Evaluation of the Spanish Population Coverage of a Prospective HLA Haplobank of Induced Pluripotent Stem Cells.

Belen Alvarez-Palomo (belen.alpa@gmail.com)
Banc de Sang i Teixits https://orcid.org/0000-0002-3894-044X

Iris Garcia-Martinez
Banc de Sang i Teixits

Jorge Gayoso
Organizacion Nacional de Trasplantes

Angel Raya
Institut d'Investigacio Biomedica de Bellvitge

Anna Veiga
Institut d'Investigacio Biomedica de Bellvitge

María Luisa Abad
Axencia Galega de Sangue, Organos e Tecidos

Adolfo Eiras
Fundacion Publica Galega de Medicina Xenomica

Maria Guzman-Fulgencio
Centro de Transfusion de la Comunidad Valenciana

Mar Luis-Hidalgo
Centro de Transfusion de la Comunidad Valenciana

Cristina Eguizabal
Centro Vasco de Transfuson y Tejidos Humanos

Silvia Santos
Centro Vasco de Transfusión y Tejidos Humanos

Antonio Balas
Centro de Transfusión de la Comunidad de Madrid

Raquel Alenda
Centro de Transfusión de la Comunidad de Madrid

Francisco Sanchez-Gordo
Centro de Transfusión, Tejidos y Células de Málaga

Laura Ponce Verdugo
Centro de Transfusión, Tejdios y Células de Málaga

Juliana Villa

Fundació i Institut de Recerca Josep Carreras
Enric Carreras
Hospital de la Santa Creu i Sant Pau Institut de Recerca
Francisco Vidal
Banc de Sang i Teixits
Alejandro Madrigal
UCL Cancer Institute, Medical School
María José Herrero
Banc de Sang i Teixits
Francesc Rudilla
Banc de Sang i Teixits
Sergi Querol
Banc de Sang i Teixits

Research

Keywords: Induced Pluripotent Stem Cells, haplobank, HLA matching, homozygous

Posted Date: August 4th, 2020

DOI: https://doi.org/10.21203/rs.3.rs-51318/v1

Version of Record: A version of this preprint was published at Stem Cell Research & Therapy on April 13th, 2021. See the published version at https://doi.org/10.1186/s13287-021-02301-0.

Submission to Stem Cells Research and Therapy

Title

Evaluation of the Spanish population coverage of a prospective HLA haplobank of Induced Pluripotent Stem Cells.

Authors

Belén Álvarez-Palomo[1,2], Iris García-Martinez[1], Jorge Gayoso[3], Angel Raya[4], Anna Veiga[4], María Luisa Abad[5], Adolfo Eiras[5], María Guzmán-Fulgencio[6], Mar Luis-Hidalgo[6], Cristina Eguizabal[7,8], Silvia Santos[7,8], Antonio Balas[9], Raquel Alenda[9], Francisco Sanchez-Gordo[10], Laura Ponce Verdugo[10], Dr. Juliana Villa[11], Dr. Enric Carreras[11], Francisco Vidal[1,12,13], Alejandro Madrigal[14,15], María José Herrero[1], Francesc Rudilla[1], Sergi Querol[1,2].

1. Banc de Sang i Teixits, Edifici Dr. Frederic Duran i Jordà, Passeig Taulat 116, 08005, Barcelona, Spain.
2. Musculoskeletal Tissue Engineering Group, Vall d'Hebron Research Institute (VHIR), Passeig de la Vall d'Hebron 129-139, 08035, Barcelona, Spain.
3. Organización Nacional de Trasplantes. Ministerio de Sanidad. C/ Sinesio Delgado 6 - Pabellón 3, 28029 Madrid.
4. Programa de Medicina Regenerativa, Institut d'Investigació Biomèdica de Bellvitge. IDIBELL, Hospital Duran i Reynals. Gran Via de l'Hospitalet, 199-203, 08908 L'Hospitalet de Llobregat, Barcelona, Spain.
5. Axencia Galega de Sangue, Órganos e Tecidos. Rúa Xoaquín Díaz de Rábago 2, 15705 Santiago, Spain.
6. Centro de Transfusión de la Comunidad Valenciana, Av. del Cid, 65-acc, 46014 Valencia, Spain.

7. Osakidetza, Centro Vasco de Transfusión y Tejidos Humanos, Barrio Labeaga, s/n. 48960 Galdakao, Spain.
8. Instituto de Investigación Sanitaria Biocruces Bizkaia, Grupo de Terapia Celular, Células Madre y Tejidos, Cruces Plaza, 48903, Barakaldo, Spain.
9. Centro de Transfusión de la Comunidad de Madrid, Avda. de la Democracia, s/n. 28032, Madrid, Spain.
10. Centro de Transfusión, Tejidos y Células de Málaga, Avda. Doctor Gálvez Ginachero s/n, 29009, Málaga, Spain.
11. REDMO / Fundació i Institut de Recerca Josep Carreras. C/Muntaner, 383 2n, 08021, Barcelona, Spain.
12. Transfusional Medicine, Vall d'Hebron Research Institute, Universitat Autònoma de Barcelona (VHIR-UAB), Barcelona, Spain.
13. CIBER de Enfermedades Cardiovasculares (CIBERCV).
14. Royal Free Hospital. Pond Street, NW3 2QG, Hampstead, UK.
15. UCL Cancer Institute, Medical School, 74 Huntley St, Bloomsbury, WC1E 6DE, London, UK.

Correspondence: Belén Alvarez Palomo, PhD, abalvarez@bst.cat; Sergi Querol, MD, PhD, squerol@bst.cat, Tel: +34-93 557 35 00 (BA-P).

Abstract

Background

iPSC (Induced Pluripotent Stem Cells) banks of iPSC lines with homozygous HLA (Human Leukocyte Antigen) haplotypes (haplobanks) are proposed as an affordable and off-the-shelf approach to allogeneic transplantation of iPSC derived cell therapies. Cord blood banks offer an extensive source of HLA-typed cells suitable for reprogramming to iPSC. Several initiatives worldwide have been undertaken to create national and international iPSC haplobanks that match a significant part of a population.

