Sequencing single B cells

By Lauren Martz, Staff Writer

Researchers from **The University of Texas at Austin** have developed a high throughput method for sequencing individual B cell immunoglobulin receptors that keeps the link between their paired heavy and light variable chains intact.[1] The technique increases the efficiency of identifying disease biomarkers and designing therapeutic antibodies but requires modifications to increase throughput and accuracy.

B cell receptors help the immune system detect antigens and trigger antibody production. An individual harbors a repertoire of B cell receptors specific to a range of antigens to which the person has been previously exposed. Sequencing each B cell receptor, which consists of a variable heavy (V_H) and a variable light (V_L) chain, in a patient could be useful for generating therapeutic vaccines and antibodies as well as for designing disease diagnostics.

But this would be a challenging task for most sequencing methods because of the need for large quantities of B cell DNA as starting material. Because each B cell has a unique B cell receptor, pooling B cells results in loss of the original V_H:V_L pairing of individual B cell receptors.

Existing single cell–based sequencing approaches, such as single-cell PCR, could preempt this problem, but they typically are complex, expensive and time consuming.

Thus, George Georgiou and colleagues at the University of Texas at Austin set out to design a new single-cell approach that would preserve the original V_H:V_L pairing while increasing throughput and lowering cost.

Georgiou is chair of engineering and professor of molecular genetics and microbiology at the university. The team also included co-investigator Andrew Ellington, professor of chemistry and biochemistry at the University of Texas at Austin, and researchers from the **University of Houston**, **Charité University Medicine**, the **German Rheumatism Research Center** and **The University of Chicago**.

The team's strategy involved depositing B cells onto high-density well plates with the goal of having at most one cell per well with a 95% probability.

The plates were then covered and incubated with lysis solution and magnetic beads to release and capture the mRNA from the individual cells, respectively.

The team used RT-PCR followed by linkage PCR to amplify these genetic materials and sequenced the linked heavy and light chains' most variable regions using the MiSeq high throughput sequencing platform from **Illumina Inc.**

To show that the method could sequence endogenous B cell receptor V_H:V_L pairs, the team sequenced three different patient B cell subpopulation repertoires: paired V_H and V_L sequences of immunoglobulin-expressing B cells from a healthy donor, plasmablasts from a healthy donor immunized against tetanus toxin and peripheral memory B cells from a healthy donor after influenza immunization. The sequencing strategy identified thousands of paired V_H:V_L sequences, including those of known controls that were included in cell samples.

To validate the identified sequences and show that the sequencing data are useful for identifying and designing therapeutic antibodies, Georgiou and colleagues expressed 10 of the paired sequences from the tetanus toxin–immunized donor as immunoglobulin in human embryonic kidney cells. The resulting antibodies all bound tetanus toxin, showing the sequences were paired accurately and suggesting they might be useful as functional antibodies.

Results were published in *Nature Biotechnology*.

"Heavy-chain sequences are important because they can be used as markers for antibodies. These are useful for applications such as tracking B cells, monitoring for minimal residual disease or diagnostics. But having both chains together allows you to reconstruct monoclonal antibodies produced by the B cells to really get at their functions."

—Harlan Robins, Adaptive Biotechnologies Corp.

Sequencing applications

This single-cell method not only could help sequence B cells more accurately than traditional high throughput approaches but also should offer cost and speed advantages over other single-cell sequencing methods.

There are, however, technological hurdles as additional modifications are needed to further increase throughput and decrease the error rate.

"Heavy-chain sequences are important because they can be used as markers for antibodies. These are useful for applications such as tracking B cells, monitoring for minimal residual disease or diagnostics. But having both chains together allows you to reconstruct monoclonal antibodies produced by the B cells to really get at their functions," said Harlan Robins, cofounder of **Adaptive Biotechnologies Corp.**

For example, he said, "the paired sequences have the clinical potential of rapidly improving vaccine design and identifying neutralizing antibodies. Right now, identifying therapeutic antibodies requires a hefty set of experiments. You can also create a set of mono- or polyclonal antibodies from a person who has had a response to a pathogenic challenge in order to develop a strong therapeutic."

