WILEY

HLA
Immune Response Genetics

ORIGINAL ARTICLE OPEN ACCESS

The Registry of Unmet Need: A World Marrow Donor Association Analysis of Patients Without an HLA Match

Martin Maiers[1] | Valerie Greco-Stewart[1] | Abeer Madbouly[1] | James Robinson[2,3] | Hans-Peter Eberhard[4] | Jürgen Sauter[5] | Yoram Louzoun[6] | Sapir Israeli[6] | Yung-Tsi Bolon[1] | Julia Pingel[5] | Alexander H. Schmidt[5] | Eric Spierings[7] | Christina Leonhard-Melief[8] | Lydia Foeken[9] | Martine Schuit[9] | Alicia Venter[9] | Steven G. E. Marsh[3]

[1]CIBMTR (Center for International Blood and Marrow Transplant Research), NMDP, Minneapolis, Minnesota, USA | [2]Anthony Nolan Research Institute, London, UK | [3]UCL Cancer Institute, Royal Free Campus, London, UK | [4]Zentrales Knochenmarkspender-Register Deutschland (ZKRD), Ulm, Germany | [5]DKMS, Tübingen, Germany | [6]Department of Mathematics, Bar-Ilan University, Ramat Gan, Israel | [7]Matchis, Leiden, the Netherlands | [8]National Cord Blood Network, Cleveland, Ohio, USA | [9]World Marrow Donor Association, Leiden, the Netherlands

Correspondence: Martin Maiers (mmaiers@nmdp.org)

Received: 28 January 2025 | **Revised:** 16 April 2025 | **Accepted:** 7 May 2025

Funding: This work was supported by Office of Naval Research (N00014-23-1-2057 and N00014-24-1-2057).

Keywords: Bayes theorem | ethnicity | haematopoietic stem cells | haplotypes | HLA | registries

ABSTRACT

While the World Marrow Donor Association global database currently offers approximately 42.7 million potential donors and cord blood units to patients in need of haematopoietic cell transplant, lack of eight HLA-matched donors remains a significant barrier. The Registry of Unmet Need (RUN) Project seeks to address disparities in transplant access for patients with rare HLA genotypes, particularly those from populations that have been historically underrepresented and underserved by global donor registries. Patients eligible for this study searched for an unrelated donor for transplant between 2015 and 2017 and, at that time, lacked a potential eight-of-eight HLA-matched unrelated donor (MUD). Sixteen donor registries contributed data from 3654 patients using standardised data-collection project templates. To address this unmet need, pooled data were analysed to identify trends and inform global recruitment strategies. Patient genotypes were queried against the global inventory at later timepoints in 2018 and 2023 to determine whether potential matches had been recruited within the years since the initial search. Patient haplotypes were imputed using an open-source method referencing US population frequencies. The imputation process used five continental reference populations and 21 detailed populations derived from the NMDP database. The method provided a Bayesian inference of population membership. A control group consisting of US patients that yielded 1000 or more potential matches was used for comparison. RUN patient haplotype and genotype frequencies were substantially lower compared with controls; both the more frequent and less frequent haplotypes in RUN patients were found to be approximately 100 times less common than those in the control group. We identified 782 potential cases in which a potential MUD was recruited after the initial RUN patient search was performed; while this result is being further investigated, clear patterns of where these new matches can be found have emerged; typically, new matches are found outside the country where the patient search was initiated. Our findings demonstrate that rare haplotypes are the primary barrier to identifying a MUD; the presence of rare alleles or haplotype combinations, as with multi-race ancestry, is rarely the cause. Although strategic donor recruitment efforts will help improve MUD access, patient transplants should not be delayed in pursuit of a MUD when viable alternative options are available.

1 | Introduction

Haematopoietic cell transplant (HCT) offers a potentially curative therapy for patients with haematologic disorders and malignancies. The most prognostic factor for HCT success is the degree of HLA similarity between donor and recipient, with allele-level match at HLA-A, -B, -C and -DRB1 representing the minimum standard by which donors and recipients are matched [1–4]. However, the highly polymorphic nature of HLA genes and the stringent match criteria often preclude identification of a suitable donor within the patient's family and, in many cases, from within national unrelated donor registries. The World Marrow Donor Association (WMDA) offers the Search and Match Service (SMS), a global repository of nearly 43 million searchable donors and cord blood units (CBU) from over 130 organisations in 57 countries to collaboratively extend access to HCT to patients worldwide (http://statistics.wmda.info). However, despite this enormous inventory, there remain patients for whom an HLA-matched donor cannot be identified, limiting opportunity for life-saving transplant. Specifically, identification of a matched donor for patients with uncommon HLA genotypes, typically from populations that have been historically underrepresented and underserved by global donor registries, has created an obstacle to receiving HCT.

While transplant professionals are generally aware of the barriers in access to transplant within their own countries, the WMDA aims to address these obstacles at a global level. By exploring targeted donor recruitment strategies across ethnic diasporas, WMDA seeks to maximise the likelihood that every patient can be transplanted with a suitable donor regardless of ancestry. Previous studies have shown a difference in HLA distribution of patients with a match compared to those without a match, citing rare or uncommon alleles, rare gene associations and rare haplotypes as the reason for this disparity [5–7]. Indeed, patients with rare genotypes, potentially reflective of characteristics such as rare alleles, associations or uncommon haplotype combinations, have been demonstrated to be less likely to identify a MUD than counterparts with higher frequency genotypes [8, 9]. Patients from ethnicities underrepresented in the global inventory are also less likely to identify an MUD for transplant [10–12].

The Registry of Unmet Need (RUN) is an initiative of the WMDA, established to identify gaps in and opportunities for improvement in access to unrelated donors for allogeneic HCT. Established in 2017, RUN invited members of the global community to submit the HLA genotypes of patients for whom no potentially HLA-matched donors, defined as having the possibility of allele match at HLA-A, -B, -C and -DRB1, were present within the WMDA SMS at patient search onset. Analyses of these uncommon genotypes, presented herein, have provided insights towards the immunogenetic characteristics of patients for whom search is difficult and have illuminated both global areas of deficit in donor inventory and recent successes in strategic bolstering of the global donor inventory that could potentially unite future patients with donors for transplant.

2 | Materials and Methods

2.1 | Patient Cohort

All WMDA member organisations who facilitate patient searches were invited to participate in the project via announcement in the electronic newsletter, *Stem Cell Matters* and through email outreach with registry leadership. A standardised data template was developed to collect HLA genotypes, basic patient demographics, race or ethnicity, and country of origin. Given the global nature of this study and the local variation in the use of race and ethnicity terms, we provided no constraints on the data template to encourage participating centres to use their local terminology. Study inclusion criteria were that patient searches must have been initiated between 2015 and 2017 and that, at the time of search onset, no potential donor match within the WMDA SMS was available. This time period was chosen because it aligned with the study's launch, and a 3-year window was selected to balance cohort size with the potential variability introduced by registry growth during the enrollment period. Cord blood and mismatched-adult donor options were not considered for this study for simplicity to focus on the question of unmatched HLA types. The number of donors in the global inventory ranged from 25,976,458 to 33,581,469 over these 3 years. A potential match was defined as having the possibility of allele match at both copies of the HLA-A, -B, -C and -DRB1 loci (eight-of-eight allele match) without requiring a match probability threshold. The allele match was defined as having identical amino acid sequences in the antigen recognition domain (exons 2 and 3 for class I and exon 2 for class II HLA) [13]. Sixteen registries in 15 countries (Figure 1A) submitted data for a total of 3646 patients with reported geographical ancestries from 53 countries (Figure 1B). Three registries provided more than 75% of the patients, introducing some bias in the cohort (Table S1). The patient cohort was further refined by rerunning the WMDA SMS search as of September 2022. Analysis of the registration dates of donors and the number of potential donor matches led to the exclusion of 581 patients for a final analysis set of $N = 3054$. Exclusions were due to several potential factors including clerical errors, not searching the global donor pool, temporary unavailability of a matching donor, and not using a predictive match algorithm to identify potential matches among donors with missing or ambiguous HLA typing.

