I-A

HIV Epitope A-list: The 2021 Update

Sandra Silva-Arrieta[a], Tuixent Escriba[a], Anuska Llano[a], Katherine A. Belobrajdic[b], Elizabeth-Sharon David-Fung[b], Christian Brander[a,c,d]

I-A-1 Introduction

Over the past two years, the definition of well-characterized T-cell epitopes in diverse viral infections has gained renewed attention and utility, partially in response to numerous approaches taken for the development of SARS-CoV-2 epitope-based vaccine candidates. Due to the successful development of both preventive and therapeutic epitope-based vaccine candidates, strategies for vaccination now include the use of T cell epitopes that appear to be (a) under especially potent immune-selection pressure, (b) highly conserved and/or (c) restricted to particular HLA(s). In addition, different approaches are being pursued that are based on common length of defined epitopes, such as potential T-cell epitope (PTE)-informed design [Fischer *et al.*, 2008; Theiler *et al.*, 2016]. With the advent of novel platforms for deep analysis of T cell receptor repertoires, the need to know the precise nature of targeted epitopes has grown again [Pavlović *et al.*, 2022]. This is critical for therapeutic HIV T cell vaccines, where it may be important to discriminate pre-existing from newly induced T cell specificities and clonotypes so that the clinical efficacy of different T cell subsets in controlling *in vivo* viral replication can be understood. With these needs in mind and an awareness of gaps in past HIV T-cell epitope listings, we have further refined the "HIV Best Defined CD8+/CTL Epitope List" (commonly called the 'A-list') by revising high-resolution HLA restriction and epitope mapping data.

For this year's update then, we spent considerable effort in revisiting not only epitope mapping data but also defining high-resolution typing of the restricting HLA allele(s) for cases where this was not previously available. The recent upgrade of Los Alamos HIV Immunology database's usage of all HLA notation to current HLA convention as per the HLA Informatics Group `http://hla.alleles.org/nomenclature/naming.html` has made this effort possible.

We have requested high-resolution HLA typing data from many laboratories for epitopes that had previously only been defined at low, 2-digit resolution. This is important as 4-digit subtypes of the same 2-digit allele can differ profoundly in the presented epitope repertoire and consequently, have a dramatically different outcome on HIV disease progression. A clear example of this is the HLA-B58 allele, where HLA-B*58:01 and B*58:02 have been shown to restrict different immunodominant epitopes and show opposite associations with viral loads in individuals expressing one or the other allele [Ngumbela *et al.*, 2008]. Similar observations have been made for other closely related alleles, for instance in the HLA-B7 supertype family [Leslie *et al.*, 2006] and even for subtypes of alleles generally associated with superior control of viral infection [Kloverpris *et al.*, 2012; Yu *et al.*, 2007]. By reviewing serial papers from different labs, we were able to resolve fine restriction for other epitopes, with the exception of a handful of candidates where no information or patient samples were available.

For this reason, we recommend that individual research groups maintain consistent patient identifiers (patient ID or patient code) in their publications and/or maintain a publicly accessible log of matching patient IDs, following existing data protection regulations. As a result, investigators could take advantage of the accumulated data on specific patients published across different publications. The Los Alamos HIV Immunology Database has taken a step in this direction with its Patient Database, where individual patient data can be accessed and which in many cases match attributes such as HLA, risk factors, cohort inclusion and other parameters across multiple records and publications and also across the LANL HIV Immunology as well as Sequence Databases. (`https://www.hiv.lanl.gov/mojo/immunology/patient/form.html`)

The current HIV epitope A-list has grown considerably as a result of our efforts to determine their high-resolution restricting HLA allele(s) and due to some epitopes being recognized in the context of several HIV subtypes. We have decided to include related alleles as individual presenting

[a]IrsiCaixa - AIDS Research Institute, Badalona, Spain
[b]HIV Immunology Database, Theoretical Biology and Biophysics, Los Alamos National Laboratory, Los Alamos, N.M., U.S.A.
[c]Universitat de Vic-Universitat Central de Catalunya (UVic-UCC), Vic, Spain
[d] Institució Catalana de Recerca i Estudis Avançats (ICREA), Barcelona, Spain

In *HIV Molecular Immunology 2021.* Elizabeth-Sharon David-Fung, Bette T. M. Korber, Christian Brander, Dan Barouch, Rob de Boer, Barton F. Haynes, Richard Koup, John P. Moore, Bruce D. Walker, and David I. Watkins, editors. 2021. Publisher: Los Alamos National Laboratory, Theoretical Biology and Biophysics, Los Alamos, New Mexico. p. 3–4.

molecules since, as mentioned above, different subtypes can have different outcomes on virus control and potentially have a marked impact on presented epitope repertoires.

