Open Access

OBM Transplantation

LIDSEN Publishing Inc.

Communication

Calculated Human Leucocyte Antigens Evolutionary Divergence (cHED)

Bruno A Lima *

Oficina de Bioestatistica, Ermesinde, Portugal; E-Mail: balima78@gmail.com

* **Correspondence:** Bruno A Lima; E-Mail: balima78@gmail.com

Academic Editors: Vikash C. Mishra, Muhammad Mujtaba and Abbas Ghazanfar

Special Issue: HLA in Transplantation: Typing, Matching and Outcomes

OBM Transplantation
2024, volume 8, issue 1
doi:10.21926/obm.transplant.2401208

Received: December 15, 2023
Accepted: March 05, 2024
Published: March 11, 2024

Abstract

Human Leucocyte Antigens (HLA) constitute a highly polymorphic set of genes pivotal to the immune response. The HLA heterozygous advantage hypothesis assumes that heterozygous individuals at the HLA level have a wider range of peptides for T cell recognition than homozygous individuals. Consequently, they possess an enhanced capacity to trigger a targeted immune reaction. The divergent allele advantage hypothesis is an extension of heterozygous advantage, considering the excessive sequence divergence between alleles of the same HLA locus. The HLA Evolutionary Divergence (HED) score has been proposed to quantify this divergence between alleles of the same HLA locus. Presented here is the calculated HED (cHED), an open-source web application designed for the computation of HED scores about the 5 classical HLA genes (HLA-A, -B, -C, -DRB1, and -DQB1) when delineated at two-field resolution.

Keywords

HLA; evolutionary divergence; heterozygote advantage; Grantham's distance

1. Introduction

The two classes of proteins with the highest degree of polymorphism in the human genome are Human Leucocyte Antigens (HLA) class I and II. HLA-A, -B, and -C molecules are found in nearly all cells except red blood cells, whereas class II HLA-DR and -DQ molecules are primarily found in immune cells [1]. The polymorphism of HLA is not only one of the main contributors to humanity's diversity, but with a genetically predetermined history, HLA has a crucial role in the immune response [2].

The HLA heterozygous advantage hypothesis postulated by Doherty and Zinkernagel [3] states that heterozygous genotypes for HLA *loci* are more likely to trigger an immune response in the presence of an infection. Also, according to this hypothesis, heterozygous HLA genotypes promote the presentation of more diverse tumor antigens to T cells [4].

Heterozygous individuals at the HLA level are assumed to possess a broader array of peptides for T cell recognition compared to homozygous individuals, thereby potentially enhancing their capacity to elicit a specific immune response. Since heterozygous individuals have shown greater resistance to pathogens, this may also explain the population's persistent increase of different HLA alleles.

This hypothesis of heterozygous advantage has been expanded to encompass the sequence level, leading to the emergence of the concept known as divergent allele advantage [5, 6]. The divergent allele advantage hypothesis [7] is a consequence of heterozygous advantage and results from excessive sequence divergence between alleles of the same HLA *locus* [8].

In other words, the divergent allele advantage hypothesis stipulates that diversity between HLA alleles of the same locus (i.e., with a greater number of amino acid differences between the peptide binding domains of two alleles of the same HLA locus) may increase the functional capacity of peptide antigen presentation and as such increase protection against pathogens and tumors through the ability to activate the protective immunity of T cells [9].

The HLA Evolutionary Divergence (HED) score has been proposed to quantify this divergence between alleles of the same HLA locus. An individual with a higher HED value between HLA alleles may allow the presentation of a more diverse immunopeptidome, as the different set of peptides from each allele within the highly polymorphic HLA collectively constitutes the immunopeptidome [10].

2. Objective

This paper aims to present a simple web application that allows the calculation of HED scores.

3. Methods

The calculated HED (cHED) is an application that allows the computation of HED scores for the 5 classical HLA genes (HLA-A, -B, -C, -DRB1, and -DQB1) when defined at two-field resolution for alleles annotated as ‘commons.’ This web application is available online as open-source software (https://txor.shinyapps.io/ched/).

The HED value is obtained from the Grantham distance [11] that calculates the difference of paired chains of polymorphic amino acids for different HLA alleles of the same *locus* and predicts the range of the immunopeptidome presented by the molecules encoded by that HLA locus. The Grantham distance [11] is a classical metric that allows quantifying the differences between

sequences of amino acids of proteins, considering their physiochemical properties: composition, polarity, and volume. The sum of the distances of the amino acids is normalized by the sequence length [12]. As per the definition, the value of HED is zero in instances of homozygosity.

