Title
P826: Evaluating HLA allele-disease associations in ClinGen: Development of a new curation tool

Authors
Keseler, Ingrid
Bowen, Mónica
Whirl-Carrillo, Michelle
et al.

Publication Date
2024

DOI
10.1016/j.gimo.2024.101737

Evaluating HLA allele-disease associations in ClinGen: Development of a new curation tool

P826

Ingrid M. Keseler[1], B. Mónica Bowen[1], Michelle Whirl-Carillo[1], Obed A. Garcia[2], Gloria Cheung[1], Clarissa J. Klein[1], Mark Mandell[1], Rachel Shapira[1], Matt W. Wright[1], S. Louis Bridges, Jr.[3], Marcelo Fernandez-Viña[4], Steven J. Mack[5], Teri E. Klein[1,6]

[1]Department of Biomedical Data Science, Stanford University, [2]Department of Ecology & Evolutionary Biology, University of California, Santa Cruz, [3]Department of Medicine, Hospital for Special Surgery, [4]Department of Pathology, Stanford University, [5]Department of Pediatrics, University of California, San Francisco, [6]Department of Medicine, Stanford University

Background

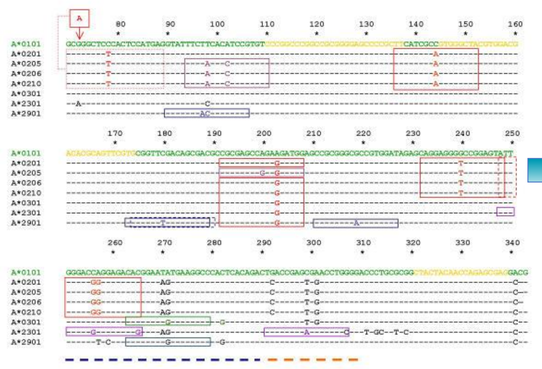
The Human Leukocyte Antigen (HLA) region encodes key components of the adaptive immune system that presents antigens to immune-monitoring cells. With 38,000 HLA alleles known, it is the most complex and polymorphic region in the human genome. The region's high structural and genetic diversity is the result of natural selection for the optimized capacity to detect pathogens and has resulted in extensive genetic variation across the human population. This selection of the HLA system for strong immunity has also contributed to human autoimmune disorders.

Within the ClinGen Consortium, curators extract essential data from individual publications for presentation to and evaluation by expert panels. The ClinGen Gene Curation Interface (GCI) enables the curation, evaluation and publication of monogenic gene-disease validity assertions. We are now developing a new curation tool, the HLA Curation Interface (HCI), that enables curation of associations between HLA alleles or haplotypes and complex autoimmune and immune-mediated disorders such as rheumatoid arthritis, ankylosing spondylitis and scleroderma.

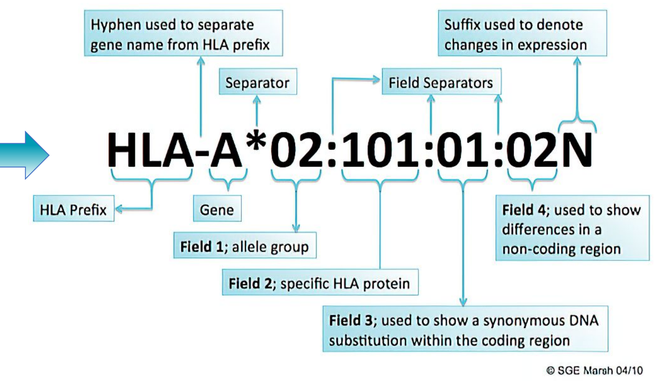
HLA Nomenclature

IPD-IMGT/HLA Database (https://www.ebi.ac.uk/ipd/imgt/hla/)

The IPD-IMGT/HLA Project Official Repository is managed by the WHO Nomenclature Committee for Factors of the HLA System. More than 100 releases of HLA allele name and sequence data of more than 38,000 HLA alleles have been produced since 1998.