Methods

To create an iPSC haplobank that serves the Spanish population (IPS-PANIA), we have searched the Spanish Bone Marrow Donor Registry (REDMO) to identify the most frequently estimated haplotypes. From the top ten donors identified, we estimated the population coverage using the criteria of zero mismatches in HLA-A, -B and -DRB1 with different stringencies: high resolution, low resolution and beneficial mismatch.

Results

We have calculated that ten cord blood units from homozygous donors stored at the Spanish cord blood banks can provide a HLA-A, HLA-B and HLA-DRB1 matching for 28,23% of the population.

Conclusion

We confirm the feasibility of using banked cord blood units to create an iPSC haplobank that will cover a significant percentage of the Spanish and international population for future advanced therapy replacement strategies.

Keywords

Induced Pluripotent Stem Cells, haplobank, HLA matching, homozygous

Introduction

Induced pluripotent stem cells (iPSC) hold great promise in the field of regenerative medicine field by being similar to embryonic stem cells, self-renew, differentiate into any cell of the human body, free of ethical concerns and allowing autologous application (1,2). There has been extensive research and development in the field to create safe iPSC and protocols to differentiate them into clinically relevant cells for cell therapy applications. This research effort culminated in 2015 with the first successful clinical trial in Japan that used iPSC derived retinal pigmented epithelial cells (RPE) in an age-related macular degeneration (AMD) patient (RIKEN trial) (3). However, the high cost and lengthy process of iPSC production could make the use of IPSC prohibitive for many applications. Consequently, the idea of using clinically matched iPSC for HLA-A, HLA-B and HLA- DRB1 as an allogeneic treatment became more widespread (4).

Although MHC (Major Histocompatibility Complex) class I and II molecules are potentially immunogenic, HLA-DRB1, followed by HLA-B and -A are the strongest determinants of rejection of an allogeneic transplant (5). The effect of HLA-A, -B and -DRB1 mismatch on solid organ transplant rejection has been extensively documented (6). It is important to note that histocompatibility assessment for solid organ transplantation mainly takes into account HLA-A, HLA-B and HLA-DRB1 loci at a low-resolution level (7).

The level of HLA-matching requirements ranges depending on the type of transplant, with a minimum score of 9/10 (HLA-A, -B, -C, -DRB1, -DQB1) for bone marrow transplantation of unrelated donors, a minimum score of 6/8 (HLA-A, -B, -C, -DRB1) for cord blood (8) to kidney transplants in which several degrees of

106 “beneficial match” are considered, from one mismatch in HLA-A or -B to just
107 matching HLA-DRB1 alone (9). Currently, for hematopoietic progenitor
108 transplantation, the common use is high resolution sequencing (two fields) of five
109 HLA loci (10) using Next Generation Sequencing (NGS) for typing.

110 The relevance and stringency of HLA-matching for iPSC-derived cells for clinical
111 transplantation has been discussed extensively (11). Unlike transplantation of
112 hematopoietic progenitors, iPSC derivatives do not have contaminating immune
113 cells and therefore the level of compatibility required is measured only for host vs.
114 graft because the graft does not have in this case the ability to react against the host
115 mismatched HLA antigens. HLA-A, -B and -DRB1 matching have shown to confer a
116 clear advantage over totally allogeneic transplant, with different degrees of
117 immunological responses observed indicating the need for immunosuppression
118 (12,4).

119 Banking of HLA-typed pluripotent cells for matching a wide proportion of a
120 population was first proposed for embryonic stem cells (13). The authors also
121 proposed the use of homozygous donors for common HLA-A, -B and -DRB1
122 haplotypes as a way to provide HLA match for a reasonable percentage of the target
123 population with a limited number of cell lines. Later, with the appearance of iPSC,
124 Nakatsuji and colleagues proposed the use of banked cord blood as a source of
125 HLA-typed cells for the construction of HLA homozygous iPSC banks (haplobanks)
126 (14). It was estimated that a haplobank with only 30 iPSC lines would be able to
127 cover 82,2% of the Japanese population and 50 lines, 90,7% (14). Other studies
128 have calculated the coverage of haplobanks for the UK population (15), South Korea
129 (16), China (17) and the U.S. (18). iPSC haplobanks created from cord blood and
130 peripheral blood donors are already a reality in South Korea and Japan (16,19,20).

The HLA haplotype landscape has been investigated before with small cohorts of patients and healthy individuals (21,22,23) and a larger cohort of 5,458 units of cord blood from the Barcelona Cord Blood Bank, HLA-typed in high resolution (24).

The Spanish project IPS-PANIA aims at creating an iPSC haplobank of at least seven clinical-grade lines to provide maximum coverage to the Spanish population. To identify the haplotypes providing maximum coverage (probability of zero mismatches in HLA-A, -B and -DRB1) we have searched a large cohort of 32.000 adult bone marrow donors and calculated the estimated coverage for a study population of 418.981 individuals including cord blood donors plus bone marrow donors from REDMO (Red de Donantes de Medula Osea) registry. We have concluded that a haplobank of seven lines would cover 23.69% of the Spanish population and ten lines would cover 28.23%.

Materials and methods

Study cohort and ethics approval

The REDMO includes the HLA typing of all cord blood donations and all the adult bone marrow donors of the Spanish population. For population coverage studies the target population consisted of all the cord blood plus all the adult bone marrow donors in REDMO. The consultation of the HLA data was approved by the Ethics Committee for Research with Medicines from Vall d'Hebron Hospital (Barcelona, Spain) and the Transplantation and Regenerative Medicine Commission of the Spanish National Health System.

HLA typing and haplotype frequency determination

High-resolution typing was performed by Sanger sequencing in an ABI PRISM 3130xl Genetic Analyzer (Thermo Fisher) and/or Next Generation Sequencing

(NGS) in a MiSeq platform (Illumina) or in Ion GeneStudio S5 System (Thermo Fisher) for *HLA-A, -B, -C, -DRB1, -DRB3/4/5, -DQB1 and -DPB1* genes. The resulting sequences were analyzed using Assign 4.7.1 (CareDX), Type Stream Visual (One Lambda) and NGSengine (GenDX, 2.16), depending on the used procedure. Low resolution was performed using a PCR-SSO (Luminex) based method for HLA-A, B, C, DRB1, DQB1 genes.

