Topological perspective on HIV escape:

A network analysis explains variability in immune control of HIV infection

Andrew J. McMichael[1], **Mary Carrington**[2]

[1]Nuffield Department of Medicine, University of Oxford, NDM Research Building, Old Road Campus, Oxford OX3 7FZ, UK.

[2]Basic Science Program, Frederick National Laboratory for Cancer Research, Frederick, MD 21702, USA.

Untreated HIV-1 infection has a variable outcome. Some patients progress to AIDS within a few months, whereas others survive for years; a very few even resist progression for decades. The rate of AIDS progression is affected by variation in the genes that encode a patient's human leukocyte antigen (HLA) proteins (1–3). This genetic association between HLA and AIDS progression has been attributed to differences in T cell recognition of HLA-bound HIV-derived peptide epitopes. The virus can variably escape immune detection by generating mutations within the peptide epitopes (4). On page 480 of this issue, Gaiha *et al.* (5) explain variation in HIV immune escape by examining each epitope in the context of HIV protein structure, using network theory analysis. The authors found that epitopes containing amino acids with multiple interactions within the three-dimensional structure of an HIV protein are constrained from escape and tend to bind to protective HLA types.

Class I HLA molecules expressed by infected cells display and present HIV-derived peptides to CD8[+] T cells. This enables CD8[+] T cells to recognize these HIV peptide epitopes and destroy infected cells. Because some epitopes are constrained and cannot mutate without affecting viral fitness, HLA molecules that bind to these particular peptides are more effective in controlling infection. Some regions of HIV proteins may be conserved because of such fitness costs. Thus, there is interest in designing vaccines based on conserved regions of the virus proteome. However, some protein regions show conservation simply because they are not targeted by the immune response and therefore are not subjected to selective pressure. How and why certain peptides dominate T cell responses is not fully understood, but clearly, some regions of HIV proteins are more immunogenic than others. The most protective HLA types (such as B*57, B*27, and B*81) tend to bind to more conserved peptide epitopes, which may incur fitness costs if they mutate (6, 7). However, by no means are all infected people with these "lucky" HLA types protected, and some who attain prolonged immune control of their infection do not have advantageous HLA types.

On the basis of crystal structures of various HIV-1 proteins, Gaiha *et al.* analyzed each amino acid within peptide epitopes, typically 8 to 11 amino acids long, for its contacts with other amino acids in the protein. The contacts included hydrogen bonds, van der Waals

andrew.mcmichael@ndm.ox.ac.uk; carringm@mail.nih.gov.

contacts, salt bridges, disulphide bonds, cation-π interactions, metal coordinated bonds, and local hydrophobic packing. This enabled calculation of network metrics to give a quantitative measure of the topological importance of each residue. The authors validated this approach with non-HIV proteins of known structure and function and showed highly significant inverse correlations between network score and experimentally determined mutational tolerance. Therefore, they could argue that network scores give an accurate measure of fitness costs of a mutation at any given site within an epitope. Gaiha *et al.* analyzed the protective value of all well-defined peptide epitopes and correlated this with the strength of T cell responses to that peptide and the control of viremia (see the figure). The findings provide an explanation for protection against AIDS progression on the basis of HLA type and clarify why some people without "good" HLA molecules can have good immune control over infection, whereas others with the same protective HLA types, whose T cells focus on less constrained epitopes, do not control HIV-1.

Previous quantitative fitness landscape analyses predicted vulnerable positions in the HIV proteome (8–10). Rather than the structure-based approach used by Gaiha *et al.*, these studies used statistical correlations to map interacting networks—that is, the rate of immune escape was related to the amino acid sequence landscape of the virus. This enabled the identification of amino acids outside the epitope that were agonistic or antagonistic to escape from immune attack. A combination of epistatic and topological network analyses could be even more informative because variation at contact residues is factored in.

Another parameter to consider is immunodominance, in which the epitope that stimulates the strongest T cell response exerts the strongest selective pressure (11). Dominant T cells compete with and suppress less frequent T cells, but it is not clear why particular T cell responses dominate and why this varies between individuals (12). In the context of HIV infection and epitope fitness, it clearly matters which epitopes become immunodominant.

Gaiha *et al.* argue that understanding the topological network score of an epitope will help vaccine design. Only one experimental simian immunodeficiency virus (SIV)/HIV–specific T cell–stimulating prophylactic vaccine has been shown to be capable of enabling complete clearance of SIV infection soon after experimental challenge (in >50% of monkeys) (13). However, there are many unusual features of the CD8+ T cell responses elicited by this vaccine, including restriction of responses by major histocompatibility complex (MHC) class II and by nonclassical MHC Mamu-E, as well as the exceptional breadth of the T cell responses (vertebrate MHC includes human HLA and monkey Mamu). It is unlikely that the infecting SIV could escape from T cells that recognize multiple SIV epitopes at the same time. Thus, a focus on networked or conserved epitopes might not be necessary if an HIV vaccine elicits such broad T cell responses. On the other hand, vaccine-stimulated CD8[+] T cell responses against specific highly networked epitopes of HIV-1 may be valuable in therapeutic strategies, allowing the patient to eradicate latent virus reservoirs that become reactivated (14). A vaccine that focuses T cell responses on highly networked epitopes, which are often not naturally immunodominant, could be particularly effective.

ACKNOWLEDGMENTS

REFERENCES AND NOTES

1. Carrington M, Walker BD, Annu. Rev. Med 63, 131 (2012). [PubMed: 22248321]

2. Fellay J et al., Science 317, 944 (2007). [PubMed: 17641165]

3. International HIV Controllers Study et al., Science 330, 1551 (2010). [PubMed: 21051598]

4. Phillips RE et al., Nature 354, 453 (1991). [PubMed: 1721107]

5. Gaiha GD et al., Science 364, 480 (2019). [PubMed: 31048489]

6. Pereyra F et al., J. Virol 88, 12937 (2014). [PubMed: 25165115]

7. Smith SM, Retrovirology 1, 8 (2004). [PubMed: 15169568]

8. Dahirel V et al., Proc. Natl. Acad. Sci. U.S.A 108, 11530 (2011). [PubMed: 21690407]

9. Ferguson AL et al., Immunity 38, 606 (2013). [PubMed: 23521886]

10. Barton JP et al., Nat. Commun 7, 11660 (2016). [PubMed: 27212475]

11. Liu MK et al., J. Clin. Invest 123, 380 (2013). [PubMed: 23221345]

12. Chen W et al., Immunity 12, 83 (2000). [PubMed: 10661408]

13. Hansen SG et al., Nature 502, 100 (2013). [PubMed: 24025770]

14. Deng K et al., Nature 517, 381 (2015). [PubMed: 25561180]

Connections analysis

Quantification of interconnectivity between amino acids within each HIV protein provides a topological network score for each residue. The score reflects constraints on mutation, which affects viral fitness and the ability of the immune system to clear the mutated virus.

HLA polymorphism

HIV-1–infected cell
(bearing HLA-peptide complex)

CD8+ T cells
(immunodominant)

Topological network analysis of HIV epitopes and mutations

p24
(an HIV-1 protein)

Low network score
(Leu83 –> Met mutation
in epitope AV9)

High network score
(Pro34 –>Ala mutation in epitope KF11)

Fitness cost

Escape detection by
CD8+ T cells

No escape from
CD8+ T cells

Decreased
virulence

Poor viral control

Viral control