2.2 | Rare Allele Analysis

The HLA alleles from the patients in the final analysis set were analysed in terms of categorising them according to two systems of common and well-documented (CWD) alleles [8, 9]. The CWD 2.0.0 list was based on an analysis performed in 2012 and the CIWD 3.0.0 list, which introduces an 'intermediate' category, is based on an analysis published in 2020. The CIWD 3.0.0 list is more recent and includes non-US populations. Missing and ambiguous alleles were excluded and the percentage of the remaining alleles that fell into the rare category was computed for both.

In the 2012 CWD list [8], common alleles were assigned based on presence in multiple populations with known frequencies generally above 1 in 1000. In the 2020 CIWD list [9]

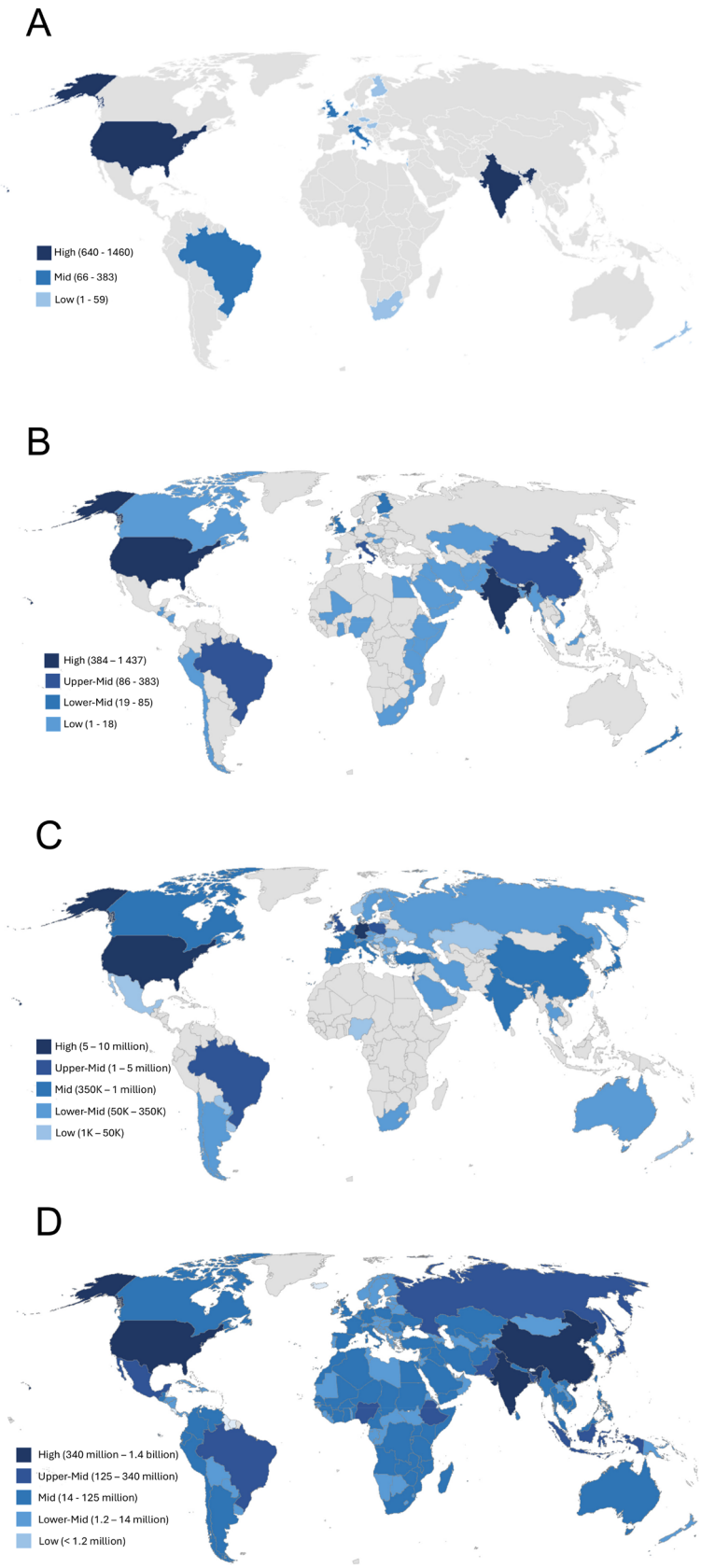

FIGURE 1 | Maps of the global distribution of (A) the countries submitting patients to the RUN cohort, (B) the RUN patient cohort country of origin, (C) the WMDA donors and (D) the global population.

this category was extended to also include alleles observed at 1 in 10,000 within a single population. The term 'intermediate' was added in 2020 to apply to alleles observed at a rate of below 1 in 10,000 but above 1 in 100,000. The definition of 'well-documented' was revised slightly between the most recent two publications as 'alleles that have been observed five or more times in unrelated individuals'. All other alleles are considered rare.

2.3 | Race, Ethnicity and Ancestry Analysis

The data collection template (Table S2) included columns for:

- Patient Country of Residence
- City/State/Region
- Patient Race or Ethnicity
- Patient Ancestry (country of origin)
- Origin details (Territory, region, state, cast, other)
- Patient Primary Language.

Recovered templates were manually consolidated into two reduced lists. The first two residence-related variables were consolidated into one of 12 regions defined by the allelefrequencies.net database [14] (Table S3). The remaining variables were used to assign one of the 21 detailed and five broad race/ethnicity categories defined by the NMDP [15] (Table S4). Although there is no universal consensus on race/ethnicity terminology in the scientific literature [16] we used the US race/ethnicity categories because, although these population categories were defined for describing donors in the US, they are the only set that both approximate global population diversity and are linked to haplotype frequency data, which is essential for the HLA analysis described in the next section. US race/ethnicity categories have been shown to offer the strongest correlation between self-identification and HLA variation [17].

For example, a patient with country of residence of New Zealand (NZ) and Race of Chinese was assigned to the 'Oceania' (OCE) region and the Chinese (NCHI) race/ethnicity group within the Asian/Pacific Islander (API) category. The location of the submitting registry was used to define the country of residence if an alternate was not provided. For instance, a patient whose search was initiated through a registry in India who reported 'United Arab Emirates' as their country of origin was assigned the region code Western Asia (WAS) and the race/ethnicity code 'Middle Eastern/North African' (MENAFC).

As a measure of the uncertainty in the assignments, the squared Euclidean norm was computed, with a higher value indicating less uncertainty. This measure is akin to the typing resolution score [18] applied to the probability distribution for race classification instead of genotypes.

