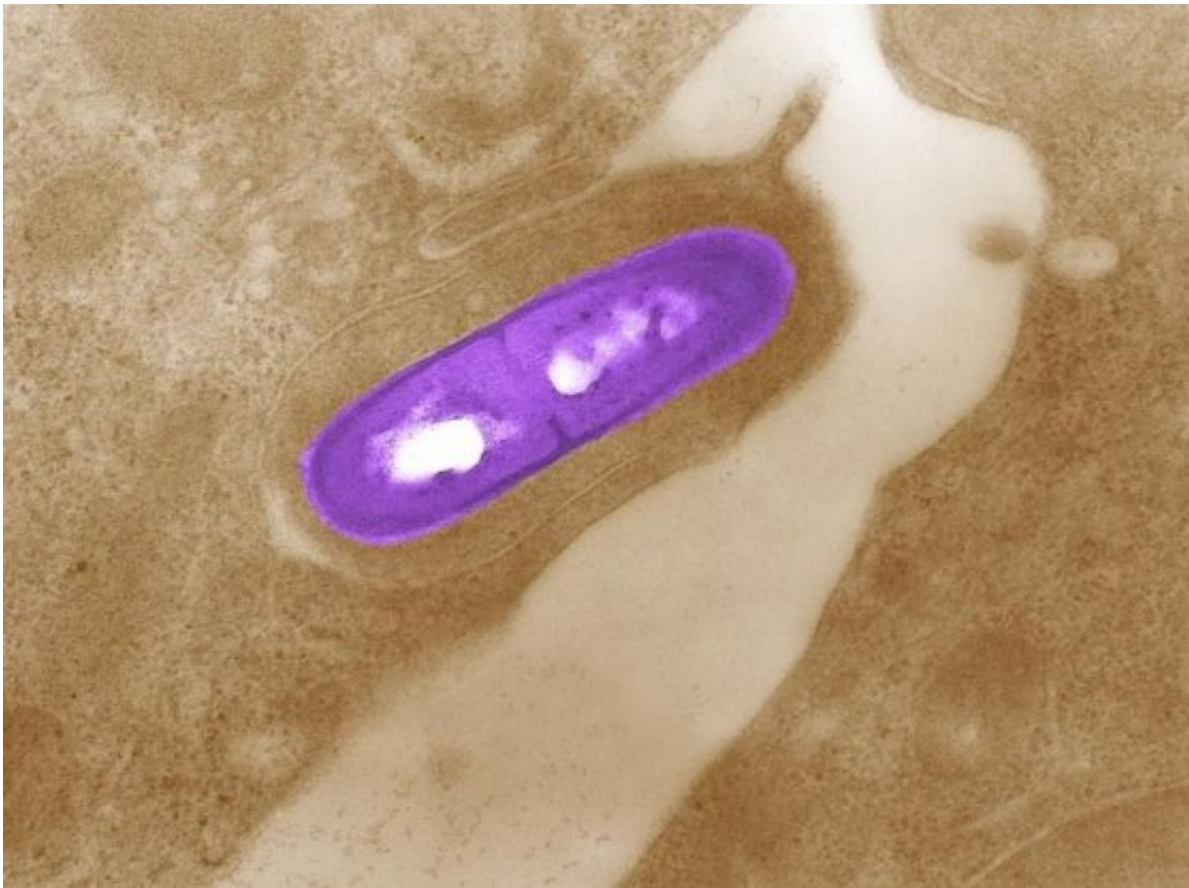# Machine learning tool predicts the potential of peptides as immune activators

October 23 2018, by Tom Ulrich

Listeria monocytogenes. Credit: CDC PHIL

The immune system keeps T cells under control by regulating precisely when they can respond to a pathogen. For instance, helper T cells only turn "on" if other immune cells, such as antigen-presenting cells (APCs)

present bacterial peptides (protein fragments) on their surface in a protein complex called MHC class II (MHC II).

However, not every bacterial peptide is immunodominant (gets loaded into MHC II and presented to T cells); nor is every peptide bound to this complex antigenic (capable of provoking an immune response). The rules that govern these dynamics are not yet fully known, muddling efforts to better understand the relationships between us as hosts, the pathogens that infect us, and our microbiomes.

To bring some clarity, a team led by Daniel Graham, Chengwei Luo, and core institute member Ramnik Xavier in the Broad's Infectious Disease and Microbiome Program have developed a deep neural network-based algorithm called BOTA (Bacteria Originated T cell Antigen) capable of predicting, based on a bacterial genome data, peptides with the highest chance of triggering an immune response.

As they reported in *Nature Medicine*, Graham and the team (which included members of Massachusetts General Hospital's Center for the Study of Inflammatory Bowel Disease and Center for Computational and Integrative biology, as well as Massachusetts Institute of Technology's Center for Microbiome Informatics and Therapeutics) built and trained BOTA to recognize potential antigens by running a "peptidomic" study of MHC II, collecting and characterizing every MHC II-bound peptide natively found in APCs in mice and formulating a list of features underlying immunodominance and antigenicity.

Graham then benchmarked BOTA in two other mouse models, of Listeria monocytogenes infection and of colitis, assessing its predictions using a high-throughput, single-cell RNA-sequencing screening test that measured whether T cells could "see" predicted peptides and how strongly they reacted.

The algorithm, the team found, accurately predicted which bacterial

peptides bound to MCH II in both models. Their RNA-sequencing data also helped identify the peptides that sparked the strongest T cell responses in their Listeria model.

The team's findings suggest that BOTA could help researchers in a number of scenarios, from discovering previously unknown bacterial antigens to improving vaccine design, and from illuminating how the microbiome tunes the immune system to understanding how that tuning breaks down in inflammatory conditions.

More information: Daniel B. Graham et al. Antigen discovery and specification of immunodominance hierarchies for MHCII-restricted epitopes, *Nature Medicine* (2018). DOI: 10.1038/s41591-018-0203-7

Provided by Broad Institute of MIT and Harvard