The expectation-maximization algorithm implemented in the Arlequin software (version 3.5.2.2) (25) was used to estimate maximum-likelihood haplotype frequencies, considering the high-resolution (4-digit) allelic frequencies of three (A, B and DRB1) and five HLA genes (A, B, C, DRB1 and DQB1), from 30,000 and 27,000, respectively, randomly selected adult subjects of the REDMO project.

Screening and selection of HLA haplotype homozygous donors

All the cord blood units in REDMO were studied for the identification of potential HLA-A, HLA-B and HLA-DRB1 homozygous cord blood donors. Selection and classification of homozygous units were performed by simple counts with Microsoft Excel.

Calculation of match coverage

To estimate the Spanish population HLA matching coverage of a 10 iPSC haplobank we counted the number individuals with zero mismatches in HLA-A, HLA-B and HLA-DR when compared with the top ten haplotypes selected by frequency. The population cohort was composed of the combined data of the adult bone marrow donors and cord blood donors in the REDMO collection. The calculation was done either in two digits HLA typing (low plus high resolution originally) on 418.980 individuals or in four digits (only high resolution) on 56.798 individuals. To assess the haplotype match benefit (coverage) in the whole REDMO sample (including adult bone marrow donors and cord blood donors), an iterative function

was generated with R software. Briefly, in each round, this function identified the haplotype from all estimated haplotypes that matched the highest number of subjects from the whole sample and then, extracted both the haplotype and the matching individuals, and recomputed again with the remaining data. The identification of the best haplotype in every round was assessed by calculating the total matching occurrences between each of the alleles of each locus for all estimated haplotypes with one of the two alleles of each locus in every subject.

To calculate the percentage of coverage, all raw counts were divided by the corresponding total sample size of each analysis and multiplied by 100.

The calculation of a "beneficial match" coverage by allowing partial match (namely HLA-A and HLA-DRB1, HLA-B and HLA-DRB1 and only HLA-DRB1) was done with Access software (Microsoft).

Results

Haplotype frequencies

A cohort of 30.000 randomized high-resolution HLA typing adult bone marrow donors from the REDMO identified a total of 8,478 different haplotypes for HLA-A, HLA-B and HLA-DRB1 (Fig. 1A and Supplementary Excel file for the complete list). Five of them presented a frequency 1% or higher, namely: A*29:02~B*44:03~DRB1*07:01 (3,12%), A*01:01~B*08:01~DRB1*03:01 (2,48%), A*30:02~B*18:01~DRB1*03:01 (1,99%), A*03:01~B*07:02~DRB1*15:01 (1,34%) and A*33:01~B*14:02~DRB1*01:02 (1,00%). Haplotype frequencies show a steep decline after the first five (Fig. 1B).

We performed the same analysis for five genes, namely HLA-A, HLA-B, HLA-C, HLA-DRB1 and HLA-DQB1 with a 27.000 cohort from the REDMO registry. Only the

top six positions corresponded with those haplotypes identified in the three genes analysis and only four had a frequency above 1%: A*29:02~ C*16:01~B*44:03~ DRB1*07:01~DQB1*02:02 (2.95%), A*01:01~ C*07:01~B*08:01~ DRB1*03:01~DQB1*02:01 (2.52%), A*30:02~ C*05:01~B*18:01~ DRB1*03:01~DQB1*02:01 (1.89%), and A*03:01~ C*07:02~B*07:02~DRB1*15:01~DQB1*06:02 (1.34%) (Sup. Fig. 1).

Identification of homozygous donors

To identify homozygous CB units available in the Spanish banks that can be candidates as source cells to create the iPSC haplobank, we looked into the Spanish registry, which comprises 52,220 cord blood donations in Spain. HLA was typed in low resolution (serological and PCR based) for 42.801 units and high resolution (NGS) for 9,419 units. We identified 322 cord blood units homozygous for HLA-A, HLA-B and HLA-DRB1 (0.62% of the units) representing 111 different haplotypes (Sup. Fig. 2). So far, 109 of the homozygous units were verified by high-resolution HLA typing and represented 43 different haplotypes. Not to deplete the cord blood banks of any haplotype for hematopoietic stem cell transplantation, we only considered those haplotypes with two or more units banked as candidates for the iPSC bank (Fig. 2). Thirty-one haplotypes were represented at least twice in low resolution and 11 in high resolution. The top ten positions in the most frequent haplotypes from the adult bone marrow donor study were all represented in at least two units genotyped in low resolution and nine of them have confirmed units by high resolution.

Recipients matching coverage

We tested the first 10 most frequent haplotypes for three genes, HLA-A, HLA-B and HLA-DRB1 and we found that the cumulative coverage was 27.84% of the population in high resolution and 31.27% in low resolution (Fig. 3A, B). We also

explored the “maximum coverage” approach for 10 haplolines selecting not the most frequent haplotypes but those that would optimize the coverage. The choice of haplotypes with this approach did not change for the top four positions, and the following six, although in a different order, only one new haplotype was considered: A*02:01~B*18:01~DRB1*03:01. The cumulative coverage in high resolution was slightly increased to 28.23% in high resolution and to 31.87% in low resolution (Fig. 3A, B). Since we did not identify any homozygous CB unit in the Spanish banks with A*02:01~B*18:01~DRB1*03:01 we estimated the optimized coverage considering only the available units and then the accumulated coverage in high resolution is 27.95% and 32.58% in low resolution. To calculate how many haplolines would be needed to cover close to 100% of the Spanish population, we repeated this iterative process with all the estimated haplotypes in our study population. We found that 100 haplolines would cover 65.94% of the Spanish population and that 630 haplolines would be necessary to cover 90% of the potential recipients (Fig 3C). We also did this calculation with 5 genes: HLA-A, HLA-B, HLA-C, HLA-DRB1 and HLA-DQB1, and then the numbers shift to 100 units to cover 50% of the population and 897 units needed to cover 90% (Sup. Fig. 3).

Beneficial match

As different iPSC-derived cell types and different engraftment sites might present different HLA matching requirements, we introduced the approach of “beneficial match” and estimated the coverage in the case that either the HLA-B allele (A-DRB1) or the HLA-A (B-DRB1) or both (DRB1) did not match -considering HLA-DRB1 as the most immunogenic of the alleles-. The cumulative coverage for the top 10 most frequent HLA-A, HLA-B, HLA-DRB1 haplotypes showed a substantial increase in population coverage to 44.5% for ADR and 40.38% for BDR; and, as expected, a dramatic increase to 69.14% when only matching HLA-DRB1 (DRB1)

(Fig. 4A). When the “maximum coverage” list was analyzed the % increased to 46.54 (A-DRB1), 40.38 (B-DRB1) and 64.67 (DRB1).