2.4 | Haplotype and Genotype Analysis

The HLA typings were analysed using an imputation algorithm [19] capable of combined imputation of phased HLA genotypes. This method accomplished three things: it confirmed or re-assigned the patient's population membership among the 21 detailed and five broad categories in the reference haplotype frequency data, used the imputed population(s) to estimate the HLA genotype frequency (without considering phase) at HLA-A, -B, -C and -DRB1, and provided an estimate of the pair of haplotypes that determine the genotype. This output was important since the haplotype was not determined experimentally and provided an informative perspective for studying HLA at the population level. Since the output of imputation is a probability distribution of either phased or unphased genotypes, the degree of ambiguity can be simply calculated using the typing resolution score [18] metric. The haplotype frequency distributions were compared by considering the median of the distribution for the more common and less common haplotypes between the RUN cohort and the controls.

The reference HLA haplotype frequencies used were from the US NMDP [20], developed as an update and extension to the previously published US NMDP frequencies [15], since this is the largest and most comprehensive dataset available for analysis.

2.5 | Analysis of New Matches

Global searches were performed for all patients in the RUN cohort against the WMDA SMS inventory as of February 2024 ($N = 41{,}007{,}014$) [statistics.wmda.info] to determine whether donors added since the time of the original search were matched. In cases where the donor status was ambiguous, potentially having been missed at the time of the original patient search, the patient was excluded to err on the side of ensuring that the analysis cohort was demonstrably lacking a match at the time the cohort was assembled.

Based on this rerun of the global search, patients with new matches were noted for further analysis. New potential matches with a probability eight-of-eight matches of $< 50\%$ were not considered as new matches since they were considered unlikely to match.

2.6 | Data Permissions and Software Availability

The patient data were provided by the 16 participating registries via spreadsheets based on a template developed for this project. No personally identifiable information was collected, and the permission to use the data for this study was granted by each registry.

The donor data were provided by the WMDA. Permission to use the donor data for this study was obtained by contacting each organisation providing donors to WMDA and requesting permission. A single organisation comprising 81,379 donors opted out of participation and their donors were excluded from the analyses.

The software used for analyses in this manuscript is available upon request. The imputation method used is available on github (https://github.com/nmdp-bioinformatics/py-graph-imputation) or pypi (https://pypi.org/project/py-graph-imputation/). The haplotype frequencies used are available at zenodo

(https://zenodo.org/doi/10.5281/zenodo.4474441) [20]. The CWD and CIWD lists are published [8, 9] and are available in a computable form (https://github.com/nmdp-bioinformatics/CIWD). HLA validation and resolution normalisation were performed using py-ard [21] (https://github.com/nmdp-bioinformatics/py-ard).

3 | Results

3.1 | Rare Allele Analysis

The ambiguity rate was near 4% for each locus (Table 1). Analysis of alleles based on the CIWD 3 catalogue found only 12 cases of rare alleles across HLA-A, -B, -C and -DRB1, representing an overall proportion of 0.051%. The rare alleles are provided (Table 2) with the largest number (5) observed at HLA-A, three at HLA-B, and two at HLA-C and HLA-DRB1, respectively.

Since this cohort predates the CIWD 3 publication, we also considered the CWD 2 list and observed 57 cases of alleles classified as rare (Table S5) with seven rare alleles observed twice and one rare allele (HLA-B*37:04) observed in four patients from three different populations. The number of CWD 2 rare alleles per locus ranged from 19 for HLA-C down to 5 for HLA-DRB1. The overall percentage of rare alleles according to the CWD 2 list was 0.245% or nearly five times higher than according to the CIWD 3.0 list.

3.2 | Race and Ethnicity Analysis

Figure 2 shows a three-level Sankey diagram of the patient cohort with the regional origin of the patients (left), the self-identified or presumed race category (middle) and the category assigned by the HLA-based classifier (right). The category 'multiracial' was assigned by the classifier for 15 (0.04%) patients at the broad race category level but for 110 (3.7%) at the detailed category (Tables S6 and S7 respectively). Patients from South Asia primarily mapped to the Asian or Pacific Islander region. However,

TABLE 2 | CIWD 3 rare alleles in the RUN cohort patients ($N = 12$).

CIWD 3—rare alleles	Count	Patient country of origin
A*03:240	1	US
A*11:190	1	US—European American
A*24:135	1	IN
A*66:06	1	NL
A*68:75	1	US—African American
B*38:08	1	BR
B*39:16	1	US—African American
B*40:128	1	IN
C*07:304	1	US
C*14:17	1	IN
DRB1*14:158	1	UK
DRB1*14:31	1	IN

Note: Each allele was observed once.

TABLE 1 | Analysis of RUN cohort HLA alleles by their categorisation according to the CWD 2 and CIWD 3 listings of common, intermediate, well-documented and rare.

	A	B	C	DRB1	Total
Missing	0	0	27	0	27
Typings	3054	3054	327	3054	12,189
Alleles	6108	6108	6054	6108	24,378
Ambiguous	247	243	254	263	1007
PCT ambiguous	4.04%	3.98%	4.20%	4.31%	4.13%
Unambiguous	5861	5865	5800	5845	23,371
CIWD 3—common	5835	5841	5773	5824	23,273
CIWD 3—intermediate	10	14	14	15	53
CIWD 3—well-documented	11	7	11	4	33
CIWD 3—rare	5	3	2	2	12
CIWD 3—PCT rare	0.085%	0.051%	0.034%	0.034%	0.051%
CWD 2—common	5796	5693	5382	5798	22,669
CWD 2—well-documented	48	156	399	42	645
CWD 2—rare	17	16	19	5	57
CWD 2–PCT rare	0.291%	0.274%	0.329%	0.086%	0.245%

