Synthetic-endogenous RSS comparison tool

Arguably one of the most fascinating feats of the immune system in humans, and really in any jawed vertebrate, is not only its ability to produce a seemingly endless array of antibody types to identify any invasive bacterium or virus-infected cell, but also the elegance of making antibody assembly a matter of cutting and pasting different regions of DNA together. This assembly called V(D)J recombination involves taking DNA sequences encoding parts of these antibodies and stitching them together and has proven to be one of the clever ways in which organisms keep the genome from taking on an unwieldy number of antibody genes while producing the necessary antibody on demand. The initial cutting part of V(D)J recombination involves the recombination-activating gene (RAG) protein, as anthropomorphized into hands in the figure, creating a loop in the DNA by grabbing onto two regions of the DNA, shown with purple and blue backbones, adjacent to these antibody-encoding segments so that they may be later joined. These two regions exhibit a series of sequence patterns that is amenable to RAG binding. RAG will eventually cut the DNA between these sites and the antibody-encoding portions for other proteins to join the antibody gene segments to yield an antibody combination.

While RAG manages to attach to and cut certain regions of the DNA because of the recognizable sequence patterns, these identifiable patterns are interspersed among DNA sequence positions that have less obvious patterns. In a study that was recently published in *Nucleic Acids Research*, we set out to understand how switching out nucleotides at one or multiple positions of these binding sites from a common initial sequence alters the extent to which RAG will hold the DNA and cut it. We illustrate some of our findings in this interactive figure, which is modified from a page in the Supplementary website that accompanies our publication. In this interactive figure, we show data from three naturally occurring binding site sequences and compare them to data from the initial binding site sequence and data collected on a single nucleotide change that accounts for the sequence difference between the initial sequence and the naturally-occurring one. Through the dropdown menu, one can select one of these three binding site sequences to reveal its sequence, with positions where the nucleotide is changed from the initial sequence colored in. The plot immediately below it shows the frequency that RAG creates a DNA loop with

For simplicity, we also make it easier to identify how a single nucleotide change compares to the naturally occurring sequence of which it contributes to the sequence change. Hovering the mouse over one of the colored nucleotides will send the rest of the data into the background for ease of visualizing data from the single nucleotide change and the naturally occurring sequence.

The data we present here shows that some nucleotides can have a dominating influence on how well RAG can hold or cut the DNA.