Discussion

Currently, clinical-grade iPSC haplobanks are already available in South Korea and Japan (16,19,20,26) and similar initiatives are being undertaken in other countries such as Australia (27). The effectiveness of the approach will come from a global collaborative to share haplolines and this implies an effort to standardize the production and quality control by the different banks (28,29).

The IPS-PANIA project aims at developing a haplobank that can serve a significant percentage of the Spanish population with HLA-matching iPSC lines that can be used as starting material for iPSC derived cell therapies. We have estimated that a minimum of seven lines can be developed in the initial phase.

A previous study of the estimated most frequent haplotypes in the Spanish bone marrow donor database identified five haplotypes with a frequency ranging from higher than 1% to the 6th position in the list were the same as previously described for the Barcelona cord blood bank study (24). The differences in the lower positions might reflect the higher accuracy of the larger database or geographical or generational differences. The haplotype with the top frequency is; A*29:02~B*44:03~DRB1*07:01 and the third one; A*30:02~B*18:01~DRB1*03:01 are very common in the western Mediterranean region these haplotypes are less common in Italy creating a border with south-eastern Europeans that express lower frequently (30). A*29:02~B*44:03~DRB1*07:01 is also a common haplotype in the Hispanic population in the USA, but less frequent for European ancestry and extremely rare for Asian ancestry (17, 31). The second and fourth most frequent

285 haplotypes; A*01:01~B*08:01~DRB1*03:01 and A*03:01~B*07:02~DRB1*15:01
286 respectively, are among the most common in the northern and central European
287 population. From the top 10 most frequent haplotypes in Spain, seven are in the top
288 30 described for the UK (15). As expected, they are not identified with the most
289 frequent haplotypes identified for Korea and Japan (16,14).

290 The identification of homozygous units confirmed that the Spanish public cord banks
291 contained enough homozygous units that could be used to build an iPSC haplobank
292 to cover the top haplotypes without compromising the availability of any lifesaving
293 unit needed for hematopoietic progenitor transplantation. We found a 0.62% of the
294 cord blood units were homozygous, a significantly lower number as compared for
295 that reported for the South Korean population (0.79%), which might reflect the wider
296 heterogeneity of the HLA genetics in the Spanish population.

297 The estimated population coverage of the most frequent haplotypes seems to also
298 reflect a higher HLA genetic variability in the Spanish population as compared to
299 other populations: the top two haplotypes in Spain cover 6.3% and cumulative
300 11.2%, while for Korea the percentages are 9.2 and 14.5 and for UK 16.9% and
301 26.4%. A ten cell line haplobank would cover 28% in four-field resolution and 32% in
302 two-field resolution, while in UK or Japan, the reported coverages for ten cell lines
303 are closer to 50% of the population. The “maximum coverage” approach for the
304 selection of the top ten candidates, opposed to following just the higher haplotype
305 frequency order, improved the coverage for the lower positions but the effect was
306 almost lost when the ten haplolines were considered. When calculating how many
307 haplotypes would be needed for almost complete coverage of the Spanish
308 population, we found that 631 would be needed to cover 90%, again revealing a
309 much wider variety of existing haplotypes as compared to other populations
310 analyzed in other studies, such as Korea that found 90% covered with less than 200
311 cell lines (16).

Wider coverage of iPSC haplolines by lowering the HLA stringency might be worth consideration for an iPSC bank. It is still early days to know what the real HLA matching requirements of iPSC derived products will be as the first clinical trial carried out at the Kobe City Medical Center in collaboration with Osaka University, using allogeneic iPSC-derived retinal pigmented epithelial cells of a HLA-A,-B,-DRB1-matching haplotype has not been reported yet. It is clear that HLA matching is advantageous and will reduce the degree of immunosuppression although this will be determined by the type of cell and the transplantation sites. Studies performed in non-human primates matching MHC antigens equivalent to HLA-A, -B and -DRB1 have revealed from mild immune infiltration in iPSC-RPE implanted in the retina (4) to a significant reaction to the allograft requiring immunosuppression in the central nervous system (10). Unlike hematopoietic progenitors and organ transplants, iPSC derived cells will be free of contaminating T-cells, except for a small possibility in the case of hematological lineage derivation, and no graft-versus-host disease is expected, making sense to consider less HLA-match stringent scenarios closer to solid organ transplants like the kidney. The population coverage when considering haplotypes in two-field resolution was not much increased as compared to four-field resolution, highlighting the predominance of certain subgroups of haploptypes. When the beneficial match was considered, as expected, the gains were much greater when allowing for HLA-A mismatch or HLA-B mismatch, and to match up to 67% when considering only match for HLA-DRB1. The choices of HLA matching stringency requirements will be determined by future evidence for the different cell types derived from iPSC and the different transplantation sites.

Taken together, a 30% population match for a ten cell line haploblank is a significant proportion of potential patients who may benefit from the cell bank. This justifies the construction of the Spanish haplobank, more so when considering a worldwide effort to share cells internationally with other banks that might contain less frequent

haplotypes. Also, as the production of clinical-grade iPSC lines is optimized and streamlined it will become easier and more affordable to increase the number of haplolines in the bank to reach several dozens and cover a much wider proportion of the population. With this study, we have investigated the feasibility to provide homozygous cord blood units to create an iPSC bank of a reduced number of haplolines that will serve a significant percentage of the Spanish and international population. Besides RPE cells for the treatment of AMD, several other iPSC-derived cells are presently being tested in clinical trials to treat conditions such as spinal cord injury, Parkinson's disease, Graft-versus-host-disease, heart failure, or cancer. Clinical-grade iPSC are intended to be used as starting material for future clinical trials and cell therapy products, accelerating the application of iPSC-based therapies soon.

Conclusion

With the present study we have been able to confirm that HLA homozygous cord blood units stored in Spanish cord blood banks can provide for the construction of an iPSC bank that is useful for a significant part of the Spanish population (about 28%). The haplotypes identified as providing the greatest coverage, may also be useful for other populations such as Europe and North America.

