



OPEN ACCESS

EDITED AND REVIEWED BY
Malek Kamoun,
University of Pennsylvania, United States

*CORRESPONDENCE

Chester A. Alper,
✉ chester.alper@childrens.harvard.edu
Roger L. Dawkins,
✉ rldawkins@cyo.edu.au
Jerzy K. Kulski,
✉ kulski@me.com
Charles E. Larsen,
✉ charles.larsen@childrens.harvard.edu
Sally S. Lloyd,
✉ slloyd@cyo.edu.au

SPECIALTY SECTION

This article was submitted to Evolutionary and Population Genetics, a section of the journal Frontiers in Genetics

RECEIVED 08 January 2023
ACCEPTED 17 January 2023
PUBLISHED 26 January 2023

CITATION

Alper CA, Dawkins RL, Kulski JK, Larsen CE and Lloyd SS (2023), Editorial: Population genomic architecture: Conserved polymorphic sequences (CPSs), not linkage disequilibrium.
Front. Genet. 14:1140350.
doi: 10.3389/fgene.2023.1140350

COPYRIGHT

© 2023 Alper, Dawkins, Kulski, Larsen and Lloyd. This is an open-access article distributed under the terms of the [Creative Commons Attribution License \(CC BY\)](https://creativecommons.org/licenses/by/4.0/). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

Editorial: Population genomic architecture: Conserved polymorphic sequences (CPSs), not linkage disequilibrium

Chester A. Alper^{1,2*}, Roger L. Dawkins^{3*}, Jerzy K. Kulski^{4*}, Charles E. Larsen^{1,2*} and Sally S. Lloyd^{3*}

¹Program in Cellular and Molecular Medicine, Boston Children's Hospital, Boston, MA, United States, ²Department of Pediatrics, Harvard Medical School, Boston, MA, United States, ³CY O'Connor ERADE Village Foundation, North Dandalup, WA, Australia, ⁴Department of Molecular Life Sciences, Division of Basic Medical Science and Molecular Medicine, Tokai University School of Medicine, Isehara, Japan

KEYWORDS

haplotype, genomic architecture, polymorphism, major histocompatibility complex, linkage disequilibrium, recombination, pedigree

Editorial on the Research Topic

[Population genomic architecture: Conserved polymorphic sequences \(CPSs\), not linkage disequilibrium](#)

The human extended major histocompatibility complex (MHC) region, ~7.7 megabases (Mb) in the middle of the short arm of chromosome 6, is one of the most gene-dense and polymorphic regions of the human genome (Horton et al., 2004; Shiina et al., 2009; Kulski et al., 2022). MHC haplotype characterization originated largely from genetic studies in the 1960s–1990s by Ceppellini et al. (1967) and the research groups of Chester Alper (Alper et al., 1983; Awdeh et al., 1983; Yunis et al., 2003) and Roger Dawkins (Degli-Esposti et al., 1992; Degli-Esposti et al., 1995; Dawkins et al., 1999).

In this Research Topic, we introduced the term “conserved polymorphic sequences” (CPSs) to encompass both Mb-length haplotypes, labeled currently either as conserved extended haplotypes or ancestral haplotypes, and their haplotypic fragments (“fixed” or “frozen” blocks). The relevance of CPSs includes: 1) an improved understanding of MHC genomic architecture—laying the foundation for other regions; 2) their use in testing genetic models for both meiotic recombination and trait association; and 3) providing insight into population genetics and evolution. Linkage disequilibrium (LD) and other expectation-maximization (EM) analyses of unrelated genotypes fail to detect most of the larger CPS variants but often detect their small (5–200 kb) blocks (Walsh et al., 2003; Alper et al., 2006). The technical explanation(s) for these differences remain unresolved. The Research Topic includes original work, a minireview, a commentary, and a mechanistic model for CPSs at high population frequency. The CPS mechanistic model may apply to any genomic region or vertebrate species (Dawkins and Lloyd).

