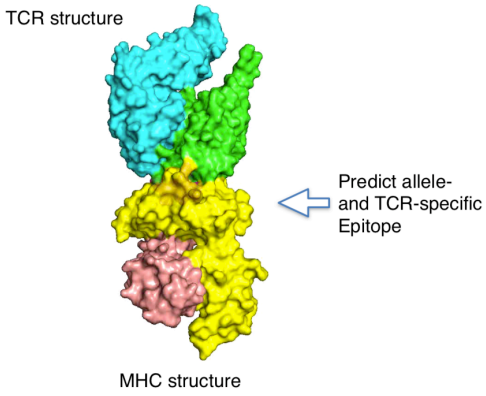
Systems Immunology Lab (Genome informatics)

Two areas where we are actively working both collaboratively and developing new methods are B/T cell repertoires and protein-nucleotide interactions. Both projects rely heavily on multiple sequence alignment (MSA), structural modeling and machine learning. We are actively developing the MAFFT multiple alignment software, which has numerous features for supporting specific use-case scenarios (Katoh, K. and Standley, DM *Mol. Biol. Evol.* (2013)).

B/T Cell Repertoires

We are currently working intensively on high-throughput analysis of B cell receptors (BCRs) and T cell receptors (TCRs). BCR and TCR "repertoires" can be sequenced from a routine blood sample and provide a highly sensitive biomarker for any perturbation to the immune system. Following the well-established paradigm in protein evolution that "structure is more conserved than sequence," our lab is developing methods to quantify predict the epitopes of T cell receptors. As a necessary first step, we have developed a high-throughput BCR and TCR modeling platform, Repertoire Builder, that is both fast and accurate when compared with other modeling tools, including our own earlier tools (Shirai, H. et al *Proteins* (2014); Yamashita, K. et al. *Bioinformatics* (2014)). We are applying our modeling tools understand autoimmunity and infections diseases.



Protein-Nucleotide Interactions

The interaction between proteins and nucleotides (DNA or RNA) is critical for proper regulation of immune responses as well as for direct detection and elimination of viral infections. Protein-nucleotide interactions can also be hijacked by pathogens or tumors to thwart detection by the immune system. Because of their importance in various aspects of immunology, we have developed a tool called aaRNA to identify RNA binding sites on RNA-binding proteins (Li, S. et al. *Nucleic Acids Res*(2014)). We have extended aaRNA to the prediction of DNA-binding sites (aaDNA) and incorporated the binding propensities in flexible docking simulations. As demonstrated by several studies, the predicted nucleotide binding sites agree well with experiment and provide valuable insight into the molecular mechanisms of protein-nucleotide interactions (Hanieh, H. et al. *Eur. J. Immunol.*; Nyati, K. K. et al. *Nucleic Acids Res* (2017); Yokogawa et al. *Si Rep* (2016); Masuda, K. et al. *J Exp Med* (2016); Mino, T. et al. *Cell* (2015)).