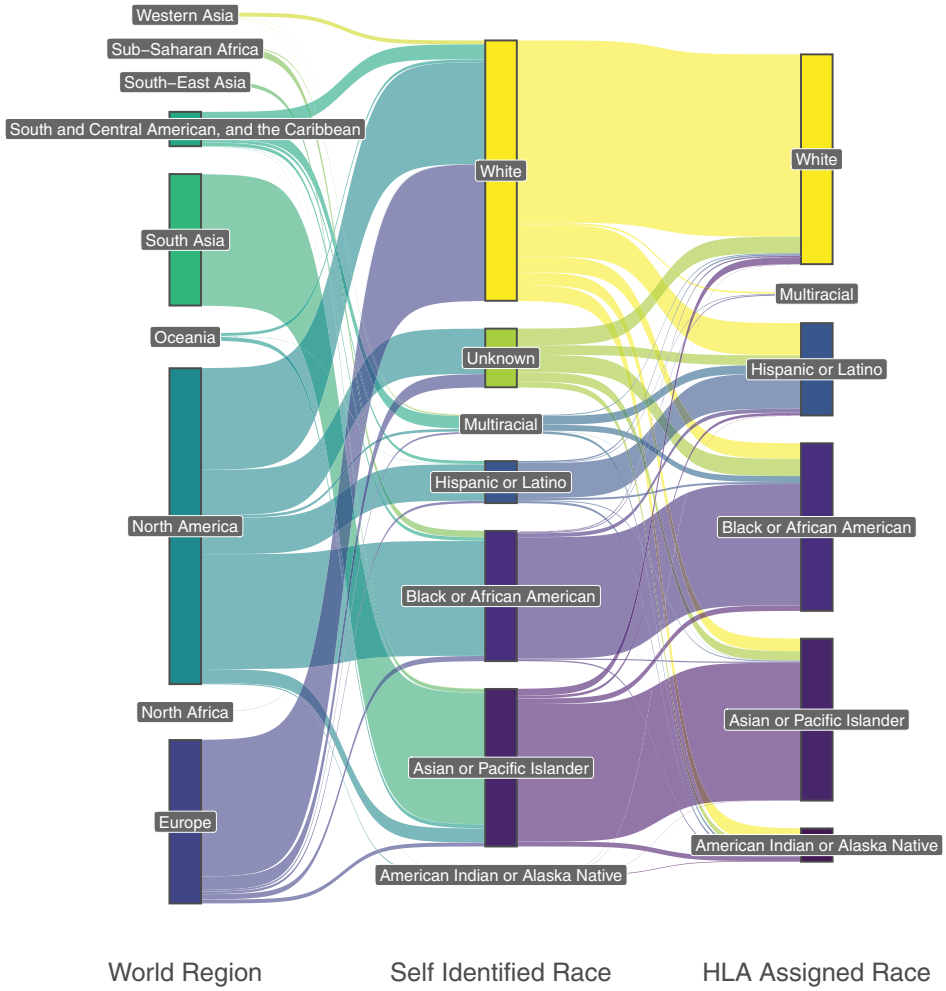

FIGURE 2 | A three-level Sankey diagram of (left) the regional origin of the RUN patients, (middle) the self-assigned or presumed race category and (right) the population category assigned using an HLA-based classifier (15 patients assigned to multiple population categories are not displayed).

patients from North America mapped to several groups, with the largest being Black or African American. Notably, the HLA-based classifier placed more RUN patients in Hispanic or Latino and Black or African American categories and a smaller fraction in White than self-identification, suggesting a higher underlying proportion of non-European HLA and more complex genetic admixture than what is self-reported.

3.3 | Haplotype and Genotype Analysis

Figure 3 shows a density plot of the genotype frequency distribution for the RUN cohort (red) and a control set of US patients with an eight-of-eight match. The density of the genotype frequency distributions for both cohorts forms a bell curve on a log scale indicating a power law distribution for the genotypes. The distributions overlap, but the RUN cohort is skewed substantially lower with a peak at approximately 1/1000 of the distribution for the controls. The green dotted line represents $1/N$ where N is the global registry population. Note that most of the RUN distribution is to the left of this line, which is consistent with this cohort lacking a match in the global donor pool. The black dashed line represents $1/M$ where M is the world population, with a substantial tail of the RUN genotype frequency distribution to the left of this line indicating genotypes that have an extremely low expected frequency and therefore are very unlikely to have a match.

Figure 4 shows the haplotype frequencies for the run patients (red) and a control set of US patients with multiple eight-of-eight matches (blue). The top two graphs are density plots of the haplotype frequencies for the higher and lower frequencies of the two haplotypes, respectively. Notably, both distributions are shifted lower for the RUN patients by a factor of 17 for the lower frequency haplotype and by a factor of 22 for the higher frequency haplotype. This indicates that, although there are some exceptions, most RUN patients have two low frequency haplotypes. A single uncommon haplotype is not sufficient to prevent matches from being found based on the matched cohort having a median lower frequency haplotype of 1 in 84,134. Further evidence for the claim that the haplotype pair drives the matching can be found in the lower plot in Figure 4. The blue line is a regression of the frequency of the pair of haplotypes for the control cohort, and the red line is a regression of the frequency of the pair of haplotypes for the RUN patients. The green dotted lines indicate a threshold of 1 in 1000 (common) and the red dotted line indicates the upper bound of rare starting at 1 in 10,000,000. The grey dotted line is the midpoint of the two extremes at 1 in 100,000. The control cohort genotype regression line (in blue) is predominantly

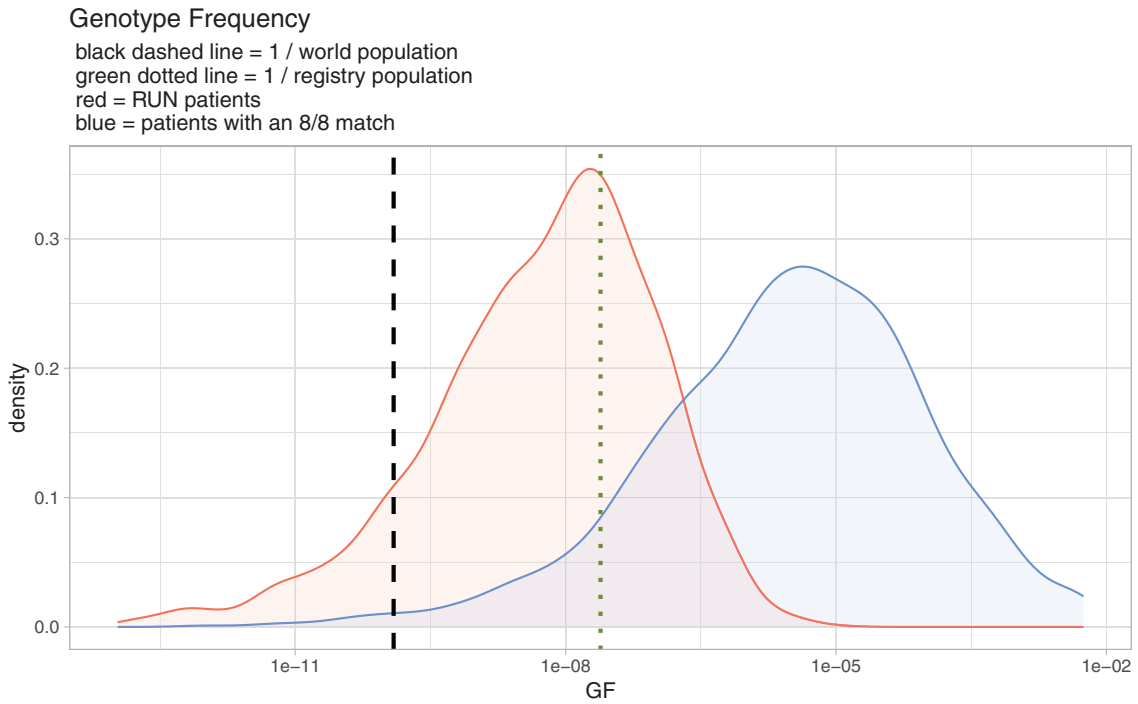

FIGURE 3 | Genotype frequencies for the RUN cohort (red) and a control set of US patients with an eight-of-eight match (blue). The vertical dashed line represents 1/the world population. The green dotted line represents 1/the WMDA global donor registry population.

in the common range for the higher frequency haplotype, with a wide range for the lower frequency haplotype but with very few extending into the rare range. The RUN cohort (in red) has almost no genotypes where either haplotype is common, and the regression line curves, indicating two modes in the data: a subset close to the diagonal where both haplotypes are above the midpoint (grey dotted lines) at 1 in 100,000 where matching may be possible with registry growth and a long tail of rare genotypes driven by low frequencies in both haplotypes, where the lower frequency haplotype is below 1 in 100,000.