The respective protein sequences of the peptide binding domain (exon 2 and 3 for HLA class I, exon 2 for HLA class II) were obtained from the international immunogenetics project's HLA database [13]; exon annotation was performed with the Ensembl database [14].

This application was built with shiny [15] in R [16] and using the package {histoc} [17] for HED computation. The entire code is accessible via its dedicated GitHub repository [18], and any reporting of bugs, requests for new features, or other forms of feedback are encouraged by submitting an issue through that platform.

4. Results

The cHED web application returns the HED values for each of the 5 main HLA's *loci*: HLA-A, -B, and -C of class I and HLA-DRB1 and -DQB1 of class II. It allows two forms of input: picking the alleles or uploading a data file.

Therefore, it is feasible to choose a pair of common alleles from the specific locus for which an outcome is required (Figure 1).

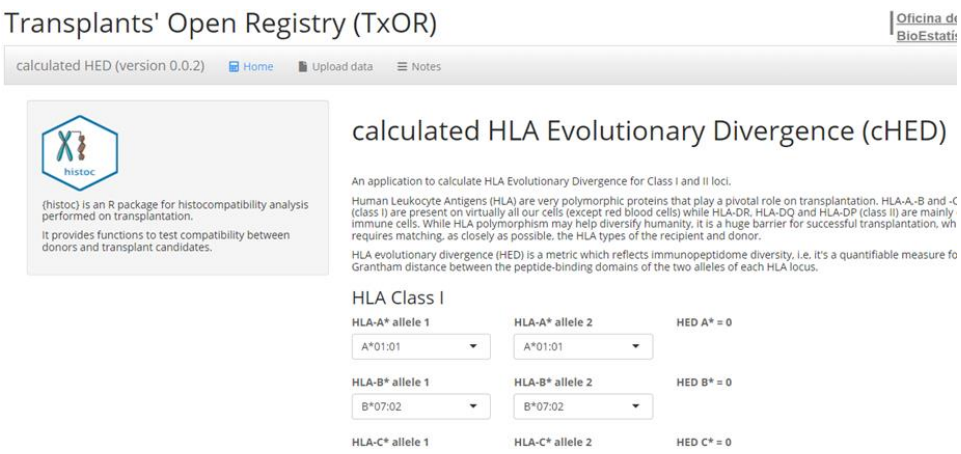

Figure 1 Calculated HED given as input a pair of alleles for each *loci*.

Alternatively, we can provide a text file containing HLA genotypes for multiple individuals as input (Figure 2). When using a text file as input, this should have 11 columns; the first corresponds to an identifier, and the remaining 10 to the 5 pairs of alleles for HLA genotypes. These columns should be named as: 'ID', 'A_1', 'A_2', 'B_1', 'B_2', 'C_1', 'C_2', 'DRB1_1', 'DRB1_2', 'DQB1_1' and 'DQB1_2'.