Alper reviewed CPS conceptual development. The identification of high population frequency Mb-length MHC haplotypes resulted from observed shorter (140 kb) MHC-encoded complement gene haplotypes (“complotypes”; Alper et al., 1983) in strong LD with HLA class I and class II haplotypic variants. Later publications demonstrated both the intervening genetic “fixity” (essentially fixed DNA sequence) among each of those long-range population haplotypes and their extension from the telomeric end of the HLA class I region to the centromeric end of HLA class II (e.g.,

Dawkins et al., 1999; Smith et al., 2006; Lam et al., 2015). Alper concluded that the CPS concept offers new genomic insights (Larsen et al., 2014; Vadva et al., 2019) and genetic/epigenetic models for complex genetic traits (Alper et al., 2019).

Okano et al. described MHC CPSs in the domestic cat and highlighted the differences between human and cat MHC haplotype organization, consistent with prior work showing divergent vertebrate MHC haplotype structures (Kulski et al., 2002; Shiina et al., 2017). Using new genotyping methods and separate 3-generation cat pedigrees, the authors identified 14 unrelated long-range (within the 3 Mb domestic cat MHC reference sequence) FLA-I/FLA-DRB allelic variant founder haplotypes. Eight unique variants were identified, only three of which were singletons. One CPS variant consisted of three founder haplotypes, and the remaining four existed twice.

Specific CPS variant frequencies differ considerably among human ethnic populations, and genetic studies of a wider range of populations are needed (Sirugo et al., 2019). Two papers expanded CPS studies to understudied human populations. Tay et al., using segregation analysis to phase long-range HLA haplotypes in families with at least one type 1 diabetes (T1D) proband from the United Arab Emirates (UAE), reported the classical association of HLA-DR4 and HLA-DR3 haplotypes in UAE citizens affected by T1D. UAE T1D patient haplotypes with those HLA class II specificities were enriched in two Mb-length CPSs previously reported in northern India. This report preceded another pedigree-phased HLA haplotype analysis of 41 families (Alnaqbi et al., 2022) that identified three novel UAE population MHC CPSs.

Among twelve minority ethnic populations in China, Cun et al. reported on HLA class II haplotypes containing *HLA-DRB1* and five polymorphic Alu insertions (POALINs; a structural variant type of transposable element (TE) genetic marker) covering 850 kb. These HLA class II TE markers of diversity and historic crossing over showed strong association among these populations with known language family, migration, and sociality characteristics. The results were compared with those obtained from dominant populations from Japan and Australia (Kulski et al., 2010) and the Yunnan province of China (Shi et al., 2014).

Mb-length CPSs constitute nearly half of European Caucasian MHC haplotypes (Szilágyi et al., 2010). However, to identify the diversity of all haplotypes, relatively dense genetic markers are needed to detect genomic alterations (e.g., meiotic recombination, insertion-deletion polymorphisms). As part of an investigation into putative associations and functions of TEs in generating human (Kulski et al., 1997; Kulski et al., 1999; Kulski et al., 2000) and non-human primate (Kulski et al., 2004) MHC haplotypes, two reports by Kulski, Suzuki and Shiina used a set of 95 extended MHC haplotype sequences from a publicly available database (Norman et al., 2017) to study meiotic recombination resulting in haplotype segmental shuffling or crossovers in the MHC class I and II regions. Both studies demonstrated that 1) SNP-density crossovers are associated with putative ancestral recombination sites widely spread across the odor type receptor and MHC class I and class II regions; 2) MHC homozygous cell line genomic sequences are useful for analyzing haplotype blocks, ancestral haplotypic landscapes and markers, CPSs, and SNP-density crossover junctions without the need for probabilistic statistical imputation; and 3) TEs are useful genetic markers of recent recombination events and for elucidating population phylogenetics and genetic interrelationships (Kulski et al., 2011; 2019; Wang et al., 2017; Abeid et al., 2019). There was substantial haplotype shuffling between different polymorphic

blocks suggesting the presence of numerous putative ancestral recombination sites between specific class II genes. TEs, in addition to being a useful class of haplotypic markers, may have had a critical impact on CPS diversity between individuals and population groups.