3.4 | Analysis of New Matches

A total of 341 out of 3054 (11.2%) patients that met the strict criteria for inclusion in the cohort (no potential eight-of-eight allele match as of December 2017) were found to have at least one potential match in the global inventory over 6 years later (February 2024). Of these 341, a total of 253 have one potential match and 88 have two or more potential matches. These 341 patients have a median genotype frequency that is 11.6 times higher than that of the patients that still do not have a match (5.39×10^8 vs. 4.65×10^9). They also have a higher proportion with two haplotypes with a frequency of at least one in 10,000 each (25.8% vs. 8.2%).

4 | Discussion

4.1 | Population Genetics

In the RUN cohort, possessing two uncommon haplotypes was the primary barrier to identifying an eight-of-eight MUD. Although rare alleles and haplotype combinations (e.g., multi-race ancestry) were observed, they affected fewer cases than rare genotypes. Rare alleles were infrequent, with 99.6% of alleles classified as 'common' (observed in at least one in 10,000 individuals), indicating that uncommon or rare alleles do not explain the lack of matches.

Most RUN patients without an eight-of-eight MUD possessed at least one uncommon haplotype. Given the linkage disequilibrium among alleles in the MHC region and the prevalence of well-defined ancestral HLA haplotypes linked to different world geographical ancestries/regions, these uncommon haplotypes are either scarce in the current global donor pool due to lower recruitment in the regions where these haplotypes are more likely to exist or are due to recent recombination.

The computed genotype frequency is a strong predictor of match likelihood and remains a valuable prognostic tool. Overall donor recruitment following existing patterns and target regions is unlikely to substantially improve MUD access, and on an individual basis, transplant should not be delayed when alternative options such as haploidentical (half-matched) related donors (MMRD), mismatched unrelated donors (MMUD) or CBU are available.

4.2 | Recruitment

Strategic recruitment in underserved populations is essential for enhancing global donor diversity. Although complete coverage of the global genotype repertoire is unattainable, focusing on populations most likely to enrich the global inventory through inclusion of uncommon and geographically or ethnogenetically-restricted genotypes will maximise the impact of recruitment efforts, particularly given the cost

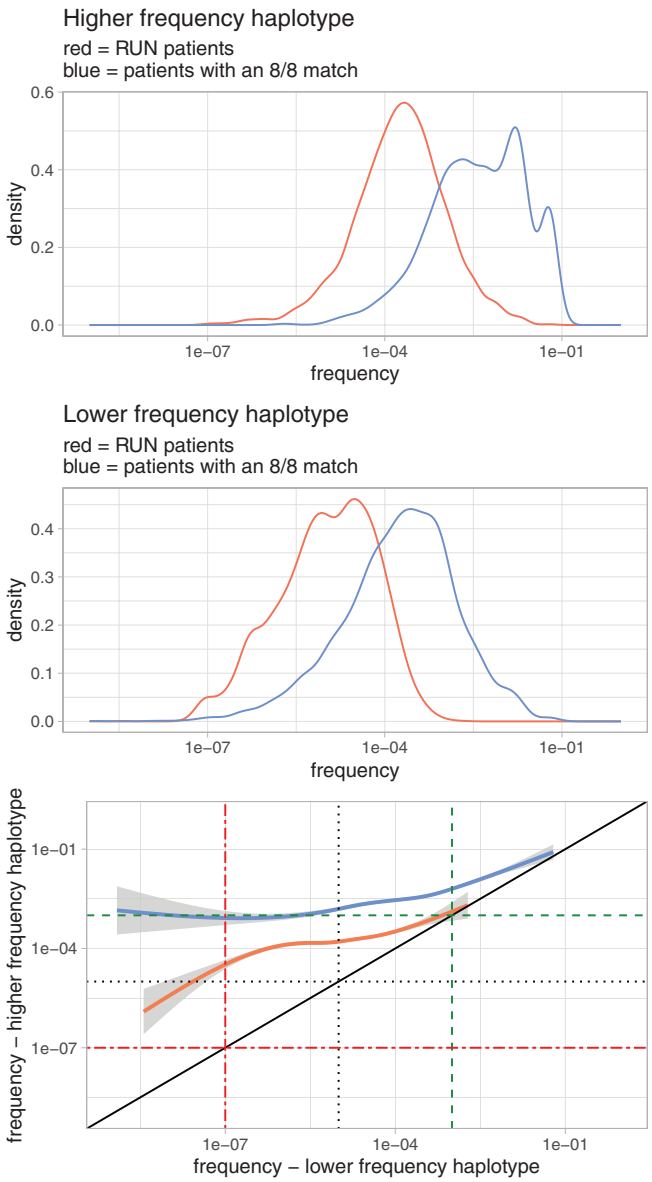

FIGURE 4 | Haplotype and genotype frequencies for the run patients (red) and a control set of US patients with an eight-of-eight match (blue). The top graph shows the frequency for the higher frequency of the two haplotypes of the patient, the middle graph shows the lower frequency haplotype, and the bottom graph is based on the haplotypes with a regression curve for the two patient cohorts. The green dotted lines indicate a threshold of 1 in 1000 (common) and the red dotted line indicates the upper bound of rare starting at 1 in 10,000,000. The grey dotted line is the midpoint of the two extremes at 1 in 100,000.

associated with these activities (i.e., recruitment events, HLA testing and donor management).

This study identified regional variation in the rate of discovery of donors that provide matches for genotypes that originally did not have a match. Targeted recruitment in geographic regions enriched for these populations will increase the likelihood of serving patients in need more so than recruitment in other regions. Considering the registry of origin of new donors who match RUN patients (Figure 5), recruitment in the South Asian and Sub-Saharan African regions appears to have the biggest impact on matching for the RUN cohort of patients. These observations can inform a strategic recruitment approach that balances potential patient impact with fiscal considerations associated with donor recruitment and management.

Moreover, the strategic recruitment plan should consider other factors that increase the likelihood of a donor being selected; for example, characteristics such as donor age and commitment (i.e., likelihood that they will donate after recruitment) are important considerations in a registry's recruitment approach.

4.3 | Importance of Self-Identification

We urge donor registries to expand patient race/ethnic categories to better identify underserved groups and to seek typing confirmation of rare alleles. Collection of self-reported race and ethnicity data (SIRE) of donors and recipients is encouraged to support registry planning by understanding the populations served and identifying gaps in coverage [16, 19]. Targeted recruitment within non-European populations can increase registries' haplotypic diversity, and the use of bioinformatics tools that assess patient HLA frequencies (e.g., search prognosis, predictive match algorithms) [22] can help set search expectations and inform donor selection strategy, especially when considering the global donor pool has substantial numbers of poorly typed donors and donor availability is a barrier.

4.4 | Search Strategy

Delayed transplant may lead to disease progression, alloimmunisation, complicating patient care and ultimately resulting in poorer outcomes. The transplant timeline can be extended when a patient search yields numerous donors with ambiguous HLA typings, requiring additional testing to establish desired histocompatibility.