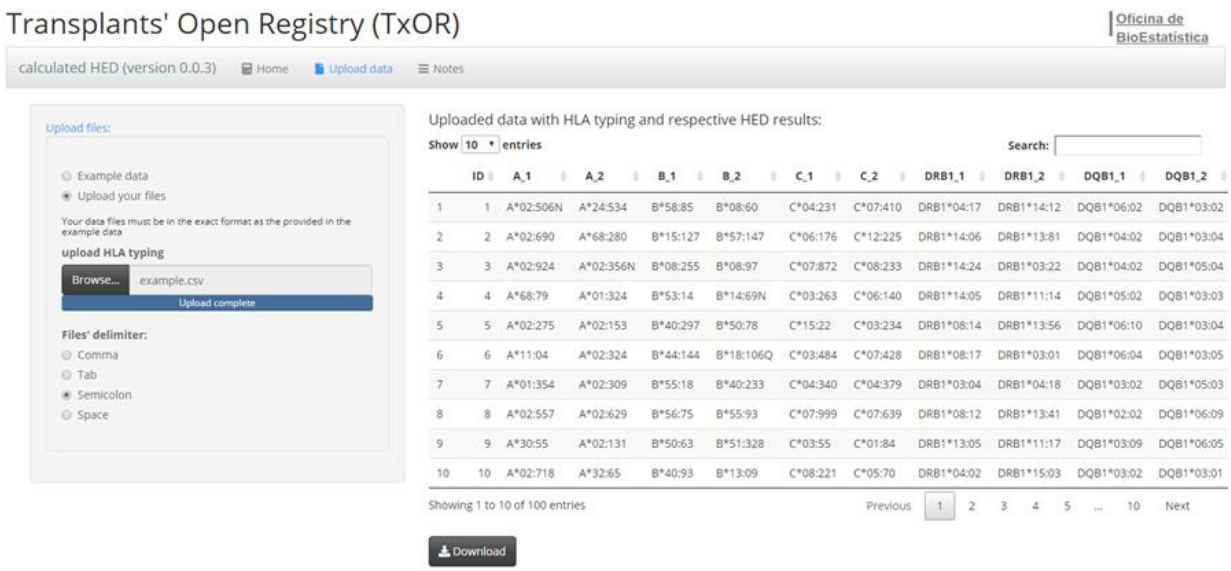

Figure 2 Inputing a file with HLA genotypes an.

For each *loci*, the valid alleles must be on two-field resolution. As shown in the example data, each allele must begin with its locus' name followed by an asterisk, the first two digits of the first field, a colon, and the digits of the second field. In the output, 5 additional columns are generated alongside the uploaded data, providing the HED results for each HLA *loci*. Moreover, the tabled results can be downloaded locally.

This web application is a straightforward calculator, offering the key advantage that individual data is neither stored nor processed, ensuring it remains inaccessible to anyone other than the user.

5. Discussion

A previously documented correlation between the genetic divergence of two alleles within the same *locus* and the combined number of peptides they are linked to has confirmed the divergent alleles advantage for the genes located within the classical HLA *loci* [6].

Likewise, higher HED values may permit the presentation of a greater variety of the immunopeptidome and, as such, facilitate the recognition of T cells and a more adapted immune response [2]. Employing the notion of allelic divergence permits a continuous and functionally oriented strategy for delineating HLA variability, in contrast to the binary classification intrinsic to basic pathogenic zygotes [8].

We can already find in the literature several studies that describe the association of HED values with some forms of cancer and their treatment. The effectiveness of T-cell-based immunotherapy approaches, for instance, is significantly influenced by HED values, which serve as a gauge of immunopeptidome variety [10]. Additionally, it was found that HED values (as a measure of sequence divergence between HLA alleles) are associated with checkpoint blockade immunotherapy response in patients treated for cancer [4]. The association between HED class I and outcome after kidney cancer treatment has also been described [12]. Similarly, an association between HED values and the effectiveness of immunotherapy for gastrointestinal cancer was established [2].

Furthermore, it has been described that HED values for both classes I and class II are associated with the survival of patients with Acute Myeloid Leukemia after hematopoietic stem cell

transplantation (HSCT) [19]. Likewise, higher HED values were associated with a better outcome after HSCT in pediatric patients [9].

In the case of infectious diseases, the functional basis of a protective association between HLA heterozygosity and HIV control has already been demonstrated [8]. While higher HED values for HLA-B could be associated with COVID-19 outcomes [20].

When HLA genotyping is conducted with a two-field resolution, the HED value can be calculated without incurring any extra expenses. Notably, this value can be computed on donors and recipients in transplantation. Suppose a link between donor HED levels and liver transplant outcome is described [21] in the future. In that case, there is potential to study the potential HED associations in kidney transplantation when the widespread utilization of HLA genotyping at a two-field resolution becomes a reality.

We can find other open-source software that allows us to calculate HED values, whether a Perl package [22] or an application that calculates HED for HLA class I loci [4, 6]. The application presented here (cHED) also allows us to calculate HED for class II loci.

In conclusion, the cHED web application is simple and freely available for anyone interested in calculating HED values. It is a part of the Transplant Open Registry (TxOR) initiative [23], which assembles several histocompatibility and transplantation-related open-source software applications and tools.

Hopefully, cHED will be a valuable instrument for researchers exploring the connection between allele advantage and the immune response to particular diseases or conditions.

Abbreviations

HED	HLA Evolutionary Divergence
HLA	Human Leucocyte Antigens
HSCT	Haematopoietic stem cell transplantation
TxOR	Transplant Open Registry

Author Contributions

Conception and design: Bruno A Lima. Drafting the article and revising it: Bruno A Lima. Providing intellectual content: Bruno A Lima. Approval of final version: Bruno A Lima.

Competing Interests

The author declares that he has no conflicts of interest to disclose.

References

1. Reits E, Neefjes J. HLA molecules in transplantation, autoimmunity and infection control: A comic book adventure. HLA. 2022; 100: 301-311.
2. Lu Z, Chen H, Jiao X, Wang Y, Wu L, Sun H, et al. Germline HLA-B evolutionary divergence influences the efficacy of immune checkpoint blockade therapy in gastrointestinal cancer. Genome Med. 2021; 13: 175.
3. Doherty PC, Zinkernagel RM. Enhanced immunological surveillance in mice heterozygous at the H-2 gene complex. Nature. 1975; 256: 50-52.