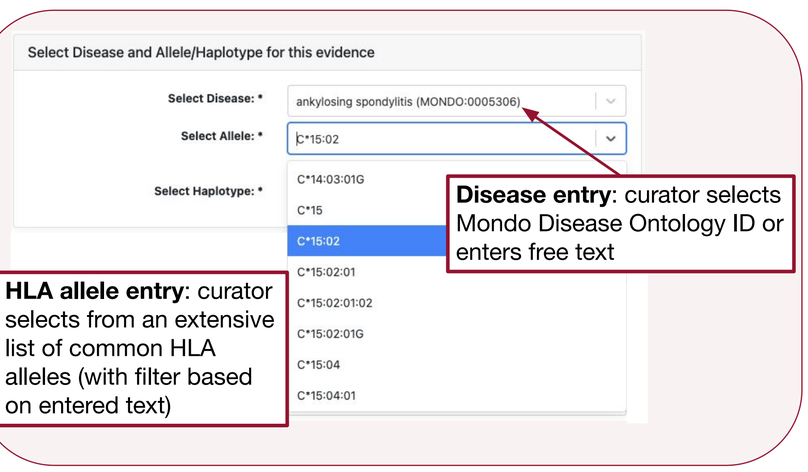


HLA nucleotide sequences are highly polymorphic and often reflect ancient polymorphisms, which makes SNP-based HLA genotyping difficult.



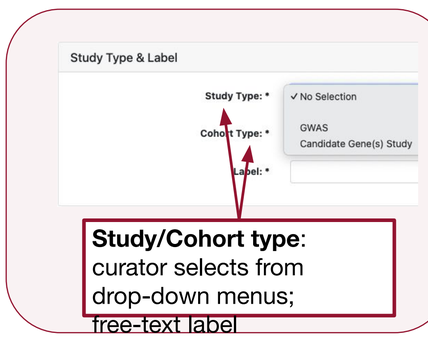
HLA Curation Interface (currently in testing phase)

Elements of the HLA Curation Interface

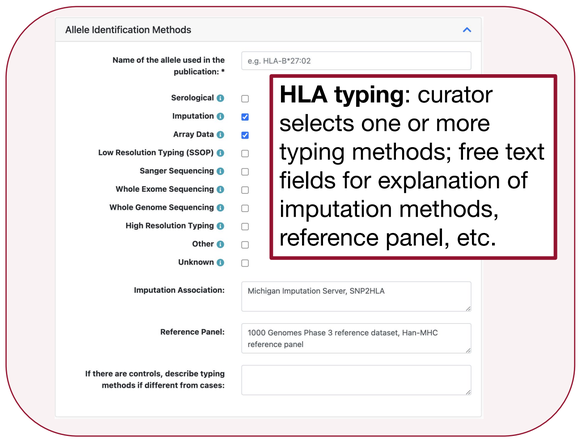


Disease entry: curator selects Mondo Disease Ontology ID or enters free text

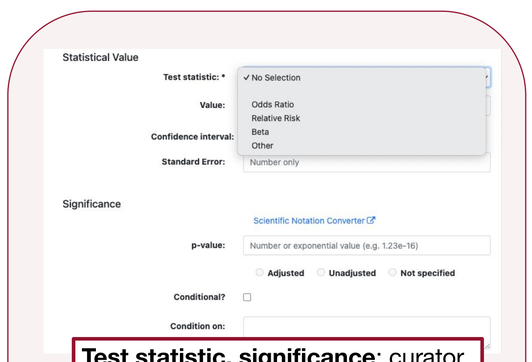
HLA allele entry: curator selects from an extensive list of common HLA alleles (with filter based on entered text)



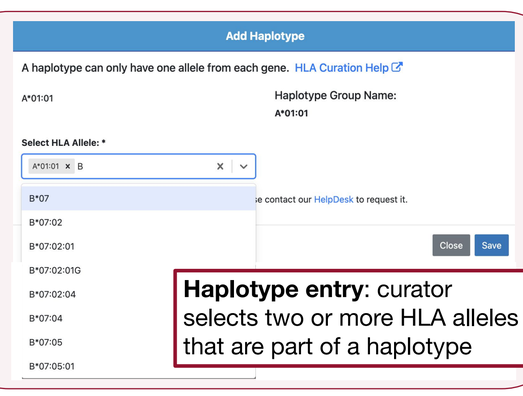
Study/Cohort type: curator selects from drop-down menus; free-text label



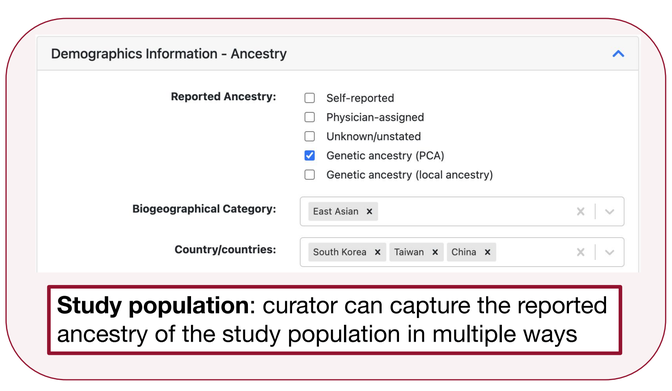
HLA typing: curator selects one or more typing methods; free text fields for explanation of imputation methods, reference panel, etc.