Although the global donor pool expanded by approximately 15 million during the interval between the initial RUN patient search and the repeat query, fewer than 10% of RUN patients now have a likely 8/8 match. From the search coordination perspective, this observation is cautionary, suggesting that, if no eight-of-eight match exists for a patient, it is unlikely that one will be identified through prolonged search activities. Moreover, while localised recruitment efforts within a community that shares a patient's ancestry might ultimately benefit a patient in need, the likelihood of identifying a matched donor among new registrants is very low.

However, a poor initial MUD search could be useful to inform patient care; alternative donors, including cord blood, haploidentical related or MMUD transplant could offer an alternative therapeutic option when a MUD is not readily available. Use of post-transplant cyclophosphamide (PTCy) as a graft versus host disease (GvHD) prophylactic strategy has extended access to transplant by reducing the stringent HLA match requirements, allowing transplantation of haploidentical related

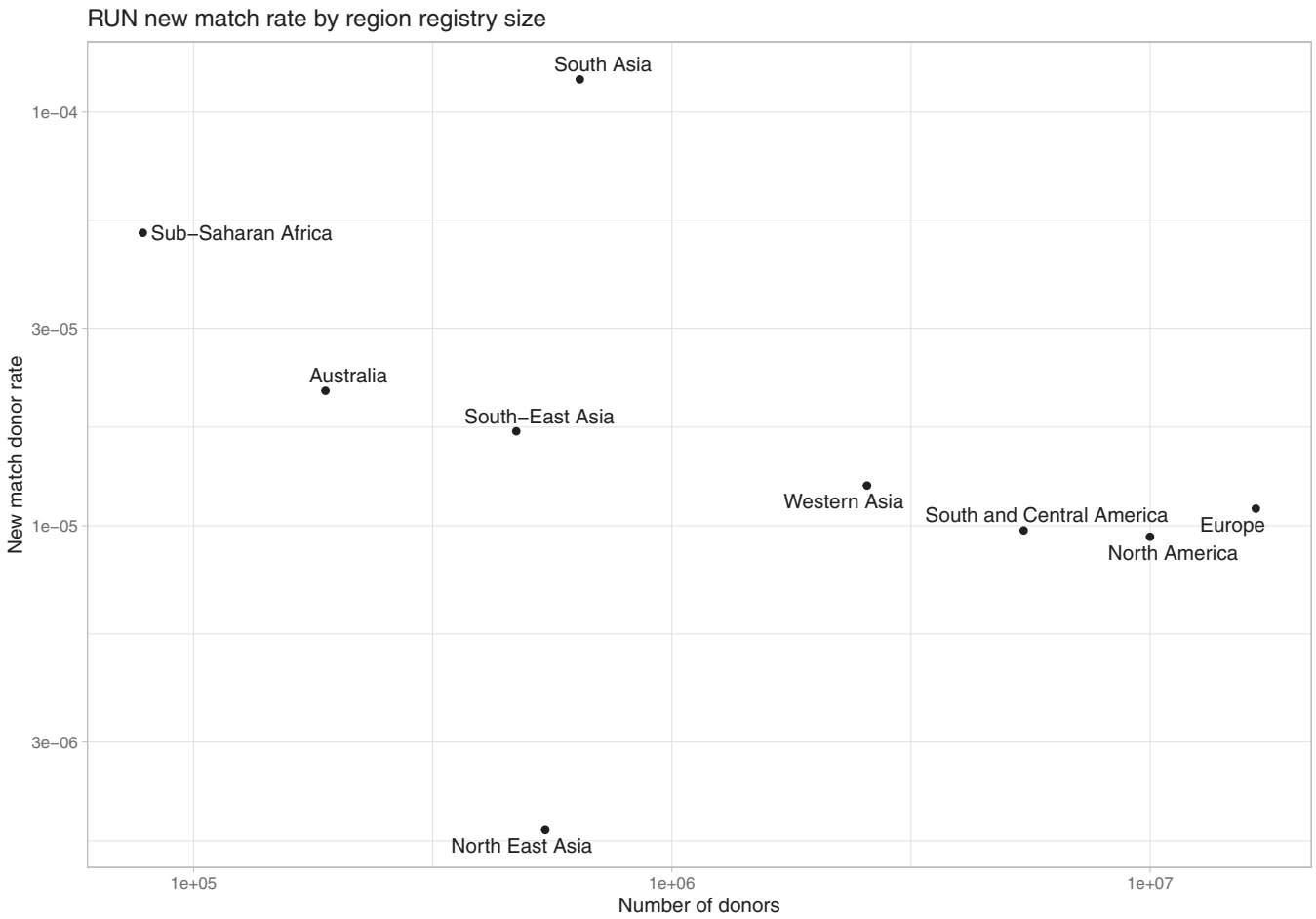

FIGURE 5 | RUN new match rate by donor region and donor registry size. Each point represents all registries from that region and the total number of donors (on the *x*-axis) and the rate at which donors from this region provided new matches for RUN patients (on the *y*-axis).

donors. While rare patient haplotypes can limit unrelated donor matches, the likelihood of a sibling sharing a haplotype follows Mendelian inheritance and remains unaffected by HLA population frequency. More recently, PTCy has been applied in the MMUD setting where outcomes comparable to transplants involving MUD were achieved [23]. Modelling studies of the US donor database have demonstrated that through use of MMUD every patient in need could find a donor [24, 25]. Use of MMUD will thus serve to extend access to underserved populations, underscoring the importance of the global donor inventory in continuing to provide options for patients in need.

Although cord blood use has declined at many transplant centres, it still accounts for 9% of unrelated donor product shipments worldwide, as reported by WMDA. Cord blood could offer a valuable opportunity to extend coverage to transplant for patients from underrepresented populations who are less likely to have an available MUD, representing a substantial number of RUN patients for whom strategic recruitment might not eliminate the gap in equity to access. Outcomes of cord blood transplant continue to improve at major transplant centres, and best practises have been disseminated to the transplant community through publication of CBU selection and transplantation guidelines [26]. Advancements in paediatric transplant have led to comparable outcomes as those involving matched sibling donors [27], and attributed to the intermediate intensity conditioning regimen have also been reported, even with utilisation of CBUs with match grades as low as three to four of eight [28]. Cord blood transplant programmes at institutions serving patients with rare haplotypes could help to address ongoing disparities in transplant access.

To further inform search strategy, search prognosis tools can be leveraged to rapidly assess patient search prospects [22]. The primary value for prognostic tools is to provide support for rapidly pursuing mismatched options. Prognostic tools use localised HF datasets; implementations using global HF could more accurately predict patient match prospects. While these tools are helpful, they are not a substitute for preliminary search of national and global databases, which should be offered free of charge through the patient's national registry [29].

Finally, the importance of global database queries cannot be understated. Our analysis revealed that 600 patients initially reported as having no potential 8/8 match actually had at least one match available at the start of the search. While this error in reporting could have occurred due to misinterpretation of inclusion criteria (i.e., reporting patients with no known match vs. potential match), in many cases donors existed outside the patient's country of search origin, speculatively suggesting that the preliminary search was restricted to either national or cooperative (i.e., EMDIS) registries (meaning registries connected

by peer-to-peer searching). To maximise the likelihood of a productive search that yields an optimal donor, concurrent search of national and international registries is strongly encouraged.