4. Chowell D, Krishna C, Pierini F, Makarov V, Rizvi NA, Kuo F, et al. Evolutionary divergence of HLA class I genotype impacts efficacy of cancer immunotherapy. Nat Med. 2019; 25: 1715-1720.
5. Wakeland EK, Boehme S, She JX, Lu CC, McIndoe RA, Cheng I, et al. Ancestral polymorphisms of MHC class II genes: Divergent allele advantage. Immunol Res. 1990; 9: 115-122.
6. Pierini F, Lenz TL. Divergent allele advantage at human MHC genes: Signatures of past and ongoing selection. Mol Biol Evol. 2018; 35: 2145-2158.
7. Potts WK, Wakeland EK. Evolution of diversity at the major histocompatibility complex. Trends Ecol Evol. 1990; 5: 181-187.
8. Arora J, Pierini F, McLaren PJ, Carrington M, Fellay J, Lenz TL. HLA heterozygote advantage against HIV-1 is driven by quantitative and qualitative differences in HLA allele-specific peptide presentation. Mol Biol Evol. 2020; 37: 639-650.
9. Merli P, Crivello P, Strocchio L, Pinto RM, Algeri M, Del Bufalo F, et al. Human leukocyte antigen evolutionary divergence influences outcomes of paediatric patients and young adults affected by malignant disorders given allogeneic haematopoietic stem cell transplantation from unrelated donors. Br J Haematol. 2023; 200: 622-632.
10. Roerden M, Nelde A, Heitmann JS, Klein R, Rammensee HG, Bethge WA, et al. HLA evolutionary divergence as a prognostic marker for AML patients undergoing allogeneic stem cell transplantation. Cancers. 2020; 12: 1835.
11. Grantham R. Amino acid difference formula to help explain protein evolution. Science. 1974; 185: 862-864.
12. Lee CH, DiNatale RG, Chowell D, Krishna C, Makarov V, Valero C, et al. High response rate and durability driven by HLA genetic diversity in patients with kidney cancer treated with Lenvatinib and pembrolizumab. Mol Cancer Res. 2021; 19: 1510-1521.
13. Robinson J, Barker DJ, Georgiou X, Cooper MA, Flicek P, Marsh SG. IPD-IMGT/HLA database. Nucleic Acids Res. 2020; 48: D948-D955.
14. Zerbino DR, Achuthan P, Akanni W, Amode MR, Barrell D, Bhai J, et al. Ensembl 2018. Nucleic Acids Res. 2018; 46: D754-D761.
15. Chang W, Cheng J, Allaire JJ, Sievert C, Schloerke B, Xie Y, et al. Shiny: Web application framework for R [Internet]. R Core Team; 2020. Available from: https://cran.r-project.org/package=shiny.
16. R Core Team. R: A language and environment for statistical computing [Internet]. R Core Team; 2019. Available from: https://www.r-project.org/.
17. Lima BA, Reis FP. Histoc: Histocompatibility analysis performed by kidney allocation systems. J Open Source Softw. 2022; 7: 4807.
18. Lima BA. txopen/cHED [Internet]. San Francisco, CA: GitHub; 2023. Available from: https://github.com/txopen/cHED.
19. Daull AM, Dubois V, Labussière-Wallet H, Venet F, Barraco F, Ducastelle-Lepretre S, et al. Class I/class II HLA evolutionary divergence ratio is an independent marker associated with disease-free and overall survival after allogeneic hematopoietic stem cell transplantation for acute myeloid leukemia. Front Immunol. 2022; 13: 841470.
20. Hernandez PV, Duffy B, Hock K, Farnsworth C, Schindler E, Liu C. HLA-B evolutionary divergence is associated with outcomes after SARS-CoV-2 infection. Hum Immunol. 2022; 83: 803-807.

21. Féray C, Taupin JL, Sebagh M, Allain V, Demir Z, Allard MA, et al. Donor HLA class 1 evolutionary divergence is a major predictor of liver allograft rejection: A retrospective cohort study. Ann Intern Med. 2021; 174: 1385-1394.
22. tobiaslenz. GranthamDist [Internet]. San Diego, CA: Source Forge Headquarters; 2020. Available from: https://sourceforge.net/projects/granthamdist/.
23. Lima B, Henriques T, Reis F, Alves H. Transplant open registry initiative. Port J Nephrol Hypertens. 2022; 36: 201-206.