons but this can alternatively be switched to donor compared with several potential recipients.

By default, HLA-Epi takes *HLA-A, HLA-B, HLA-C, HLA-DRB1*, and *HLA-DQB1* alleles as inputs. Missing or low-resolution alleles are automatically detected and upgraded with the HLA-Upgrade algorithm after selecting a reference population (ancestry). Only the most probable *HLA* genotype is kept for an individual corresponding to the best result of the HLA-Upgrade tool.[18] Since *HLA-DQA1* and *HLA-DPB1* alleles are often missing in investigation studies and because HLA-Epi is not able to impute them, these particular alleles are only proposed in the "extended" mode; clicking the "Extended" button will add them in the input section. When a genotype is not known, the field must be left blank.

Results are directly output within the interface (Figure 4). Numbers of eplet mismatches and eplet details are shown for each *HLA* allele of each gene and each class. Indeed, all eplets brought by the donor and unknown by the recipient will be considered as mismatches (see Section 2). If data is missing for one gene, then the output and the mismatch load score will be noted as "NA" in order to only keep comparable results.

Figure 4A presents the output obtained for Figure 3's input. In that example, we observe that donor 1 and donor 2 both have a total of 5 allelic mismatches (Figure 3). However, at the epitopic level, donor 2 appears to show a higher number of shared eplets with the recipient (21 eplet mismatches vs. 46 for donor 1) and might therefore be a better candidate in terms of HLA eplet similarity. In Figure 4B, we launched the same *HLA* genotypes by switching the donor and recipient status. Indeed, in the case of HSCT, it is important to evaluate similarities in both graft versus host and host versus graft ways. This resulted in a significant change on the calculated eplet mismatch score with for instance 21 eplet mismatch load to 42 eplet mismatch load for the same *HLA* genotypes pairs when only switching donor to recipient (Figure 4B).

Alternatively, a file containing genotypes to test can be used, especially if more than 30 individuals are to be considered simultaneously, using our multiple search mode. Symmetrically to the manual mode, upgrade of missing or low-resolution data can be selected and the output file contains all mismatch computation scores. File specifications are detailed in the HLA-Epi online tutorial.

3.3 | Distribution of eplet variability analysis

Using the HLA-Epi tool, we have studied the distribution and average of the eplet mismatch number for two genotypes chosen at random from the 1KG populations and from a real-life cohort of kidney transplanted patients

TABLE 1 Eplet compositions of the 2976 HLA alleles represented in the eplet epRegistry database

HLA genes	Alleles	Total eplets	Exposed	Checked
A	607	195	158	56
B	989	199	158	60
C	334	107	76	21
DRB1	858	142	70	34
DQB1	125	60	34	17
DQA1	29	25	11	3
DPB1	34	49	24	11
Total	2976	777 (560 unique[a])	531 (374 unique[a])	202 (154 unique[a])

[a]Eplet appearing only in alleles belonging to the same HLA gene are described as "unique" eplets.

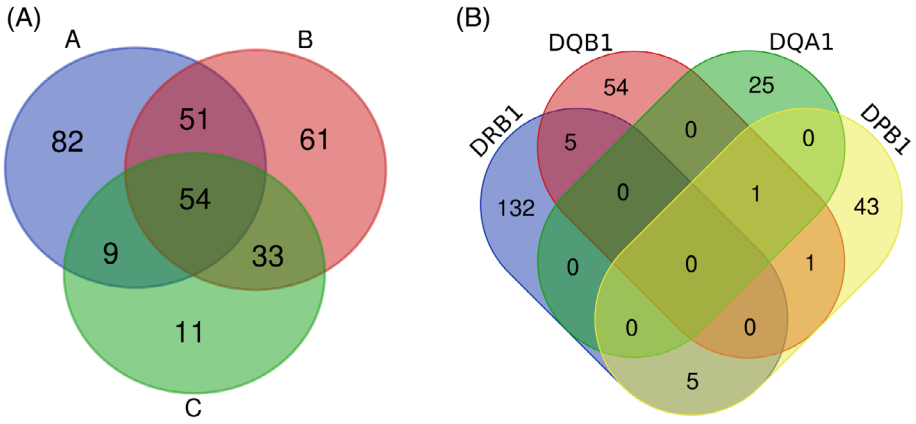

FIGURE 2 (A) Shared eplets between the HLA class I genes, HLA-A (blue), HLA-B (red) and HLA-C (green). A large number of eplets are shared between HLA class I genes, with 54 eplets shared between the 3 Class I genes. (B) Shared eplets between the HLA class II genes, HLA-DRB1 (blue), HLA-DQB1 (red), HLA-DQA1 (green) and HLA-DPB1 (yellow). Only a few eplets are shared between class II HLA genes