The HIV A-list is now a collection of well-defined epitopes, where the shortest epitope length-variant eliciting the strongest immune response is considered "optimal". Even in cases where length variants show comparable activity, preference has traditionally been given for the shorter version. Supported by growing data from epitope-elution studies [Yaciuk *et al.*, 2014], multiple length variants of the same "optimal" epitope may be presented by the same allele and it will be important to understand the impact of such mixed epitope lengths on the "optimal epitope-specific" T-cell receptor repertoire, the cross-reactivity of such T cells and, eventually, impacts on immune escape and control of viral replication.

Acknowledgments

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreements No. 681137 and 731626. The HIV Databases at Los Alamos National Laboratory are funded by NIH contract NIH: HIV Data Base and Analysis Unit (AAI22007).

We wish to thank the laboratories of S. Migueles and M. Connors (NIH) and P. Goulder (Oxford) that provided especially valuable data, and even performed high-resolution typing to define A-list restricting HLA elements.

This year's update is dedicated to the memory of Dr. Karina Yusim, who has been a constant support for our work and the driving force behind the epitope listings at the HIV Immunology Database. Her efforts have made this database one of the most powerful resources in the field of HIV immunology and beyond.

I-A-2 Table of optimal HIV-1 CTL epitopes

Updated Link to the HIV Database Best-Defined (A-list) Table of Cytotoxic T-cell epitopes, `https://www.hiv.lanl.gov/content/immunology/tables/optimal_ctl_summary.html`.

I-A-3 Map of optimal HIV-1 CTL epitopes

Updated, Complete and Interactive Maps, `https://www.hiv.lanl.gov/content/immunology/maps/maps.html`.

I-A-4 References

[Fischer *et al.*, 2008] W. Fischer, H. X. Liao, B. F. Haynes, N. L. Letvin, & B. Korber. Coping with viral diversity in HIV vaccine design: a response to Nickle *et Al. PLoS Comput Biol* **4**(1), Jan. 2008, e15, author reply e25. DOI: 10.1371/journal.pcbi.0040015. On p. 3.

[Kloverpris *et al.*, 2012] H. N. Kloverpris, A. Stryhn, M. Harndahl, *et al.* HLA-B*57 micropolymorphism shapes HLA allele-specific epitope immunogenicity, selection pressure, and HIV immune control. *J Virol* **86**(2), Jan. 2012, pp. 919–929. DOI: 10.1128/JVI.06150-11. On p. 3.

[Leslie *et al.*, 2006] A. Leslie, D. A. Price, P. Mkhize, *et al.* Differential selection pressure exerted on HIV by CTL targeting identical epitopes but restricted by distinct HLA alleles from the same HLA supertype. *J Immunol* **177**(7), Oct. 1, 2006, pp. 4699–4708. On p. 3.

[Ngumbela *et al.*, 2008] K. C. Ngumbela, C. L. Day, Z. Mncube, *et al.* Targeting of a CD8 T cell Env epitope presented by HLA-B*5802 is associated with markers of HIV disease progression and lack of selection pressure. *AIDS Res Hum Retroviruses* **24**(1), Jan. 2008, pp. 72–82. DOI: 10.1089/aid.2007.0124. On p. 3.

[Pavlović *et al.*, 2022] M. Pavlović, G. S. A. Hajj, J. Pensar, M. Wood, L. M. Sollid, V. Greiff, & G. K. Sandve. *Improving generalization of machine learning-identified biomarkers with causal modeling: an investigation into immune receptor diagnostics*. arXiv. Apr. 20, 2022. DOI: 10.48550/ARXIV.2204.09291. On p. 3.

[Theiler *et al.*, 2016] J. Theiler, H. Yoon, K. Yusim, L. J. Picker, K. Fruh, & B. Korber. Epigraph: a vaccine design tool applied to an HIV therapeutic vaccine and a pan-filovirus vaccine. *Sci Rep* **6**, Oct. 5, 2016, p. 33987. DOI: 10.1038/srep33987. On p. 3.

[Yaciuk *et al.*, 2014] J. C. Yaciuk, M. Skaley, W. Bardet, *et al.* Direct interrogation of viral peptides presented by the class I HLA of HIV-infected T cells. *J Virol* **88**(22), Nov. 2014, pp. 12992–3004. DOI: 10.1128/JVI.01914-14. On p. 4.

[Yu *et al.*, 2007] X. G. Yu, M. Lichterfeld, S. Chetty, *et al.* Mutually exclusive T-cell receptor induction and differential susceptibility to human immunodeficiency virus type 1 mutational escape associated with a two-amino-acid difference between HLA class I subtypes. *J Virol* **81**(4), Feb. 2007, pp. 1619–1631. DOI: 10.1128/JVI.01580-06. On p. 3.