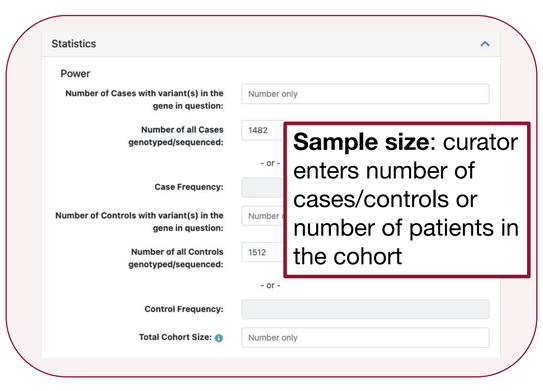
Test statistic, significance: curator selects test statistic and enters p-value with information on multiple testing correction, conditionalization on presence of other alleles



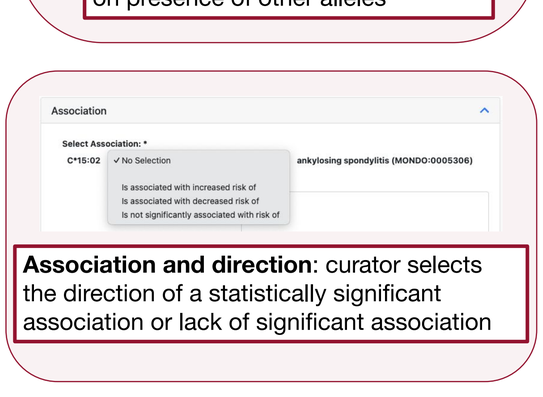
Haplotype entry: curator selects two or more HLA alleles that are part of a haplotype



Study population: curator can capture the reported ancestry of the study population in multiple ways

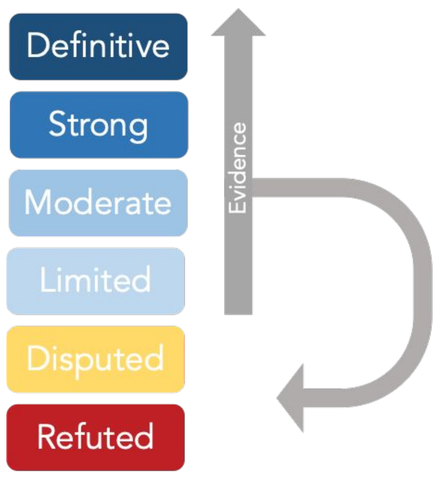


Sample size: curator enters number of cases/controls or number of patients in the cohort



Association and direction: curator selects the direction of a statistically significant association or lack of significant association

Levels of Evidence for HLA Allele-Disease Associations



Individual pieces of evidence will receive a point score based on quality and statistical significance. For example, a study that used high resolution allele typing methods, had a large cohort size and reached high statistical significance will receive a high point score.

Scores from multiple studies on the same HLA allele-disease association will be combined into an overall score, which will determine the classification.

During evidence review by a ClinGen Expert Panel, scores and classifications can be modified. If contradictory evidence is found, the HLA allele-disease association can also be moved to the Disputed or Refuted category.

Scan this QR code for more information on the **ClinGen HLA Working Group**

Example Output

HLA-B*27 has a **high level of evidence for association** for **increased susceptibility** to **Ankylosing spondylitis**. This association demonstrates an increased risk across **multiple cohorts** and in **multiple populations** worldwide.

Future Directions

We are in the process of conducting user testing of the HCI to ensure that the tool captures all necessary information to curate publications on HLA allele-disease associations. We are also refining the scoring system for data categories including HLA typing, sample size, significance, etc., to accurately reflect the evidence level, and determining the appropriate score thresholds for the overall level of evidence for an HLA-disease association.

Acknowledgements

We would like to thank our funding: NIH ClinGen Grant #U24HG009649