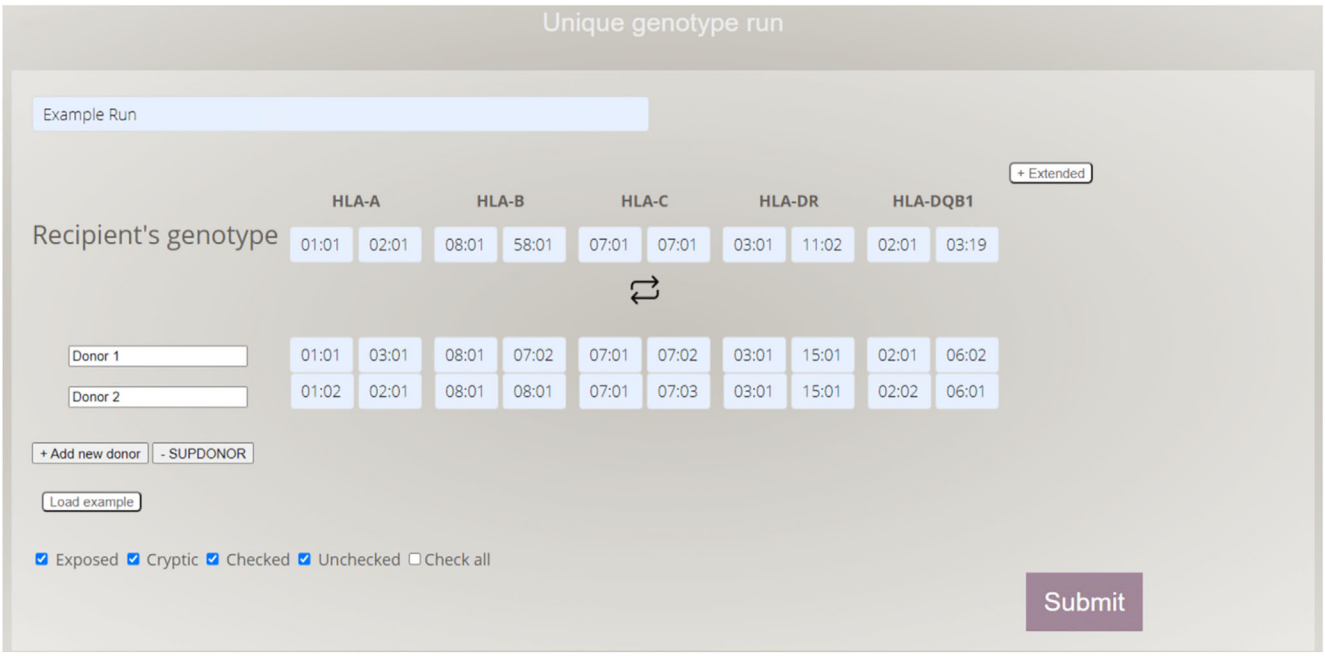

FIGURE 3 Input interface in manual entry mode. Recipient's and donor's genotypes need to be written manually in the different fields. Users can select to query all, exposed, cryptic, checked or unchecked eplets. HLA-Epi can provide recipient to one or several donors comparison but this can be easily switched to donor to several potential recipients by clicking the arrow button. Up to 30 genotypes can be added for comparison by clicking the "add new donor" button. Extended version comprising HLA DQA1 and DPB1 is finally available with the "Extended" button

(pooled results are summarized in Figure 5). Regarding the general populations, ANOVA ($p < 2e\text{-}20$) and pairwise *t* tests showed different distributions (Figure 5, gray lines) between the three simulated ancestral populations from The 1000 Genomes project (1KG). The African pairs exhibited on average 71.3 [0–152] mismatches, while the Asian pairs showed an average of 69.7 [0–166]. European pairs exhibited the highest mean with 73.6 [0–160] eplet mismatch load on average. When comparing with the computations run in real kidney transplant pairs from the French population, we observed lower values with 56.9 mismatches on average [0–158] (red distribution on Figure 5) illustrating the beneficial effect of the donor allocation system partly based on *HLA* allelic mismatches showing on average a decrease of 15 epitopic mismatch compared with random allocation ($p < 2e\text{-}16$). In the same line, we observed a bimodal distribution in the kidney transplantation cohort with a peak for 0 mismatch, which corresponds to 347 fully matched pairs. Among these, 280 pairs had fully identical *HLA* genotypes (siblings) whereas 67 pairs appeared to have one or more allelic mismatches.

Figure 6 illustrates the level of disparity between the different metrics (allelic level and epitopic level). Indeed, although strong correlations were found between the number of *HLA* mismatches at allelic level to its corresponding epitopic level (R^2 = [0.5–0.6]), large repartition around