Finally, Radman offered a hypothesis, which was further contextualized in a commentary by Dawkins and Lloyd, for the existence of long-range CPSs: as a consequence of resistance to meiotic recombination both in regions with highly polymorphic sequences and in regions dense with gene families requiring maintenance of genomic integrity. Significantly, the MHC class III region (between HLA class I and class II) is both the most gene-dense sub-region of the extended MHC (Horton et al., 2004) and, along with a portion of the telomeric end of class II, the most difficult MHC region to align sequencing assemblies due to high levels of indel polymorphism (Horton et al., 2008; Chin et al., 2020). Radman's mechanistic model and this Research Topic are pleas for additional work to understand better the enigmatic relationships between CPS haplotypes, recombination mechanisms, and genetic fixity with genomic structural integrity and diversity within and between different populations and species. A database (http://cyo.edu.au/CPS_Database) is available for researchers wishing to submit CPSs of 400 nucleotides or longer shown to exist as at least three variants within a specific genome at a well-defined genomic location.

Author contributions

All authors listed have made a substantial, direct, and intellectual contribution to the work and approved it for publication.

Funding

CAA and CEL were funded by institutional funds from the Program in Cellular and Molecular Medicine, Boston Children's Hospital.

Acknowledgments

The guest editors thank all the authors, reviewers, Associate Editors and Frontiers support staff for their valuable contributions to this Research Topic.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

