要　旨

It is well established that protein structure is more conserved than sequence on an evolutionary timescale. This fact allows functional inferences to be drawn from proteins that share the same fold, even when their sequence similarity is quite low. In the case of B cell receptors, the relationships between sequence and structure and function are more complex. Most BCRs look similar globally but differ in the details of their antigen-binding regions. These differences are due to the fact that each BCR is assembled from a patchwork of genes, which are combined randomly and can diverge further by random mutations upon antigen encounter. Traditionally, bioinformatics analysis of BCR sequences involves clustering those that arise from the same genes into "lineages", in order to identify BCRs in a given donor that target a common antigen (Fig.1A).

The diversity of BCRs has been estimated to exceed 1013 in humans, which means that it is very unlikely that any two donors will display the same repertoire of BCRs, even after exposure to the same antigen. Nevertheless, x-ray crystallographic studies have demonstrated that structurally and sequentially similar BCRs targeting common antigens can arise in different donors using different genes. Our hypothesis is that clusters of BCRs targeting the same antigen are more likely to have sequence and structural features in common than BCRs targeting different antigens. High-throughput sequencing methodologies can now deliver paired (heavy-light chain) sequence datasets on the order of 104 sequences per experiment, and are expected to improve rapidly in the near future. Clearly, x-ray crystallography will not be able to cope with so many emerging BCR sequences in a high-throughput manner. Thus, there is a strong motivation to leverage structural bioinformatics in order to infer structure and functional similarities. In this presentation, I will show results from our high-throughput BCR and TCR structural modeling platform (Repertoire Builder). Using multiple alignment and 3D rendering methods developed in our lab, we could reduce the time required to build an atomic-resolution BCR model to just seconds, corresponding to over 17,000 atomic resolution models per day on a single CPU (Fig. 2). We then show that human BCRs acquired post flu vaccination indeed display strong structural convergence (Fig. 1B), and even exhibit structural similarities to BCRs acquired from vaccinated mice. These findings suggest that BCR modeling, in combination with high-throughput sequencing may be able to identify diverse sequences targeting common antigens across donors and across species.