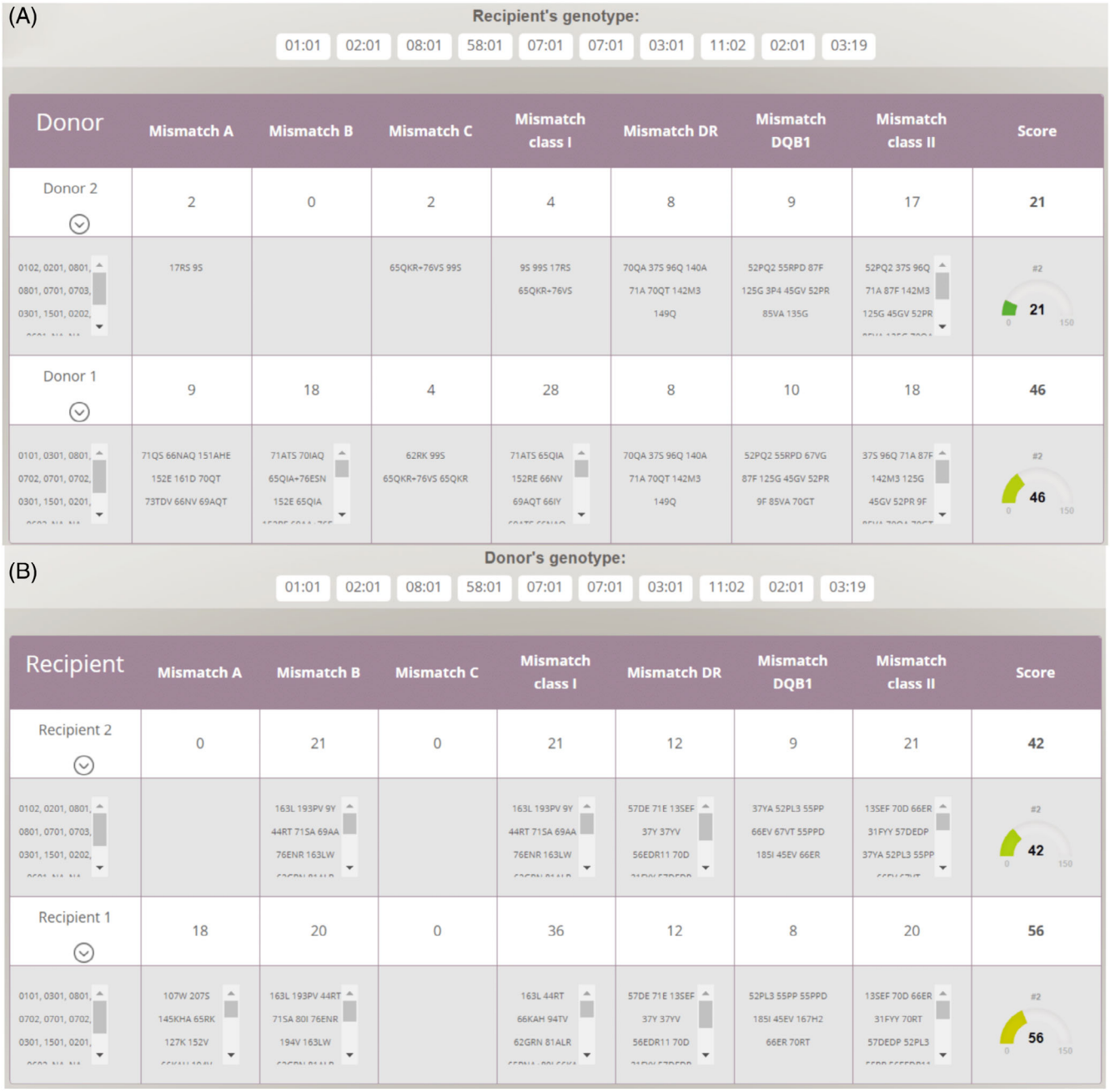

FIGURE 4 Example run and results page. For each potential donor, the different eplet mismatches calculated for each HLA gene (HLA-A, HLA-B, HLA-C, HLA-DRB1, HLA-DQB1) and for each HLA class (class I and class II) are displayed on the first line (white line). Details of the mismatched eplets names appears on the second line (gray line). The last column corresponds to the global compatibility score displayed in a gauge. (A) Results page for the recipient to donors direction; (B) Results page for the donor to recipients way. Same genotypes as A were kept but the reference recipient becomes a reference donor and potential donors become potential recipients

median value were also found. For instance, in HLA class I regression curve (Figure 6A) for 6 allelic mismatches, epitopic mismatches range from 9 to 109. Moreover, slope appears to be higher for class I *HLA* (Figure 6A, coefficient = 8.5), compared with class II *HLA* (Figure 6B, coefficient = 7.4). Finally, no association was found between HLA interclass (allelic mismatches class I vs. class II, or epitopic mismatches class I vs. class II).

3.4 | HLA-Epi comparison with HLAMatchmaker and PIRCHE software

We created a dataset of 1000 randomly assigned European-ancestry donor-recipient pairs from 1KG dataset and compared our HLA-Epi tool results with HLAMatchmaker and PIRCHE. Our HLA-Epi tool shows high similarities with HLAMatchmaker (Figure 7A,B);

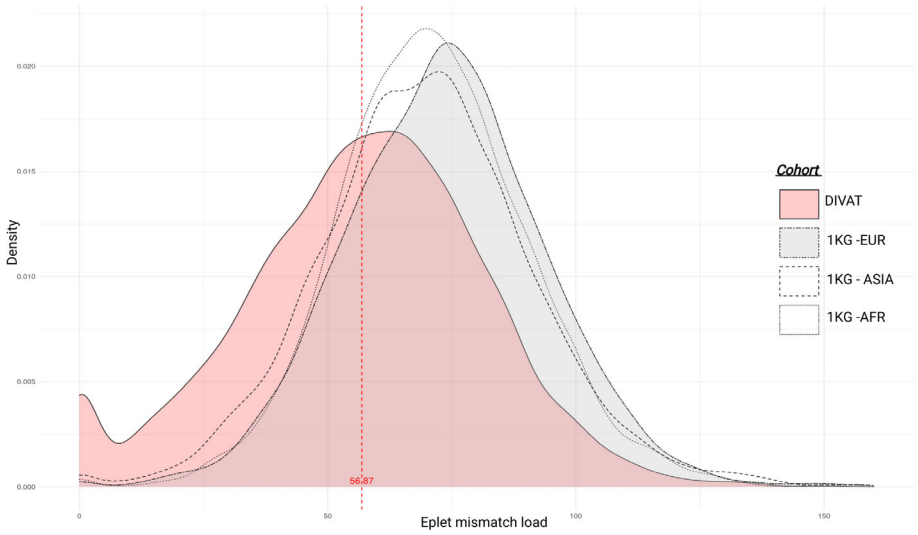

FIGURE 5 Eplet mismatch load distribution. In gray, 1KG simulations mismatch load distribution from three tested populations (European, Asian, and African), 5000 donor-recipient pairs were simulated and mismatch load calculation was performed based on their high-definition HLA genotypes. In red, real eplet mismatch load distribution (from 10,667 DIVAT cohort kidney transplanted individuals, red curve)