References

- Abeid, S. N., Motrane, M., Farhane, H., and Harich, N. (2019). Alu elements within the human major histocompatibility class I region in the Comoros islands: Genetic variation and population relationships. *Ann. Hum. Biol.* 46 (2), 169–174. doi:10.1080/03014460.2019.1620854
- Alnaqbi, H., Tay, G. K., Chehadeh, S. E. H., and Alsafar, H. (2022). Characterizing the diversity of MHC conserved extended haplotypes using families from the United Arab Emirates. *Sci. Rep.* 12, e7165. doi:10.1038/s41598-022-11256-y
- Alper, C. A., Raum, D., Karp, S., Awdeh, Z. L., and Yunis, E. J. (1983). Serum complement 'supergenes' of the major histocompatibility complex in man (complotypes). *Vox Sang.* 45, 62–67. doi:10.1111/j.1423-0410.1983.tb04124.x
- Alper, C. A., Larsen, C. E., Dubey, D. P., Awdeh, Z. L., Fici, D. A., and Yunis, E. J. (2006). The haplotype structure of the human major histocompatibility complex. *Hum. Immunol.* 67, 73–84. doi:10.1016/j.humimm.2005.11.006
- Alper, C. A., Larsen, C. E., Trautwein, M. R., and Alford, D. R. (2019). A stochastic epigenetic Mendelian oligogenic disease model for type 1 diabetes. *J. Autoimmun.* 96, 123–133. doi:10.1016/j.jaut.2018.09.006
- Awdeh, Z. L., Raum, D., Yunis, E. J., and Alper, C. A. (1983). Extended HLA/complement allele haplotypes: Evidence for T/t-like complex in man. *Proc. Natl. Acad. Sci. U. S. A.* 80, 259–263. doi:10.1073/pnas.80.1.259
- Cepellini, R., Curtioni, E. S., Mattiuz, P. L., Miggiano, V., Scudeller, G., and Serra, A. (1967). "Genetics of leukocyte antigens: A family study of segregation and linkage," in *Histocompatibility testing 1967*. Editors E. S. Curtioni, P. L. Mattiuz, and R. M. Tosi (Copenhagen: Munksgaard), 149–187.
- Chin, C.-S., Wagner, J., Zeng, Q., Garrison, E., Garg, S., Fungtammasan, A., et al. (2020). A diploid assembly-based benchmark for variants in the major histocompatibility complex. *Nat. Commun.* 11, 4794. doi:10.1038/s41467-020-18564-9
- Dawkins, R., Leelayuwat, C., Gaudieri, S., Tay, G., Hui, J., Cattley, S., et al. (1999). Genomics of the major histocompatibility complex: Haplotypes, duplication, retroviruses and disease. *Immunol. Rev.* 167, 275–304. doi:10.1111/j.1600-065x.1999.tb01399.x
- Degli-Esposti, M. A., Leaver, A. L., Christiansen, F. T., Witt, C. S., Abraham, L. J., and Dawkins, R. L. (1992). Ancestral haplotypes: Conserved population MHC haplotypes. *Hum. Immunol.* 34, 242–252. doi:10.1016/0198-8859(92)90023-g
- Degli-Esposti, M. A., Leelayuwat, C., Daly, L. N., Carcassi, C., Contu, L., Versluis, L. F., et al. (1995). Updated characterization of ancestral haplotypes using the Fourth Asia-Oceania Histocompatibility Workshop panel. *Hum. Immunol.* 44, 12–18. doi:10.1016/0198-8859(95)00044-5
- Horton, R., Gibson, R., Coggill, P., Miretti, M., Allcock, R. J., Almeida, J., et al. (2008). Variation analysis and gene annotation of eight MHC haplotypes: The MHC Haplotype Project. *Immunogenetics* 60, 1–18. doi:10.1007/s00251-007-0262-2
- Horton, R., Wilming, L., Rand, V., Lovering, R. C., Bruford, E. A., Khodiyar, V. K., et al. (2004). Gene map of the extended human MHC. *Nat. Rev. Genet.* 5, 889–899. doi:10.1038/nrg1489
- Kulski, J. K., Gaudieri, S., Bellgard, M., Balmer, L., Giles, K., Inoko, H., et al. (1997). The evolution of MHC diversity by segmental duplication and transposition of retroelements. *J. Mol. Evol.* 45, 599–609. doi:10.1007/PL00006264
- Kulski, J. K., Gaudieri, S., Martin, A., and Dawkins, R. L. (1999). Coevolution of PERB11 (MIC) and HLA class I genes with HERV-16 and retroelements by extended genomic duplication. *J. Mol. Evol.* 49, 84–97. doi:10.1007/PL00006537
- Kulski, J. K., Gaudieri, S., and Dawkins, R. L. (2000). "Transposable elements and the metamerismic evolution of the HLA class I region," in *Major histocompatibility complex*. Editor M. Kasahara (Tokyo: Springer Japan), 158–177. doi:10.1007/978-4-431-65868-9_11
- Kulski, J. K., Shiina, T., Anzai, T., Kohara, S., and Inoko, H. (2002). Comparative genomic analysis of the MHC: The evolution of class I duplication blocks, diversity and complexity from shark to man. *Immunol. Rev.* 190, 95–122. doi:10.1034/j.1600-065X.2002.19008.x
- Kulski, J. K., Anzai, T., Shiina, T., and Hidetoshi, I. (2004). Rhesus macaque class I duplicon structures, organization, and evolution within the alpha block of the major histocompatibility complex. *Mol. Biol. Evol.* 21, 2079–2091. doi:10.1093/molbev/msh216