4.5 | Limitations

This study has limitations, including potential bias due to the participating organisations. The WMDA global donor pool is biassed since recruitment rates are much higher in countries in Europe compared to sub-Saharan Africa or countries without any donor registry like Indonesia (Figure 1C,D). The use of US populations is not ideal, but a global HLA haplotype frequency dataset is not readily available and due to a long history of immigration from all world populations, the US HLA haplotype frequency data is a reasonable source for this study.

4.6 | Future Directions

Future analysis could involve mapping the rare haplotypes to further pinpoint opportunities for donor recruitment beyond the regional messages here. Another opportunity for further analysis is to learn more about the genotypes that are thought to be common but have no match or conversely the genotypes that are common but are predicted to be rare. A more detailed investigation into the role of recombination could provide insights into haplotype diversity and its impact on matching. With the rise in mismatched transplants, future studies should explore the effects of mismatching at specific HLA loci on the pool of available donors. Self-identification continues to be an opportunity for the donor registry community to improve data collection [16, 17]. HLA haplotype frequency data should be generated from the global donor pool to further inform matching algorithms and modelling studies. Since the conception and implementation of the RUN project, advancement in treatment paradigms [23, 30–32]—specifically the advent of PTCy GvHD prophylaxis—has further expanded access to transplant with unrelated donors through reduction of the HLA barrier to facilitate MMUD transplants that yield comparable results to those involving MUD. Although encouraging, strategic recruitment of donors from all ethnic diaspora is still essential to ensuring that the optimal donor will be available for every patient in need. The improved outcomes of MMUD transplants compared to MUD transplants will influence the cost-effectiveness of recruitment strategies, particularly those targeting rare HLA or underrepresented populations. This underscores the need for registries to integrate non-HLA factors, such as donor age and commitment, into recruitment strategies to effectively meet both domestic and international transplant demands.

In summary, this study highlights the continued need for altruistic unrelated donations to provide life-saving transplants for patients. We have identified deficits in HLA coverage among specific populations within the global donor inventory but also highlighted successes in strategic recruitment that have led to matches that were not possible at the time of the patient initial search. Importantly, we have characterised the HLA characteristics that may preclude identification of a MUD to inform global recruitment strategy and support best practises in patient unrelated search.

Author Contributions

M.M. and L.F. conceived of the study. A.V., A.M. and M.M. helped prepare the data. M.M., H.-P.E., S.I. and Y.L. performed analysis of HLA. M.M., V.G.-S. and A.M. drafted the manuscript. M.M., V.G.-S., A.M., J.R., H.P.-E., J.S., Y.L., S.I., Y.-T.B., J.P., A.H.S., L.F., M.S., A.V. and S.G.E.M. reviewed and commented on the manuscript and approved the final version.

Acknowledgements

This study was conducted on behalf of the RUN Project Group of the Bioinformatics & Innovation Committee of WMDA. We thank the staff of the WMDA and WMDA office (Leiden, the Netherlands) Machteld Oudshoorn, Monique Jöris, Jane Prior, Louise Cho, Pam Robinett, Claudia Rutt, Adam Watson, Caroline Räther, Loren Gragert, Nikki Garret, Sigal Manor, Medhat Askar, Steve Devine and Jeff Auletta for input and discussion. We also thank the registries that participated in this study. Thanks to Michael Wright for proofreading the manuscript. The research presented herein reflects the analyses and consensus perspectives of the RUN project team members and does not necessarily reflect those of the authors' parent organisations. M.M., V.G.-S., A.M., Y.-T.B., S.I. and Y.L. are supported by N00014-23-1-2057 and N00014-24-1-2057 from the Office of Naval Research.

Conflicts of Interest

The authors declare no conflicts of interest.

Data Availability Statement

The data that support the findings of this study are available on request from the corresponding author. The data are not publicly available due to privacy or ethical restrictions.

References

1. J. Dehn, M. Setterholm, K. Buck, et al., "HapLogic: A Predictive Human Leukocyte Antigen–Matching Algorithm to Enhance Rapid Identification of the Optimal Unrelated Hematopoietic Stem Cell Sources for Transplantation," *Biology of Blood and Marrow Transplantation* 22, no. 11 (2017): 2038–2046, https://doi.org/10.1016/j.bbmt.2016.07.022.

2. A. Little, A. Akbarzad-Yousefi, A. Anand, et al., "BSHI Guideline: HLA Matching and Donor Selection for Haematopoietic Progenitor Cell Transplantation," *International Journal of Immunogenetics* 48, no. 2 (2021): 75–109, https://doi.org/10.1111/iji.12527.

3. J. Dehn, S. Spellman, C. K. Hurley, et al., "Selection of Unrelated Donors and Cord Blood Units for Hematopoietic Cell Transplantation: Guidelines From NMDP/CIBMTR," *Blood* 134 (2019): 924–934, https://doi.org/10.1182/blood.2019001212.

4. E. Spierings, A. Madrigal, and K. Fleischhauer, "Histocompatibility," in *The EBMT Handbook*, ed. A. Sureda, S. Corbacioglu, R. Greco, N. Kröger, and E. Carreras (Springer International Publishing, 2024), 73–84, https://doi.org/10.1007/978-3-031-44080-9_9.

5. A. Balas, F. García-Sánchez, and J. L. Vicario, "Allelic and Haplotypic HLA Frequency Distribution in Spanish Hematopoietic Patients. Implications for Unrelated Donor Searching," *Tissue Antigens* 77, no. 1 (2011): 45–53, https://doi.org/10.1111/j.1399-0039.2010.01578.x.

6. V. Greco-Stewart, J. Kiernan, D. Killeen, et al., "Unrelated Donor Choices for Allogeneic Hematopoietic Cell Transplantation in Canada: An Evaluation of Factors Influencing Donor Selection," *Transfusion* 58, no. 3 (2018): 718–725, https://doi.org/10.1111/trf.14458.

7. J. M. Tiercy, G. Nicoloso, J. Passweg, et al., "The Probability of Identifying a 10/10 HLA Allele-Matched Unrelated Donor Is Highly

Predictable," *Bone Marrow Transplantation* 40, no. 6 (2007): 515–522, https://doi.org/10.1038/sj.bmt.1705787.

8. S. J. Mack, P. Cano, J. a. Hollenbach, et al., "Common and Well-Documented HLA Alleles: 2012 Update to the CWD Catalogue," *Tissue Antigens* 81, no. 4 (2013): 194–203, https://doi.org/10.1111/tan.12093.

9. C. K. Hurley, J. Kempenich, K. Wadsworth, et al., "Common, Intermediate and Well-Documented HLA Alleles in World Populations: CIWD Version 3.0.0," *HLA* 95, no. 6 (2020): 1–16, https://doi.org/10.1111/tan.13811.

10. L. Gragert, M. Eapen, E. Williams, et al., "HLA Match Likelihoods for Hematopoietic Stem-Cell Grafts in the U.S. Registry," *New England Journal of Medicine* 371, no. 4 (2014): 339–348, https://doi.org/10.1056/NEJMsa1311707.

11. V. Greco-Stewart, H. Elmoazzen, G. Morris, et al., "Improved Access to Better HLA-Matched Hematopoietic Cells for Allogeneic Transplant: Analysis of Donors and Cord Blood Units Selected for Canadian Patients in 2018," *Transfusion* 60, no. 7 (2020): 1508–1518, https://doi.org/10.1111/trf.15795.