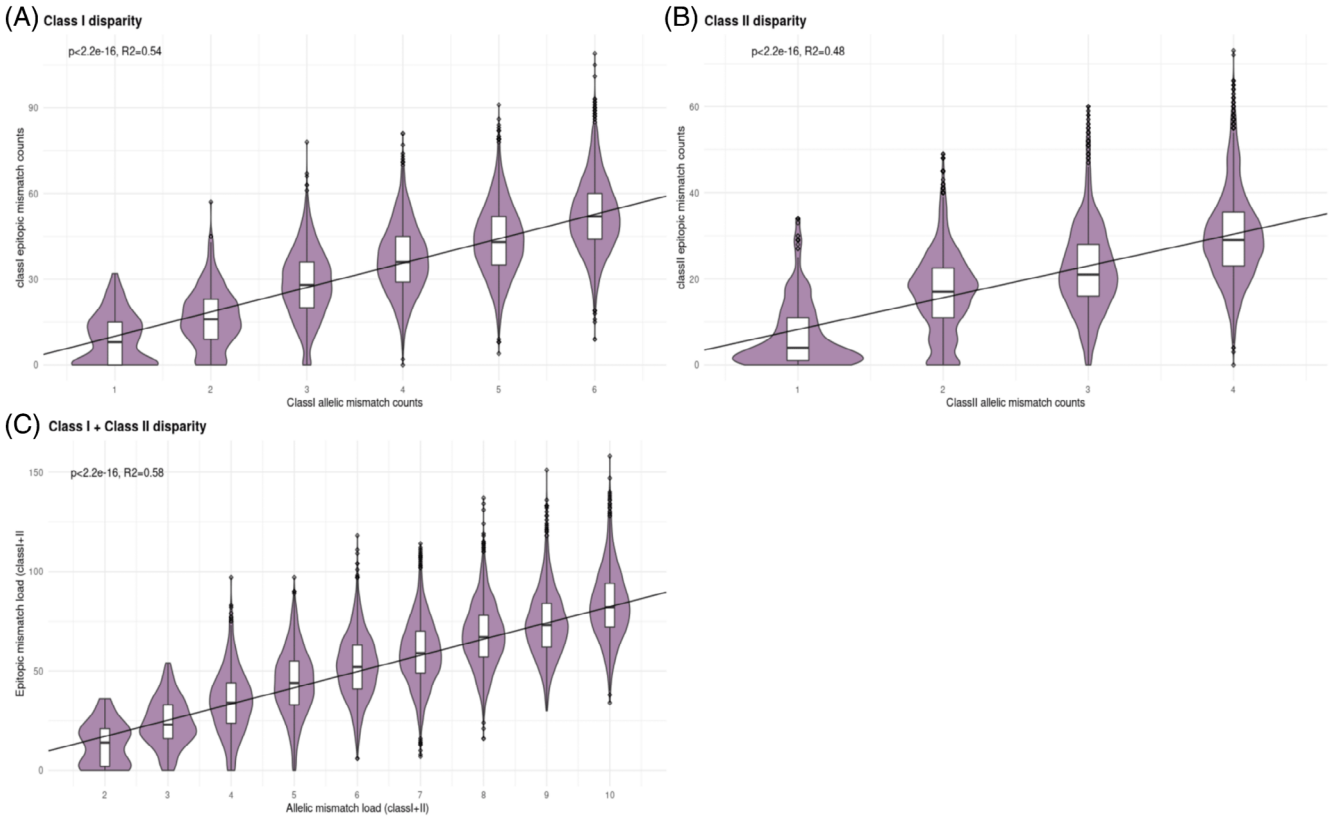

FIGURE 6 Disparity between allelic level and epitopic level. Regression curves were obtained for (A): Class I (HLA-A, HLA-B, HLA-C) allelic mismatches versus Class I HLA epitpic mismatches, (B): Class II (HLA-DRB1, HLA-DQB1) allelic mismatches versus Class II HLA epitopic mismatches and (C): allelic HLA mismatch load (Class I + Class II) versus HLA epitopic mismatch load

indeed, linear regression shows strong correlation for both class I as well as class II epitopic mismatch scores (R^2 = 0.77 and 0.84 respectively, p < 2e-16). HLA-Epi mismatch scores appear to be higher compared with the corresponding HLAMatchmaker scores (slope = 1.99 and 1.27 for class I and class II, respectively). Correlation was weaker between HLA-Epi and PIRCHE (Figure 7C, R^2 = 0.15, p < 2e-16). HLA-Epi considers epitopes composition of 1930 class I and 983 class II *HLA* alleles whereas HLAMatchmaker only considers eplet composition for 1900 class I and 700 class II *HLA* alleles. We selected 3 examples from the HLA-Epi/HLAMatchmaker

comparison for the class II results. First, for fully matched pairs (Figure 7B, case 1), HLAMatchmaker class II score is always one epitopic mismatch. Second, HLA-Epi score can be lower than HLAMatchmaker score (Figure 7B case 2, HLA-Epi = 17, HLAMatchmaker = 22) when some alleles coming from the recipient are unknown to the HLAMatchmaker algorithm. Indeed, in Figure 7B case 2, no eplet arise from the unknown *HLA* allele of recipient and consequently increase the number of unknown eplets from the donor. In addition, HLAMAtchmaker can sometimes count 2 mismatches for the same epitope, when HLA-Epi only accounts for 1 (e.g., 74SV for HLA-Epi becomes 74S and 75 V in HLAMatchmaker). Finally, third and most common case is when the HLA-Epi score is higher than HLAMatchmaker (Figure 7B, case 3, HLA-Epi = 53, HLAMatchmaker = 28), HLAmatchmaker tool does not consider the latest version of the eplet database. Newly characterized eplets are not present in