Adaptive is developing immunoSEQ, a sequencing service capable of generating sequences of the immunoglobulin heavy chain and T cell receptors (TCRs).

The method developed by Georgiou's team could sequence thousands of paired heavy and light chain regions in 10 hours over 4 days for about $550. Using single-cell PCR to sequence the same number of pairs would take more than 10 weeks and would cost $25,000.

Guy Hermans, principal scientist at **Ablynx N.V.**, said that compared with single-cell sequencing approaches for pairing heavy and light chains, "the benefits of the technology include maintaining the correct pairing with high throughput using low reagent and labor costs."

Ablynx CSO Andreas Menrad added that the new method has the potential to process small samples, which could help identify T or B cells infiltrating solid tumors or inflammatory lesions.

Ablynx's Nanobody platform involves the design of antibody-derived therapeutics based on only the heavy chain of immunoglobulin.

"A huge advantage for this new approach is that the complicated part of the assay does not have to be done with live cells at the sites of the early steps," added Robins. "This is important because it could allow a pharma company doing a clinical trial to just put the samples on the plates at the trial sites and then run the sequencing experiments later."

One important drawback, he said, is that the throughput of the current approach is not that high. Thus, Robins thinks the logistics of scaling up the approach could be difficult.

"For this technology, it requires you to have far less than one cell per well. If wells with more than one cell result, you end up with false positives and incorrect cross-pairings. Therefore, there may be a limit to how quickly you can scale it up."

He suggested that more data could be obtained simply through preparation and sequencing of additional plates.

Robins said another issue is that the researchers used as active controls B cells that were spiked with immortalized IM-9 lymphoblast cells. Those control cells "look different than the other cells, and the assays tend to favor those cells, making the results look better than they are," he said. "It is possible that this could lead the researchers to get some false positives."

Hermans noted that once the sequences are determined, the group also does not have a method in place to screen the antibodies for antigen binding. "The bottleneck of the approach is the processing and validation of the downstream sequencing results," he said.

William Robinson, associate professor of medicine in the division of immunology and rheumatology at the **Stanford University School of Medicine** and cofounder of **Atreca Inc.**, told *SciBX*, "Incomplete variable region sequences are obtained, and thus complete antibodies are not characterized in a high throughput fashion." He said that to address the problem, synthesis of additional primers, additional PCR reactions and additional sequencing reactions will be required to obtain complete variable region sequences.

Atreca's Immune Repertoire Capture technology uses barcoding to generate full-length heavy and light chain variable region sequences.

The incomplete variable regions are "only a short-term problem," noted Robins. "High throughput sequencing is evolving so quickly it should be able to catch up in a few years."

Georgiou told *SciBX* that patent applications covering the technology have been filed. His team has not yet licensed the technology and has not determined how to move forward commercially.

Martz, L. *SciBX* 6(5); doi:10.1038/scibx.2013.108
Published Feb. 7, 2013

REFERENCES

1. DeKosky, B.J. *et al. Nat. Biotechnol.*; published online Jan. 20, 2013; doi:10.1038/nbt.2492

Contact: George Georgiou, The University of Texas at Austin, Austin, Texas
e-mail: gg@che.utexas.edu

COMPANIES AND INSTITUTIONS MENTIONED

Ablynx N.V. (Euronext:ABLX), Ghent, Belgium
Adaptive Biotechnologies Corp., Seattle, Wash.
Atreca Inc., San Carlos, Calif.
Charité University Medicine, Berlin, Germany
German Rheumatism Research Center, Berlin, Germany
Illumina Inc. (NASDAQ:ILMN), San Diego, Calif.
Stanford University School of Medicine, Stanford, Calif.
The University of Chicago, Chicago, Ill.
University of Houston, Houston, Texas
The University of Texas at Austin, Austin, Texas