12. G. Nicoloso, O. Kürsteiner, F. Bussmann, M. Marbacher, and J. M. Tiercy, "A Study of Selected Hematopoietic Stem Cell Donors Provided by an Intermediate Size Registry," *European Journal of Haematology* 103, no. 4 (2019): 426–432, https://doi.org/10.1111/ejh.13307.

13. E. Nunes, H. Heslop, M. Fernandez-Vina, et al., "Definitions of Histocompatibility Typing Terms: Harmonization of Histocompatibility Typing Terms Working Group," *Human Immunology* 72, no. 12 (2011): 54–56, https://doi.org/10.1016/j.humimm.2011.06.002.

14. F. F. Gonzalez-Galarza, A. McCabe, E. J. M. D. Santos, et al., "Allele Frequency Net Database (AFND) 2020 Update: Gold-Standard Data Classification, Open Access Genotype Data and New Query Tools," *Nucleic Acids Research* 48, no. D1 (2020): gkz1029, https://doi.org/10.1093/nar/gkz1029.

15. L. Gragert, A. Madbouly, J. Freeman, and M. Maiers, "Six-Locus High Resolution HLA Haplotype Frequencies Derived From Mixed-Resolution DNA Typing for the Entire US Donor Registry," *Human Immunology* 74, no. 10 (2013): 1313–1320, https://doi.org/10.1016/j.humimm.2013.06.025.

16. A. Madbouly and Y. T. Bolon, "Race, Ethnicity, Ancestry, and Aspects That Impact HLA Data and Matching for Transplant," *Frontiers in Genetics* 15 (2024): 1375352, https://doi.org/10.3389/fgene.2024.1375352.

17. V. Damotte, C. Zhao, C. Lin, et al., "Multiple Measures for Self-Identification Improve Matching Donors With Patients in Unrelated Hematopoietic Stem Cell Transplant," *Communications Medicine* 4, no. 1 (2024): 189.

18. V. Paunić, L. Gragert, J. Schneider, C. Müller, and M. Maiers, "Charting Improvements in US Registry HLA Typing Ambiguity Using a Typing Resolution Score," *Human Immunology* 77, no. 7 (2016): 542–549, https://doi.org/10.1016/j.humimm.2016.05.002.

19. S. Israeli, L. Gragert, A. Madbouly, et al., "Combined Imputation of HLA Genotype and Self-Identified Race Leads to Better Donor-Recipient Matching," *Human Immunology* 84, no. 12 (2023): 110721, https://doi.org/10.1016/j.humimm.2023.110721.

20. L. Gragert, S. R. Spellman, B. E. Shaw, and M. Maiers, "Unrelated Stem Cell Donor HLA Match Likelihood in the US Registry Incorporating HLA-DPB1 Permissive Mismatching," *Transplantation and Cellular Therapy* 29, no. 4 (2023): 244–252, https://doi.org/10.1016/j.jtct.2022.12.027.

21. M. Maiers, P. Bashyal, Y. Bolon, and L. Gragert, "P508 PY-ARD – A Swiss Army Knife of HLA Assignments," *Human Immunology* 84 (2023): 132, https://doi.org/10.1016/j.humimm.2023.08.075.

22. K. Wadsworth, M. Albrecht, R. Fonstad, S. Spellman, M. Maiers, and J. Dehn, "Unrelated Donor Search Prognostic Score to Support Early HLA Consultation and Clinical Decisions," *Bone Marrow Transplantation* 51, no. 11 (2016): 1476–1481, https://doi.org/10.1038/bmt.2016.162.

23. B. C. Shaffer, M. Gooptu, T. E. DeFor, et al., "Post-Transplant Cyclophosphamide-Based Graft-Versus-Host Disease Prophylaxis Attenuates Disparity in Outcomes Between Use of Matched or Mismatched Unrelated Donors," *Journal of Clinical Oncology* 42, no. 28 (2024): 3277–3286, https://doi.org/10.1200/JCO.24.00184.

24. A. S. Chowdhury, M. Maiers, S. R. Spellman, T. Deshpande, Y. T. Bolon, and S. M. Devine, "Existence of HLA-Mismatched Unrelated Donors Closes the Gap in Donor Availability Regardless of Recipient Ancestry," *Transplantation and Cellular Therapy* 29, no. 11 (2023): 686.e1–686.e8, https://doi.org/10.1016/j.jtct.2023.08.014.

25. S. Fingerson, M. Maiers, Y. T. Bolon, S. M. Devine, and S. R. Spellman, "Expanding Donor Options: Haploidentical Transplant Recipients Are Also Highly Likely to Have a 7/8-Matched Unrelated Donor," *Blood Advances* 8, no. 3 (2024): 758–765, https://doi.org/10.1182/bloodadvances.2023011814.

26. I. Politikos, E. Davis, M. Nhaissi, et al., "Guidelines for Cord Blood Unit Selection," *Biology of Blood and Marrow Transplantation* 26, no. 12 (2020): 2190–2196, https://doi.org/10.1016/j.bbmt.2020.07.030.

27. A. K. Keating, J. Langenhorst, J. E. Wagner, et al., "The Influence of Stem Cell Source on Transplant Outcomes for Pediatric Patients With Acute Myeloid Leukemia," *Blood Advances* 3, no. 7 (2019): 1118–1128, https://doi.org/10.1182/bloodadvances.2018025908.

28. J. N. Barker, S. M. Devlin, K. A. Naputo, et al., "High Progression-Free Survival After Intermediate Intensity Double Unit Cord Blood Transplantation in Adults," *Blood Advances* 4, no. 23 (2020): 6064–6076, https://doi.org/10.1182/bloodadvances.2020003371.

29. W. Hwang and M. Boo, *A Gift for Life: WMDA Handbook for Blood Stem Cell Donation* (World Marrow Donor Association, 2016).

30. B. E. Shaw, A. M. Jimenez-Jimenez, L. J. Burns, et al., "Three-Year Outcomes in Recipients of Mismatched Unrelated Bone Marrow Donor Transplants Using Post-Transplantation Cyclophosphamide: Follow-Up From a National Marrow Donor Program-Sponsored Prospective Clinical Trial," *Transplantation and Cellular Therapy* 29, no. 3 (2023): 208.e1–208.e6, https://doi.org/10.1016/j.jtct.2022.12.017.

31. J. J. Auletta, J. Kou, M. Chen, et al., "Real-World Data Showing Trends and Outcomes by Race and Ethnicity in Allogeneic Hematopoietic Cell Transplantation: A Report From the Center for International Blood and Marrow Transplant Research," *Transplantation and Cellular Therapy* 29, no. 6 (2023): 346.e1–346.e10, https://doi.org/10.1016/j.jtct.2023.03.007.

32. B. E. Shaw, A. M. Jimenez-Jimenez, L. J. Burns, et al., "National Marrow Donor Program-Sponsored Multicenter, Phase II Trial of HLA-Mismatched Unrelated Donor Bone Marrow Transplantation Using Post-Transplant Cyclophosphamide," *Journal of Clinical Oncology* 39, no. 18 (2021): 1971–1982, https://doi.org/10.1200/JCO.20.03502.

Supporting Information

Additional supporting information can be found online in the Supporting Information section.