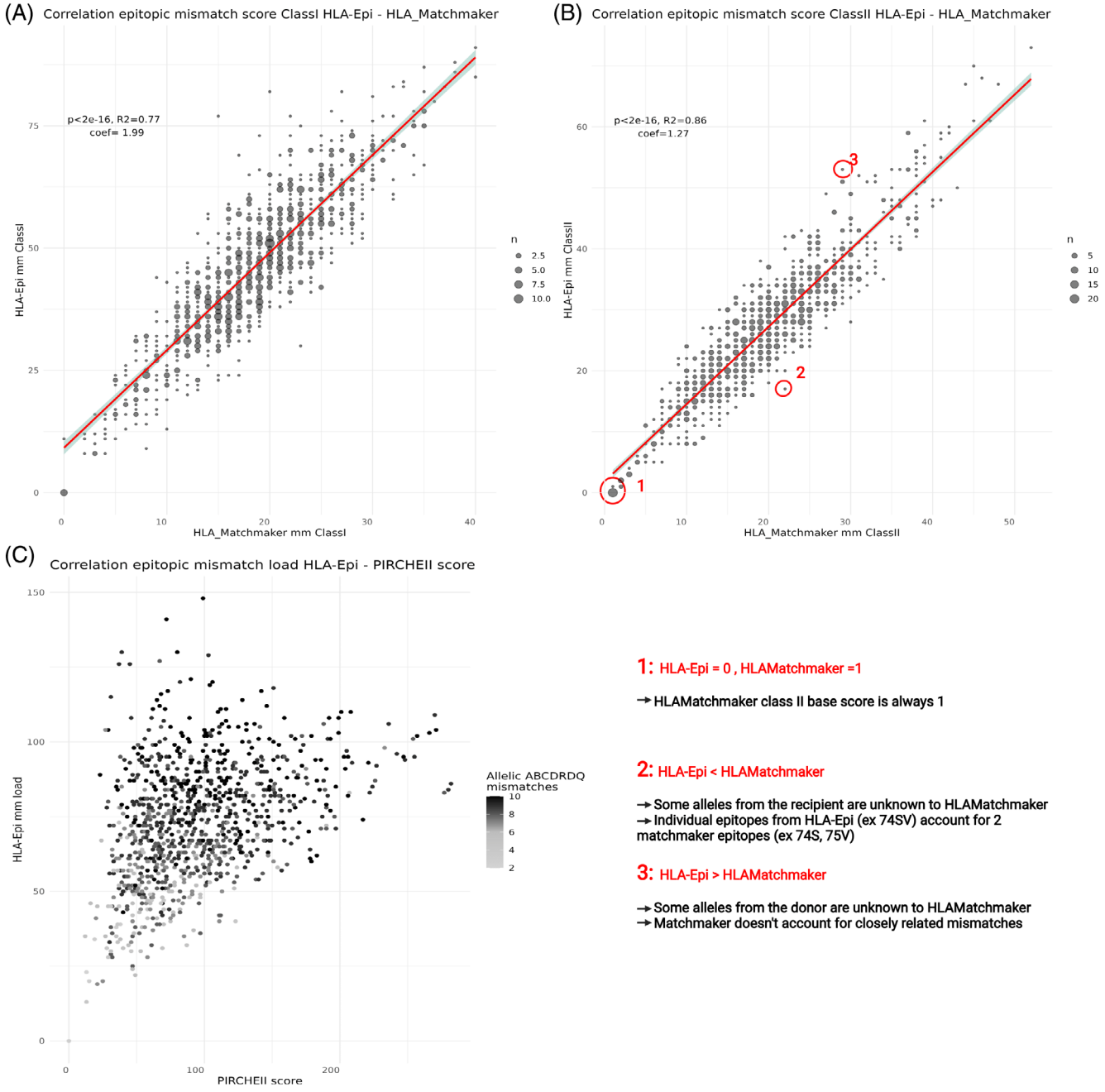

FIGURE 7 Comparison HLA-Epi to other epitope-matching related tools. Correlation between HLA-Epi scores and (A): HLAMatchmaker class I mismatch score, (B): HLAMatchmaker class II. Three scenarios were considered to further investigate the differences between tools: 1: HLA-Epi = 0, HLAMatchmaker = 1, 2: HLA-Epi < Matchmaker, 3: HLA-Epi < Matchmaker. (C): PIRCHEII epitopic mismatch score. Dots were colored according to the number of HLA allelic mismatches

HLAmatchmaker, hence, less possible eplet mismatches can be characterized. This global underestimation of eplet mismatches in HLAMatchmaker had previously been reported for version 2.1[14] and appears to remain in our current study on version 3.1. Other differences arise, epitopes showing close properties (for instance a single switch between two acid residues) are not counted as mismatches in HLAMatchmaker. Finally, unknown alleles to HLAMatchmaker coming from the donor can also give a lower score compared with HLA-Epi.

4 | DISCUSSION

We have created the HLA-Epi tool, embedded within the Easy-HLA web suite, to assess HLA epitopic compatibilities from *HLA* genotypes of different individuals. Easy-HLA is a freely accessible and user-friendly web application suite (hla.univ-nantes.fr) aiming at facilitating *HLA* genotype analyses without requiring high-level computer skills. In addition to HLA-Epi, the Easy-HLA platform currently offers three complementary tools: (1) EasyMatch-R to assist unrelated HSCT donor search from the *HLA* patient's genotype; (2) HLA-Upgrade to update missing or low-resolution *HLA* genotypes to high-resolution; (3) HLA-2-Haplo to predict pair(s) of haplotypes corresponding to an *HLA* genotype.[18]

HLA-Epi offers a simple interface with a high flexibility to adapt to researchers' needs. HLA-Epi delivers the number of eplet incompatibilities between donor-recipient pairs for each *HLA* gene and each *HLA* class. It also provides the overall eplet mismatch load, as well as details on mismatch eplets. Importantly, the possibility to impute missing or low-resolution *HLA* genotypes with HLA-Upgrade algorithms is a non-negligible asset for HLA-Epi, even if *HLA-DQA1* and *HLA-DPB1* genes are currently not covered in our reference haplotypes database. Although imputation may come with some uncertainties (86%–97% accuracy for low to high resolution imputation found in Geffard et al.,[18]), high resolution imputation is an important feature of HLA-Epi as many historical cohorts were typed in low resolution or with missing alleles, their epitopic HLA potential can now be investigated. Regarding applications, HLA-epi allows studying epitopic compatibility between a recipient's *HLA* genotype and potential(s) donors' genotype(s) in the HSCT context. Reciprocally, HLA-epi can compare epitope compatibility between a donor's genotype and potential(s) recipients' genotype(s) in the solid organ transplant context. Currently, the calculated compatibility score only considers the overall eplet mismatches, and not mismatch immunogenicity. Indeed, replacing an eplet amino acid with one showing similar physicochemical properties will have a lower impact on the HLA protein 3D microenvironment and thus on the mismatch immunogenicity, compared with a non-similar amino acid. In a near future, we therefore ambition to refine this mismatch load score in order to include eplet physicochemical properties (isoelectric point, steric hindrance).