- Kulski, J. K., Shigenari, A., Shiina, T., and Inoko, H. (2010). Polymorphic major histocompatibility complex class II *Alu* insertions and five loci and their association with *HLA-DRB1* and *-DQB1* in Japanese and Caucasians. *Tissue Antigens* 76, 35–47. doi:10.1111/j.1399-0039.2010.01465.x
- Kulski, J. K., Shigenari, A., and Inoko, H. (2011). Genetic variation and hitchhiking between structurally polymorphic *Alu* insertions and HLA-A, -B, and -C alleles and other retroelements within the MHC class I region. *Tissue Antigens* 78, 359–377. doi:10.1111/j.1399-0039.2011.01776.x
- Kulski, J. K., Mawart, A., Marie, K., Tay, G. K., and Alsafar, H. S. (2019). MHC class I polymorphic *Alu* insertion (POALIN) allele and haplotype frequencies in the Arabs of the United Arab Emirates and other world populations. *Int. J. Immunogenet.* 46, 247–262. doi:10.1111/iji.12426
- Kulski, J. K., Suzuki, S., and Shiina, T. (2022). Human leukocyte antigen super-locus: Nexus of genomic supergenes, SNPs, indels, transcripts, and haplotypes. *Hum. Genome Var.* 9, e49. doi:10.1038/s41439-022-00226-5
- Lam, T. H., Tay, M. Z., Wang, B., Xiao, Z., and Ren, E. C. (2015). Intrahaplotypic variants differentiate complex linkage disequilibrium within human MHC haplotypes. *Sci. Rep.* 5, 16972. doi:10.1038/srep16972
- Larsen, C. E., Alford, D. R., Trautwein, M. R., Jalloh, Y. K., Tarnacki, J. L., Kunnenkeri, S. K., et al. (2014). Dominant sequences of human major histocompatibility complex conserved extended haplotypes from *HLA-DQA2* to *DAXX*. *PLoS Genet.* 10, e1004637. doi:10.1371/journal.pgen.1004637
- Norman, P. J., Norberg, S. J., Guethlein, L. A., Nemat-Gorgani, N., Royce, T., Wroblewski, E. E., et al. (2017). Sequences of 95 human MHC haplotypes reveal extreme coding variation in genes other than highly polymorphic *HLA class I* and *II*. *Genome Res.* 27, 813–823. doi:10.1101/gr.213538.116
- Shi, L., Kulski, J. K., Zhang, H., Dong, Z., Cao, D., Zhou, J., et al. (2014). Association and differentiation of MHC class I and II polymorphic *Alu* insertions and HLA-A, -B, -C and -DRB1 alleles in the Chinese Han population. *Mol. Genet. Genomics* 289, 93–101. doi:10.1007/s00438-013-0792-2
- Shiina, T., Hosomichi, K., Inoko, H., and Kulski, J. K. (2009). The HLA genomic loci map: Expression, interaction, diversity and disease. *J. Hum. Genet.* 54, 15–39. doi:10.1038/jhg.2008.5
- Shiina, T., Blancher, A., Inoko, H., and Kulski, J. K. (2017). Comparative genomics of the human, macaque and mouse major histocompatibility complex. *Immunology* 150, 127–138. doi:10.1111/imm.12624
- Sirugo, G., Williams, S. M., and Tishkoff, S. A. (2019). The missing diversity in human genetic studies. *Cell* 177, 26–31. doi:10.1016/j.cell.2019.02.048
- Smith, W. P., Vu, Q., Li, S. S., Hansen, J. A., Zhao, L. P., and Geraghty, D. E. (2006). Toward understanding MHC disease associations: Partial resequencing of 46 distinct HLA haplotypes. *Genomics* 87, 561–571. doi:10.1016/j.ygeno.2005.11.020
- Szilágyi, Á., Bánlaki, Z., Pozsonyi, É., Yunis, E. J., Awdeh, Z. L., Hossó, A., et al. (2010). Frequent occurrence of conserved extended haplotypes (CEHs) in two Caucasian populations. *Mol. Immunol.* 47, 1899–1904. doi:10.1016/j.molimm.2010.03.013
- Vadva, Z., Larsen, C. E., Propp, B. E., Trautwein, M. R., Alford, D. R., and Alper, C. A. (2019). A new pedigree-based SNP haplotype method for genomic polymorphism and genetic studies. *Cells* 8 (8), 835. doi:10.3390/cells8080835
- Walsh, E. C., Mather, K. A., Schaffner, S. F., Farwell, L., Daly, M. J., Patterson, N., et al. (2003). An integrated haplotype map of the human major histocompatibility complex. *Am. J. Hum. Genet.* 73, 580–590. doi:10.1086/378101
- Wang, L., Norris, E. T., and Jordan, I. K. (2017). Human retrotransposon insertion polymorphisms are associated with health and disease via gene regulatory phenotypes. *Front. Microbiol.* 8, 1418. doi:10.3389/fmicb.2017.01418
- Yunis, E. J., Larsen, C. E., Fernandez-Viña, M., Awdeh, Z. L., Romero, T., Hansen, J. A., et al. (2003). Inheritable variable sizes of DNA stretches in the human MHC: Conserved extended haplotypes and their fragments or blocks. *Tissue Antigens* 62, 1–20. doi:10.1034/j.1399-0039.2003.00098.x