Declarations

Ethics approval and consent to participate

The consultation of the HLA data was approved by the Ethics Committee for Research with Medicines from Vall d'Hebron Hospital (Barcelona, Spain) and the Transplantation and Regenerative Medicine Commission of the Spanish National Health System.

Consent for publication

Not applicable

Availability of data and material

All the presented data is available for consultation

Competing interests

The authors declare no competing interests.

Funding

This research was funded by the Spanish Ministry of Science, Innovation and Universities, National Plan for Scientific and Technical Research and Innovation 2013-2016 RETOS COLABORACIÓN Program 2017: exp. RTC-2017-6000-1. 2017-2021.

Authors' contribution

BAP, has participated in the design of the work; the acquisition, analysis, interpretation of data; and draft of the document. SQ has participated in the design of the work, analysis, interpretation of data and draft of the document. IG-M, has participated in the acquisition, analysis, interpretation of the data and revision of the document. FR has participated in the design of the work, acquisition, analysis, and interpretation of data and revision of the work. FV, AM and MJH have participated in the design of the work, interpretation of data and revision of the document. AR and AV have participated in the design of the work, and revision of the document. JG, MLA, AE, MG-F, ML-H, CE, SS, AB, RA, FSG, LPV, EC and JV have participated in the acquisition of the data and revision of the document.

Acknowledgements

We would like to acknowledge Dr. Emma Enrich Randé for discussion of the results, Mrs. Elisabet Tahull Navarro for administrative and management help and Dr. Michael J. Edel for proof reading. We would also like to thank the cord blood donors,

the donation programs professionals, and all the cord blood banks staff for their invaluable dedication to helping bring hope to many patients in need.

References

1- Robinton DA, Daley GQ. **The promise of induced pluripotent stem cells in research and therapy.** *Nature* 2012; 481(7381):295-305. doi: 10.1038/nature10761. Review.

2- Eguizabal C, Aran B, Chuva de Sousa Lopes SM, et al. **Two decades of embryonic stem cells: a historical overview**. *Hum Reprod Open* 2019; 2019(1):hoy024. doi: 10.1093/hropen/hoy024. eCollection 2019.

3- Mandai M, Watanabe A, Kurimoto Y, et al. **Autologous Induced Stem-Cell-Derived Retinal Cells for Macular Degeneration.** *N Engl J Med* 2017; 376(11):1038-1046. doi: 10.1056/NEJMoa1608368.

4- Sugita S, Iwasaki Y, Makabe K, et al. **Lack of T Cell Response to iPSC-Derived Retinal Pigment Epithelial Cells from HLA Homozygous Donors**. *Stem Cell Reports* 2016; 7(4):619-634. doi:10.1016/j.stemcr.2016.08.011.

5- Feucht HE. **The humoral immune response towards HLA class II determinants in renal transplantation.** *Kidney International* 1996; 50:1464-1475.

6- Zachary AA and Leffell MS. **HLA Mismatching Strategies for Solid Organ Transplantation – A Balancing Act.** *Front Immunol* 2016; 7: 575. doi: 10.3389/fimmu.2016.00575

7- Copley HC, Elango M, Kosmoliaptsis. **Assessment of human leukocyte antigen immunogenicity - current methods, challenges and opportunities.** *Current Opinion in Organ Transplantation* 2018; 23(4): 477-485. doi: 10.1097/MOT.0000000000000544

8- Howard A, Fernandez-Vina MA, Appelbaum FR et al. **Recommendations for Donor HLA Assessment and Matching for Allogeneic Stem Cell Transplantation: Consensus Opinion of the Blood and Marrow Transplant Clinical Trials Network (BMT CTN).** *Biol Blood Marrow Transplant* 2015; 21(1): 4–7. doi: 10.1016/j.bbmt.2014.09.017

9- Lim WH, Chadban SJ, Clayton P et al. **Human leukocyte antigen mismatches associated with increased risk of rejection, graft failure, and death independent of initial immunosuppression in renal transplant recipients.** *Clin Transplant* 2012;26(4):E428.

10- Fürst D, Müller C, Vucinic V et al. **High-resolution HLA matching in hematopoietic stem cell transplantation: a retrospective collaborative analysis.** *Blood.* 2013 Oct 31;122(18):3220-3229. doi: 10.1182/blood-2013-02-482547.

11- Taylor CJ, Bolton EM and Bradley JA. **Immunological considerations for embryonic and induced pluripotent stem cell banking.** *Phil. Trans. R. Soc. B* 2011; 366, 2312–2322. doi:10.1098/rstb.2011.0030

12- Aron Badin R, Bugi A, Williams S et al. **MHC matching fails to prevent long-term rejection of iPSC-derived neurons in non-human primates.** *Nat Commun* 2019; 10: 4357. Published online 2019 Sep 25. doi: 10.1038/s41467-019-12324-0

13- Taylor CJ, Bolton EM, Pocock S et al. **Banking on human embryonic stem cells: estimating the number of donor cell lines needed for HLA matching.** *Lancet* 2005; 366: 2019-2025.

14- Nakatsuji N, Nakajima F, Tokunaga K. **HLA-haplotype banking and iPS cells.** *Nature Biotechnology* 2008; 26(7):739-740.

15- Taylor CJ, Peacock S, Chaudhry AN, et al. **Generating an iPSC Bank for HLA-Matched Tissue Transplantation Based on Known Donor and**

Recipient HLA Types. *Cell Stem Cell* 2012; 11(2):147-152. doi:10.1016/j.stem.2012.07.014

16- Lee S.; Huh JY, Turner, DM et al. **Repurposing the Cord Blood Bank for Haplobanking of HLA-Homozygous iPSCs and Their Usefulness to Multiple Populations.** *Stem Cells* 2018; 36:1552–1566, doi:10.1002/stem.2865.

17- Lin G, Xie Y, Ouyang Q et al. **HLA-matching potential of an established human embryonic stem cell bank in China.** *Cell Stem Cell* 2009; 5(5):461-465. doi: 10.1016/j.stem.2009.10.009.

18- Gourraud PA, Gilson L, Girard M et al. **The role of human leukocyte antigen matching in the development of multiethnic "haplobank" of induced pluripotent stem cell lines.** *Stem Cells* 2012; 30(2):180-186. doi: 10.1002/stem.772.