Tools targeting HLA epitopic mismatches were previously developed. Among them HLAMatchmaker, HLA-EMMA, or PIRCHE can be mentioned.[16–18] All of them share similarities and present differences with HLA-Epi. Regarding epitopic targets, HLA-Epi and HLAMatchmaker both focus on eplet mismatches and rely on different versions of the same database (Epregistry). As a consequence, strong correlations between the two tools were observed (Figure 7A,B). However, most of the time, HLA-Epi counted more mismatches compared with HLAMatchmaker. HLA-Epi covers more HLA alleles than HLAMatchmaker (1930 vs. 1900 class I alleles, and 983 vs. 700 class II alleles, respectively) resulting in more potential mismatches. Second, HLAMatchmaker does not count some closely related epitopes as mismatches (e.g., 71AK → 71AR) contrary to HLA-Epi. Our tool therefore addresses some of HLAMatchmaker limitations. First, its spreadsheet-based interface makes it very heavy and slow as soon as more than 10 pairs are simultaneously evaluated. Second, HLAMatchmaker is limited to donor-recipient pairs only and epitopic mismatches are only made in the Host versus Graft way. Finally, only high-resolution genotypes can be evaluated and missing alleles will be counted as mismatches when our tool also offers the possibility to upgrade low-resolution HLA genotypes. The PIRCHE algorithm aims at predicting presented HLA peptides after HLA proteins lysosomal digestion by APC cells. HLA-Epi mismatch load and PIRCHEII score showed weak correlations (Figure 7C), which was expected as both tools investigate HLA mismatches in different immune recognition contexts. Both tools therefore appear complementary as illustrated by Figure 1B. Contrary to the paid-license PIRCHE, HLA-Epi has the advantage of being a free tool, simply available online.

HLA-Epi uses the latest scientific evidence regarding tissular compatibility. HLA epitopes are both involved in the direct and indirect recognition by the host's immune system. Indeed, after transplantation, donor's APC from the graft will migrate to the recipient's lymphatic nodes. Host CD8+ T-cell will recognize these foreign donor cells via their HLA class I and foreign presented peptides. This mechanism, named direct alloantigen recognition, can lead to acute rejection.[24] On the other hand, in indirect recognition, alloantigens (including HLA derived peptides) from the graft are internalized, processed and presented at the surface of the recipient's APC. This can lead to T cell-mediated chronic rejection. As summarized in Figure 1B, while PIRCHE algorithm addresses this last concept, HLA-Epi deals with the direct

recognition mechanism as well as the semi-direct recognition. In the context of semi-direct recognition host APC can present peptides through an intact HLA molecule arising from the donor (Figure 1B). Therefore, the same molecule will be found at the cell surface as in direct recognition, only the APC belong to the host. Thus, as we only consider the *MHC* region, our calculations consider both direct and semi-direct recognition patterns in the same way. Furthermore, both PIRCHE and HLA-Epi tools cover the concept of graft-versus-host disease in HSCT.

As shown in Figure 2, several eplets are shared between different *HLA* genes, especially for class I *HLA* genes whereas class II *HLA* genes only share a very limited number of eplets. This observation could explain why class II HLA eplet mismatches appear to be more immunogenic than class I eplet mismatches.[25] At the opposite, class I *HLA* carry higher number of eplets than class II *HLA*, this could explain the higher slope observed in Figure 6: class I *HLA* have higher number of eplets for a given number of mismatches compared with class II *HLA*. Although the epitope matching concept appears promising for assessing graft allocation, some information is still lacking, especially regarding immunogenicity of each individual eplet, which is not completely captured by the current definition of exposed vs. cryptic eplets (the latter being less immunogenic). Two donor-recipient combinations with the same number of epitopic mismatches may indeed exhibit very different immunogenicity schemes. Moreover, it is highly probable that immunogenicity of an individual eplet will vary from one recipient to another and thus must be put in the context of the recipient. As a major objective of the 18th international HLA and Immunogenetics Workshop taking place in May 2022, intensive research is currently carried out assessing eplet mismatches immunogenicity and trying to validate new metrics using physicochemical properties of eplets amino acids.[26,27]

Finally, we ran HLA-Epi to assess the epitopic mismatches distribution between two *HLA* genotypes in three large general populations and one population of kidney-transplanted patients. On average, around 70 mismatches were observed between 2 individuals randomly picked within each ancestral general population. In our European cohort of kidney transplanted patients, we observed the lowest number of mismatches with an average of 57 epitopic mismatches between recipient/donor genotype pairs. This lower number was expected in a context of kidney transplantation where the allocation system notably aims at limiting the number of *HLA* allelic mismatches as much as possible. Moreover, only in the real-transplanted data, a zero peak can be observed corresponding to 347 fully epitopic-matched pairs. Among those, 280 pairs, likely siblings, had fully identical HLA genotypes whereas 67 pairs appeared to have one or more allelic mismatches. This illustrates that in some cases perfect epitopic matching can be achieved without having a perfect allelic matching and thus emphasize the interest of epitopic matching. Surprisingly, Asian and African-ancestry simulated pairs exhibited significantly less mismatches on average compared with simulated European population. Indeed, very rare *HLA* genotypes (frequency less than 10e-6) were less present in African and Asian populations compared with Europeans, mechanically decreasing diversity. This may reflect a bias in the sampling of 1KG individuals. As a consequence, *HLA* genotypes matching resulted in less possible combinations and less eplet mismatches.

This analysis has allowed us to (1) characterize for the first time *HLA* epitopic diversity and populational level epitopic compatibilities in three different ancestry populations and (2) set an average threshold to better interpret the calculated epitopic mismatch score. Indeed, comparing an epitopic mismatch load to the distribution obtained from different 1KG non-transplanted populations (Figure 5) enables to contextualize a transplanted patient as being part of the "high," "low," or "medium" range in a population of interest. Allelic matching provides only a partial and low-resolution solution to HLA matching in transplantation. Indeed, very limited patients will get a 100% compatible graft in terms of *HLA* alleles and as illustrated in Figure 6 individual allelic mismatches can range from a very few corresponding epitopic mismatches to many, probably greatly impacting the risks of immunization against the graft. With a very limited number of solid organs compared with transplant candidates, it is important to optimize the allocation system in regards to the most up to date available data regarding tissular compatibility. The implemented concept of epitopic matching aims at assessing and improving overall transplantation quality.[10,28] This system could complete the current allocation system based on *HLA* alleles matching alone, especially for patients with rare *HLA* genotypes or highly immunized. The most appropriate donor for a patient would then be the one exhibiting the lowest number of epitope mismatches. In addition, this may allow to anticipate alloreactivity between donor and recipient *HLA* alleles. It has been previously shown that epitopic donor/recipient matching could reduce the development of de novo DSA after solid organ transplantation, specifically at the HLA class II level.[25,28] Other potential outcomes such as rejections or overall survival could be investigated through the study of HLA epitopes. Finally, besides transplantation, genetic association studies could also evaluate the presence/absence of particular eplets as risk factors for different autoimmune pathology. This may improve the understanding of certain HLA alleles association in some autoimmune conditions. With our new tool, directed toward research applications at this stage, it will

now be easier to retrospectively evaluate epitopic compatibility between donor/recipient pairs to study different outcomes such as DSA occurrence. When the epitope matching predictive power toward different adverse outcomes (DSA formation, overall graft survival ...) will be confirmed, this should be added to the graft allocation system. In the future, HLA-Epi's epitopic compatibility score might help monitor de novo DSA development therefore limiting acute antibody-mediated rejection and graft failure events and might also limit the needs in immunomodulatory compounds.[8]

ACKNOWLEDGMENTS

Acknowledgments to *Données Informatisées et VAlidées en Transplantation Cohort* (DIVAT) Consortium. Nicolas Vince has received funding from the European Union's Horizon 2020 research and innovation program under the Marie Skłodowska-Curie grant agreement No. 846520. This work is supported by the ATIP-Avenir Inserm program, the Region Pays de Loire ConnectTalent. Analysis and interpretation are those of the author.

CONFLICT OF INTEREST

The authors declare no conflicts of interest.

AUTHOR CONTRIBUTIONS

Conception and design of study: Léo Boussamet, Estelle Geffard, Nicolas Vince. *Analysis and/or interpretation of data*: Léo Boussamet, Estelle Geffard, Nicolas Vince. *Drafting the manuscript*: Léo Boussamet, Estelle Geffard, Nicolas Vince, Sophie Limou. *Revising the manuscript critically for important intellectual content*: Nicolas Vince, Sophie Limou, Florent Delbos, Alexandre Walencik, Pierre-Antoine Gourraud. *Approval of the version of the manuscript to be published*: Léo Boussamet, Estelle Geffard, Nicolas Vince, Sophie Limou, Alexandre Walencik, Florent Delbos, Pierre-Antoine Gourraud.

DATA AVAILABILITY STATEMENT

The data that support the findings of this study are openly available in Easy-HLA at https://gitlab.univ-nantes.fr/crtiteam5/easy-hla.