19- Rim YA, Park N, Nam Y, et al. **Recent progress of national banking project on homozygous HLA-typed induced pluripotent stem cells in South Korea.** *J Tissue Eng Regen Med.* 2018 Mar;12(3):e1531-e1536. doi: 10.1002/term.2578. Epub 2017 Nov 10.

20- Umekage M, Sato Y1, Takasu N1. **Overview: an iPS cell stock at CiRA.** *Inflamm Regen* 2019 Sep 2;39:17. doi: 10.1186/s41232-019-0106-0. eCollection 2019.

21- Balas A, García-Sánchez F, Vicario JL. **Allelic and haplotypic HLA frequency distribution in Spanish hematopoietic patients. Implications for unrelated donor searching.** *Tissue Antigens* 2011; 77(1):45-53. doi: 10.1111/j.1399-0039.2010.01578.x.

22- Montero-Martín G., Mallempati K, Gangavarapu S et al. **High-resolution characterization of allelic and haplotypic HLA frequency distribution in a Spanish population using high-throughput next-generation sequencing.** *Human Immunology* 2019; 80(7): 429-436.

23- Roura S, Rudilla F, Gastelurrutia P et al. **Determination of HLA-A, -B, -C, -DRB1 and -DQB1 allele and haplotype frequencies in heart failure patients.** *ESC Heart Fail* 2019; 6(2):388-395. doi: 10.1002/ehf2.12406.

24- Enrich E, Campos E, Martorell L, Herrero MJ, Vidal F, Querol S, Rudilla F. **HLA-A, -B, -C, -DRB1, and -DQB1 allele and haplotype frequencies: An analysis of umbilical cord blood units at the Barcelona Cord Blood Bank.** *HLA.* 2019 Oct;94(4):347-359. doi: 10.1111/tan.13644. Epub 2019 Aug 6

25- Excoffier L, Laval G, Schneider S. **Arlequin (version 3.0): an integrated software package for population genetics data analysis.** *Evol Bioinform Online.* 2007; 1:47-50.

26- Ortuño-Costela MDC, Cerrada V, García-López M et al. **The Challenge of Bringing iPSCs to the Patient.** *Int J Mol Sci*. 2019; 20(24). pii: E6305. doi: 10.3390/ijms20246305.

27- Abberton KM and Elwood NJ. **Establishing an Australian Bank of Cord Blood-Derived Induced Pluripotent Stem Cell Lines: Ethics, Re-consent, and Progress Towards Cellular Therapies.** *Stem Cells Transl Med* 2019; 8 (Suppl 1): S22. doi: 10.1002/sctm.12565.

28- Sullivan, S.; Stacey, G.N.; Akazawa, C et al. **Quality control guidelines for clinical-grade human induced pluripotent stem cell lines.** *Regen Med* 2018, 13:859–866, doi:10.2217/rme-2018-0095.

29- Alvarez-Palomo B, Vives J, Casaroli-Marano RP et al. **Adapting Cord Blood Collection and Banking Standard Operating Procedures for HLA-Homozygous Induced Pluripotent Stem Cells Production and Banking for Clinical Application.** *J Clin Med* 2019; 8(4). pii: E476. doi: 10.3390/jcm8040476.

30- Alcoceba M, Marín L, Balanzategui A et al. **Frequency of HLA-A, -B and -DRB1 specificities and haplotypic associations in the population of**

Castilla y León (northwest-central Spain). *Tissue Antigens* 2011; 78(4):249-255. doi: 10.1111/j.1399-0039.2011.01766.x.

31- Pappas DJ, Gourraud PA, Le Gall C et al. **Proceedings: Human leukocyte antigen haplo-homozygous induced pluripotent stem cell haplobank modeled after the california population: Evaluating matching in a multiethnic and admixed population.** *Stem Cells Transl Med* 2015; 4:413–418, doi:10.5966/sctm.2015-0052.

Figure legends

Figure 1. A. Top 30 ranking HLA-A,-B,-DRB1 estimated haplotypes by % of frequency at the Spanish Bone Marrow Donor Registry. B. Graphic representation of the % of frequency of the top 100 HLA-A, -B, -DRB1 estimated haplotypes.

Figure 2. HLA haplotypes and number of units (N.) of HLA-A, HLA-B and HLA-DRB1 homozygous found in the Spanish registry of public cord blood Banks and present at least in two units. LR = low resolution; HR= high resolution.

Figure 3. Estimated percentage and cumulative percentage of HLA-matched individuals in the Spanish population with a panel of 10 homozygous donors in high resolution (A) and low resolution (B). Estimated numbers of iPSC lines homozygous for HLA-A, -B, -DRB1 (haplolines) to cover the Spanish population (C).

Figure 4. Estimated cumulative percentage of matched individuals in the Spanish population considering a “beneficial match” with a panel of 10 homozygous donors

523 from Spanish cord blood Banks in high resolution allowing for a mismatch in HLA-B
524 (A-DRB1), HLA-A (B-DRB1) or HLA-A and HLA-B (DRB1).

525

526 Supplementary figure 1. Top 30 ranking HLA-A, -B, -C, -DRB1 and -DQB1
527 estimated haplotypes by % of frequency at the Spanish Bone Marrow Donor
528 Registry.

529

530 Supplementary figure 2. HLA types and number of units (N.) of HLA-A, HLA-B and
531 HLA-DRB1 homozygous found in the Spanish registry of public cord blood banks.

532

533 Supplementary Figure 3. Estimated numbers of iPSC lines homozygous for HLA-A, -
534 C, -B, -DRB1 and -DQB1 (haplolines) to cover the Spanish population.

Figures

A

	Haplotypes	Freq
1	A*29:02-B*44:03-DRB1*07:01	0,031147
2	A*01:01-B*08:01-DRB1*03:01	0,02621
3	A*30:02-B*18:01-DRB1*03:01	0,019625
4	A*03:01-B*07:02-DRB1*15:01	0,013517
5	A*33:01-B*14:02-DRB1*01:02	0,009568
6	A*23:01-B*44:03-DRB1*07:01	0,008988
7	A*01:01-B*57:01-DRB1*07:01	0,008506
8	A*02:01-B*44:03-DRB1*07:01	0,008328
9	A*02:01-B*07:02-DRB1*15:01	0,007709
10	A*02:01-B*51:01-DRB1*11:01	0,006861
11	A*24:02-B*07:02-DRB1*15:01	0,005897
12	A*02:01-B*18:01-DRB1*03:01	0,005894
13	A*02:01-B*18:01-DRB1*11:04	0,005741
14	A*03:01-B*35:01-DRB1*01:01	0,005563
15	A*02:01-B*44:02-DRB1*13:01	0,005297
16	A*30:01-B*13:02-DRB1*07:01	0,005078
17	A*02:01-B*51:01-DRB1*13:01	0,004811
18	A*02:01-B*08:01-DRB1*03:01	0,004659
19	A*02:01-B*44:02-DRB1*04:01	0,004565
20	A*11:01-B*35:01-DRB1*01:01	0,004559
21	A*02:01-B*07:02-DRB1*01:03	0,004226
22	A*02:01-B*51:01-DRB1*07:01	0,004176
23	A*02:01-B*50:01-DRB1*07:01	0,004147
24	A*02:01-B*44:02-DRB1*01:01	0,003983
25	A*02:01-B*51:01-DRB1*08:01	0,003942
26	A*24:02-B*35:02-DRB1*11:04	0,003848
27	A*25:01-B*18:01-DRB1*15:01	0,003741
28	A*02:01-B*14:02-DRB1*01:02	0,003733
29	A*11:01-B*27:05-DRB1*01:01	0,003486
30	A*02:05-B*50:01-DRB1*07:01	0,003453

B

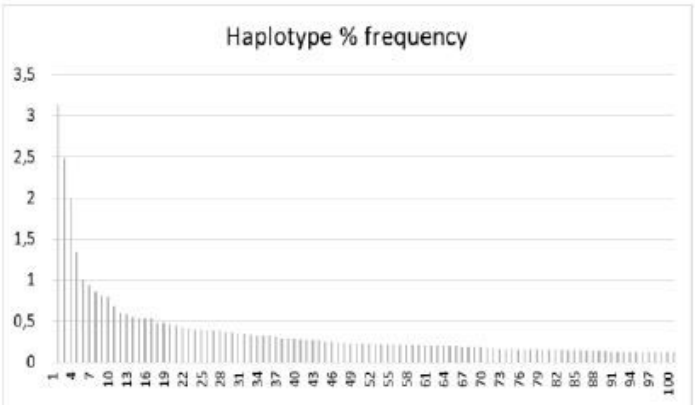

Figure 1

A. Top 30 ranking HLA-A,-B,-DRB1 estimated haplotypes by % of frequency at the Spanish Bone Marrow Donor Registry. B. Graphic representation of the % of frequency of the top 100 HLA-A, -B, -DRB1 estimated haplotypes.

Haplotype	N. LR	Haplotype	N. HR
A*29-B*44-DR*07	61	A*29:02-B*44:03-DR*07:01	29
A*30-B*18-DR*03	44	A*30:02-B*18:01-DR*03:01	23
A*01-B*08-DR*03	32	A*01:01-B*08:01-DR*03:01	18
A*03-B*07-DR*15	17	A*03:01-B*07:02-DR*15:01	9
A*02-B*44-DR*07	9	A*02:01-B*44:03-DR*07:01	5
A*02-B*51-DR*11	9	A*02:01-B*51:01-DR*11:01	2
A*02-B*44-DR*04	7	A*02:01-B*44:02-DR*04:01	1
A*23-B*44-DR*07	6	A*23:01-B*44:03-DR*07:01	1
A*24-B*35-DR*11	6		0
A*33-B*14-DR*01	6	A*33:01-B*14:02-DR*01:02	2
A*02-B*07-DR*15	5	A*02:01-B*07:02-DR*15:01	2
A*02-B*18-DR*11	4	A*02:01-B*18:01-DR*11:04	3
A*02-B*44-DR*01	4		0
A*02-B*51-DR*04	4	A*02:01-B*51:01-DR*04:01	1
A*01-B*57-DR*07	3	A*01:01-B*57:01-DR*07:01	3
A*02-B*15-DR*04	3	A*02:01-B*15:09-DR*04:03	1
A*02-B*44-DR*13	3		0
A*02-B*51-DR*13	3	A*02:01-B*51:01-DR*13:01	1
A*11-B*27-DR*01	3	A*11:01-B*27:05-DR*01:01	1
A*25-B*18-DR*15	3	A*25:01-B*18:01-DR*15:01	2
A*02-B*07-DR*01	2		0
A*02-B*18-DR*15	2		0
A*02-B*44-DR*11	2		0
A*02-B*48-DR*09	2		0
A*02-B*49-DR*04	2	A*02:01-B*49:01-DR*04:05	1
A*02-B*50-DR*07	2		0
A*03-B*44-DR*07	2		0
A*11-B*35-DR*14	2	A*11:01-B*35:01-DR*14:54	1
A*11-B*52-DR*15	2		0
A*24-B*07-DR*15	2	A*24:02-B*07:02-DR*15:01	2
A*26-B*38-DR*13	2	A*26:01-B*38:01-DR*13:01	1

Figure 2

HLA haplotypes and number of units (N.) of HLA-A, HLA-B and HLA-DRB1 homozygous found in the Spanish registry of public cord blood Banks and present at least in two units. LR = low resolution; HR= high resolution.