ORCID

Nicolas Vince https://orcid.org/0000-0002-3767-6210

REFERENCES

1. European Directorate for the Quality of Medicines and Health Care. International figures on donation and transplantation 2020. *Newsletter Transplant*; 2021.
2. Hurley CK. Naming HLA diversity: a review of HLA nomenclature. *Hum Immunol*. 2021;82(7):457-465. doi:10.1016/j.humimm.2020.03.005
3. Robinson J, Guethlein LA, Cereb N, et al. Distinguishing functional polymorphism from random variation in the sequences of >10,000 HLA-A, -B and -C alleles. *PLoS Genet*. 2017;13(6):e1006862. doi:10.1371/journal.pgen.1006862
4. Loiseau P, Busson M, Balere ML, et al. HLA association with hematopoietic stem cell transplantation outcome: the number of mismatches at HLA-A, -B, -C, -DRB1, or -DQB1 is strongly associated with overall survival. *J Am Soc Blood Marrow Transplant*. 2007;13(8):965-974. doi:10.1016/j.bbmt.2007.04.010
5. Osorio-Jaramillo E, Haasnoot GW, Kaider A, et al. HLA epitope mismatching is associated with rejection and worsened graft survival in heart transplant recipients. *J Heart Lung Transplant*. 2019;38(4, Supplement):S87. doi:10.1016/j.healun.2019.01.199
6. European Directorate for the Quality of Medicines and HealthCare. International figures on donation and transplantation 2016. *Newsletter Transplant*; 2017.
7. Snanoudj R, Legendre C. New aspects of HLA compatibility in organ transplantation. *Nephrol Ther*. 2016;12(Suppl 1):S125-S130. doi:10.1016/j.nephro.2016.01.006
8. Wiebe C, Rush DN, Nevins TE, et al. Class II eplet mismatch modulates tacrolimus trough levels required to prevent donor-specific antibody development. *J Am Soc Nephrol*. 2017;28(11):3353-3362. doi:10.1681/ASN.2017030287
9. Snanoudj R, Kamar N, Cassuto E, et al. Epitope load identifies kidney transplant recipients at risk of allosensitization following minimization of immunosuppression. *Kidney Int*. 2019;95(6):1471-1485. doi:10.1016/j.kint.2018.12.029
10. Rimando J, Slade M, DiPersio JF, et al. HLA epitope mismatch in haploidentical transplantation is associated with decreased relapse and delayed engraftment. *Blood Adv*. 2018;2(24):3590-3601. doi:10.1182/bloodadvances.2018025437
11. Tambur AR, Claas FHJ. HLA epitopes as viewed by antibodies: what is it all about? *J Am Soc Transplant Surg*. 2015;15(5):1148-1154. doi:10.1111/ajt.13192
12. Laune D, Molina F, Ferrieres G, et al. Systematic exploration of the antigen binding activity of synthetic peptides isolated from the variable regions of immunoglobulins*. *J Biol Chem*. 1997;272(49):30937-30944. doi:10.1074/jbc.272.49.30937
13. Tambur AR. Human leukocyte antigen matching in organ transplantation: what we know and how can we make it better (revisiting the past, improving the future). *Curr Opin Organ Transplant*. 2018;23(4):470-476. doi:10.1097/MOT.0000000000000538
14. Duquesnoy RJ, Marrari M, Tambur AR, et al. First report on the antibody verification of HLA-DR, HLA-DQ and HLA-DP epitopes recorded in the HLA epitope registry. *Hum Immunol*. 2014;75(11):1097-1103. doi:10.1016/j.humimm.2014.09.012
15. Tassone G, Santis DD, Vukovic I, Downing J, Martinez OP, D'Orsogna LJ. Different eplet software programs give discordant and incorrect results: an analysis of HLAMatchmaker vs Fusion Matchmaker Eplet calling software. *HLA*. 2020;96(1):52-63. doi:10.1111/tan.13897
16. Duquesnoy RJ, Takemoto S, de Lange P, et al. HLAmatchmaker: a molecularly based algorithm for histocompatibility determination. III. Effect of matching at the HLA-A,B amino acid triplet level on kidney transplant survival. *Transplantation*. 2003;75(6):884-889. doi:10.1097/01.TP.0000055101.20821.AC
17. Kramer CSM, Koster J, Haasnoot GW, Roelen DL, Claas FHJ, Heidt S. HLA-EMMA: a user-friendly tool to analyse HLA class

I and class II compatibility on the amino acid level. *HLA*. 2020; 96(1):43-51. doi:10.1111/tan.13883
18. Geneugelijk K, Niemann M, Drylewicz J, et al. OR41 PIRCHE-II: a novel tool to identify permissible HLA mismatches in kidney transplantation. *Hum Immunol*. 2017;78:39. doi:10.1016/j.humimm.2017.06.047
19. Geneugelijk K, Niemann M, Drylewicz J, et al. PIRCHE-II is related to graft failure after kidney transplantation. *Front Immunol*. 2018;9:321. doi:10.3389/fimmu.2018.00321
20. Geffard E, Limou S, Walencik A, et al. Easy-HLA: a validated web application suite to reveal the full details of HLA typing. *Bioinform Oxf Engl*. 2020;36(7):2157-2164. doi:10.1093/bioinformatics/btz875
21. Duquesnoy RJ, Marrari M, Marroquim MS, et al. Second update of the international registry of HLA epitopes. I. The HLA-ABC epitope database. *Hum Immunol*. 2019;80(2):103-106. doi:10.1016/j.humimm.2018.11.007
22. Gragert L, Madbouly A, Freeman J, Maiers M. Six-locus high resolution HLA haplotype frequencies derived from mixed-resolution DNA typing for the entire US donor registry. *Hum Immunol*. 2013;74(10):1313-1320. doi:10.1016/j.humimm.2013.06.025
23. 1000 Genomes Project Consortium, Auton A, Brooks LD, et al. A global reference for human genetic variation. *Nature*. 2015; 526(7571):68-74. doi:10.1038/nature15393
24. Abi-Rached L, Gouret P, Yeh JH, et al. Immune diversity sheds light on missing variation in worldwide genetic diversity panels. *PLoS One*. 2018;13(10):e0206512. doi:10.1371/journal.pone.0206512
25. Kubal CA, Mangus R, Ekser B, et al. Class II human leukocyte antigen epitope mismatch predicts de novo donor-specific antibody formation after liver transplantation. *Liver Transpl*. 2018; 24(8):1101-1108. doi:10.1002/lt.25286
26. Hönger G, Niemann M, Schawalder L, et al. Toward defining the immunogenicity of HLA epitopes: impact of HLA class I eplets on antibody formation during pregnancy. *HLA*. 2020; 96(5):589-600. doi:10.1111/tan.14054
27. Kosmoliaptsis V, Chaudhry AN, Sharples LD, et al. Predicting HLA class I alloantigen immunogenicity from the number and physiochemical properties of amino acid polymorphisms. *Transplantation*. 2009;88(6):791-798. doi:10.1097/TP.0b013e3181b4a9ff
28. Wiebe C, Pochinco D, Blydt-Hansen TD, et al. Class II HLA epitope matching—A strategy to minimize de novo donor-specific antibody development and improve outcomes. *Am J Transplant*. 2013;13(12):3114-3122. doi:10.1111/ajt.12478

How to cite this article: Geffard E, Boussamet L, Walencik A, et al. HLA-EPI: A new EPIsode in exploring donor/recipient epitopic compatibilities. *HLA*. 2022;99(2):79-92. doi:10.1111/tan.14505