A

HR	Haplotypes	Recipients matched %	Cumulative %
	Max frequency		
1	A*29:02-B*44:03-DRB1*07:01	6,27	6,27
2	A*01:01-B*08:01-DRB1*03:01	4,89	11,16
3	A*30:02-B*18:01-DRB1*03:01	3,68	14,84
4	A*03:01-B*07:02-DRB1*15:01	2,79	17,63
5	A*33:01-B*14:02-DRB1*01:02	1,81	19,44
6	A*23:01-B*44:03-DRB1*07:01	1,63	21,07
7	A*01:01-B*57:01-DRB1*07:01	1,48	22,55
8	A*02:01-B*44:03-DRB1*07:01	1,86	24,42
9	A*02:01-B*07:02-DRB1*15:01	1,85	26,27
10	A*02:01-B*51:01-DRB1*11:01	1,57	27,84
	Max coverage		
1	A*29:02-B*44:03-DRB1*07:01	6,27	6,27
2	A*01:01-B*08:01-DRB1*03:01	4,89	11,16
3	A*30:02-B*18:01-DRB1*03:01	3,68	14,84
4	A*03:01-B*07:02-DRB1*15:01	2,79	17,63
5	A*02:01-B*44:03-DRB1*07:01	2,39	20,02
6	A*02:01-B*07:02-DRB1*15:01	1,88	21,91
7	A*33:01-B*14:02-DRB1*01:02	1,78	23,69
8	A*02:01-B*51:01-DRB1*11:01	1,57	25,26
9	A*02:01-B*18:01-DRB1*03:01	1,53	26,79
10	A*01:01-B*57:01-DRB1*07:01	1,44	28,23

B

LR	Haplotypes	Recipients matched %	Cumulative %
	Max frequency		
1	A*29-B*44-DRB1*07	6,30	6,30
2	A*01-B*08-DRB1*03	4,83	11,13
3	A*30-B*18-DRB1*03	3,96	15,08
4	A*03-B*07-DRB1*15	3,07	18,15
5	A*33-B*14-DRB1*01	1,95	20,10
6	A*23-B*44-DRB1*07	1,64	21,74
7	A*01-B*57-DRB1*07	1,61	23,34
8	A*02-B*44-DRB1*07	3,17	26,52
9	A*02-B*07-DRB1*15	2,11	28,63
10	A*02-B*51-DRB1*11	2,64	31,27
	Max coverage		
1	A*29-B*44-DRB1*07	6,30	6,3
2	A*01-B*08-DRB1*03	4,83	11,13
3	A*30-B*18-DRB1*03	3,96	15,08
4	A*03-B*07-DRB1*15	3,07	18,15
5	A*02-B*44-DRB1*07	3,79	21,95
6	A*02-B*07-DRB1*15	2,14	24,09
7	A*33-B*14-DRB1*01	1,91	26,00
8	A*02-B*51-DRB1*11	2,65	28,65
9	A*02-B*18-DRB1*03	1,72	30,37
10	A*01-B*57-DRB1*07	1,50	31,87

C

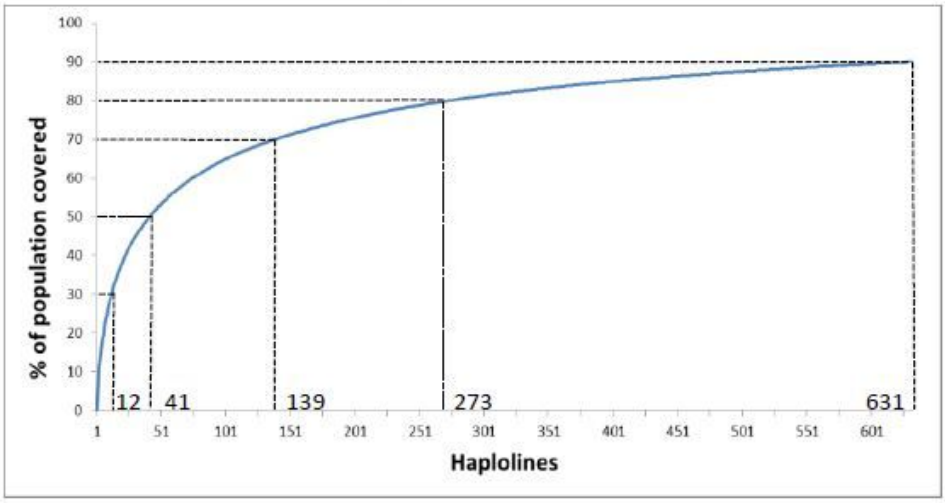

Figure 3

Estimated percentage and cumulative percentage of HLA-matched individuals in the Spanish population with a panel of 10 homozygous donors in high resolution (A) and low resolution (B). Estimated numbers of iPSC lines homozygous for HLA-A, -B, -DRB1 (haplolines) to cover the Spanish population (C).

	Haplotypes	Recipients matched (cumulative %)		
	Max frequency	A-DRB1	B-DRB1	DRB1
1	A*29:02-B*44:03-DRB1*07:01	7,44	11,74	29,18
2	A*01:01-B*08:01-DRB1*03:01	13,97	19,30	46,93
3	A*30:02-B*18:01-DRB1*03:01	17,74	25,75	46,93
4	A*03:01-B*07:02-DRB1*15:01	21,66	31,69	57,85
5	A*33:01-B*14:02-DRB1*01:02	23,64	35,69	62,32
6	A*23:01-B*44:03-DRB1*07:01	25,98	35,69	62,32
7	A*02:01-B*44:03-DRB1*07:01	29,51	37,97	62,32
8	A*01:01-B*57:01-DRB1*07:01	36,72	37,97	62,32
9	A*02:01-B*07:02-DRB1*15:01	40,77	37,97	62,32
10	A*02:01-B*51:01-DRB1*11:01	44,50	40,38	69,14
	Max coverage	A-DRB1	B-DRB1	DRB1
1	A*29:02-B*44:03-DRB1*07:01	7,44	11,74	29,18
2	A*01:01-B*08:01-DRB1*03:01	13,97	19,30	46,93
3	A*30:02-B*18:01-DRB1*03:01	17,74	25,75	46,93
4	A*03:01-B*07:02-DRB1*15:01	21,66	31,69	57,85
5	A*02:01-B*44:03-DRB1*07:01	30,72	31,69	57,85
6	A*02:01-B*07:02-DRB1*15:01	34,81	31,69	57,85
7	A*33:01-B*14:02-DRB1*01:02	36,69	35,69	57,85
8	A*02:01-B*51:01-DRB1*11:01	40,42	38,13	64,67
9	A*02:01-B*18:01-DRB1*03:01	43,87	38,13	64,67
10	A*01:01-B*57:01-DRB1*07:01	46,54	40,38	64,67

Figure 4

Estimated cumulative percentage of matched individuals in the Spanish population considering a “beneficial match” with a panel of 10 homozygous donors from Spanish cord blood Banks in high resolution allowing for a mismatch in HLA-B (A-DRB1), HLA-A (B-DRB1) or HLA-A and HLA-B (DRB1).