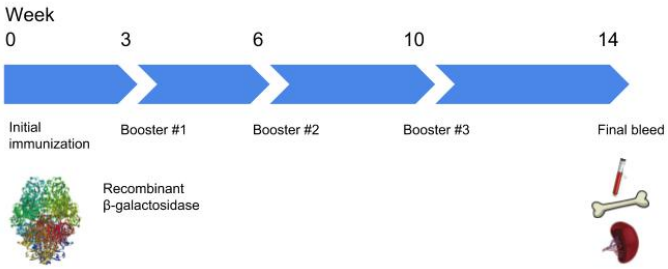
Alicanto

DIGITAL PROTEOMICS

Custom rabbit monoclonal antibody discovery from serum

Alicanto® is a custom antibody discovery service that harnesses over a decade of technology development across genomics and proteomics(*1–6*). In this case study we demonstrate how Alicanto is employed to discover high affinity rabbit monoclonal antibodies. From this project, *two antibodies with picomolar affinity to β-galatosidase were discovered* in rabbit.



Immunization

Two juvenile New Zealand white rabbits were immunized with a recombinant *Escherichia coli* β-galatosidase (lacZ) antigen in complete Freund's adjuvant. Additional boosters of incomplete Freund's adjuvant were given at weeks 3, 6, and 10 (See figure above). The serum was processed in a column with the same β-galatosidase used for the immunizations in order to enrich for target-specific antibodies. After affinity purification, the enriched polyclonal antibodies were analyzed via ELISA for desired reactivity to the target. At this point, if desired reactivity is not achieved in either rabbit, the customer may elect to discontinue the project. Otherwise, we pick the best responding rabbit and continue with antibody discovery.

Data generation

B cell selection and next-generation sequencing

Next-generation sequencing libraries were prepared from peripheral blood mononuclear cells (PBMCs) at 9 time points as well as from spleen and bone marrow. RNA was extracted from the cells or directly from tissue and reverse transcribed into cDNA. Using variable region-targeting primers, we amplified the entire variable region of the IgG and IgK transcripts. Our variable region-targeting primers anneal just outside of the variable region in order to amplify the entire region without masking potential somatic mutations at the primer sites. A library was constructed from the 9 PBMC samples and 2 tissue samples, and sequenced on the Illumina MiSeq System. In total 1.4M and 1.9M paired end reads were recovered from the heavy chain and light chain transcripts, respectively.

Repertoire construction and analysis

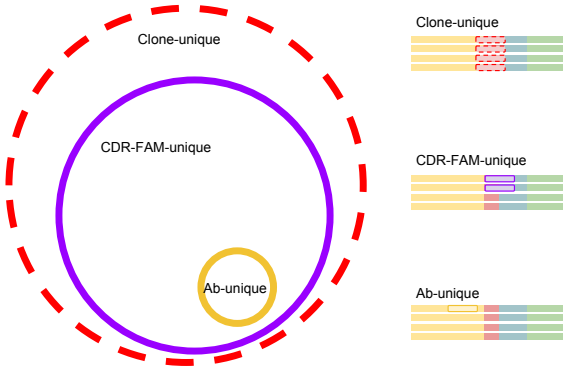
We constructed a repertoire from the reads by performing quality filtering, pair stitching, and error correction (*4*). Each input sample was processed individually and information about the source tissue of each antibody in the repertoire was retained. Roughly 70,000 unique heavy chain sequences and 80,000 unique light chain sequences were constructed from each PBMC time point and tissue. After combining all time points, and applying a minimum abundance threshold of 2, a total of 62,486 unique sequences were recovered. Each nucleotide sequence was translated to create an amino acid sequence database for analysis with mass spectrometry.

Each antibody was further analyzed to determine the germline V, D, and J genes that produced the clone, a process called V(D)J-labeling (*5*). The three complementarity-determining regions (CDRs) for each antibody were annotated. Antibodies with nearly identical CDR3's were clustered together into *Clone Clusters*.

Mass spectrometry analysis

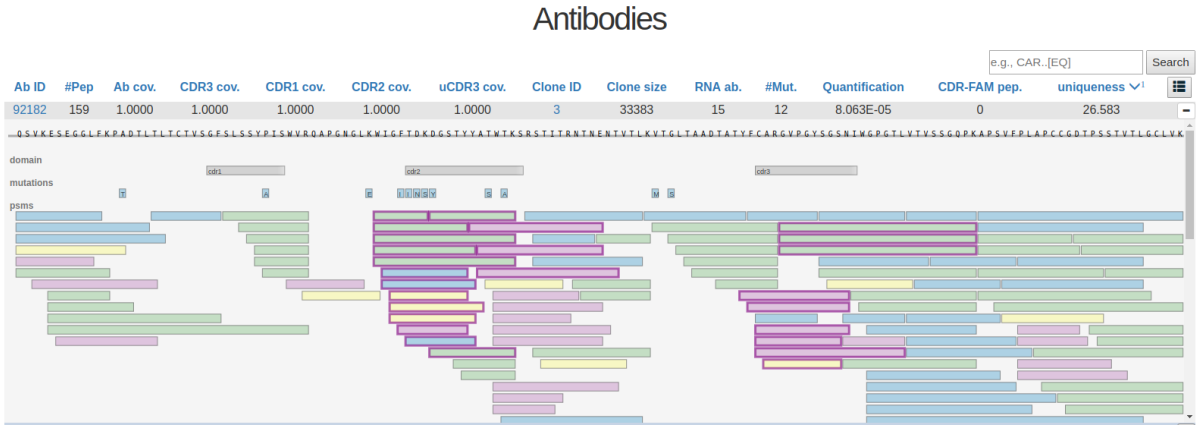
The enriched polyclonal antibodies were analyzed by LC-MS/MS. Briefly, the antibody chains were separated, and digested with 4 distinct proteases: trypsin, chymotrypsin, elastase, and pepsin. Each protease exhibits unique cleavage patterns, yielding a diverse population of peptides across the full-length of the antibodies. Each of the 8 chain x protease combinations was analyzed on an Orbitrap Fusion Lumos Tribrid, with doublet HCD/EThCD fragmentation for each precursor. In total ≈332,000 tandem mass spectra (MS/MS) were acquired across both chains.

Data analysis



MS/MS spectra from heavy and light chain polyclonal antibodies were mapped to their respective amino acid sequence databases. As is expected, many of the peptides mapped to regions of the antibody that are highly degenerate. We adopted a nomenclature (as shown in the figure to the right) for determining how informative each peptide may be. The most informative type of peptide maps to only a single antibody sequence in the database and is dubbed *Ab-unique*. A peptide which maps to multiple antibody sequences that all share identical CDR1, CDR2, and CDR3 is called a *CDR-FAM-unique* peptide. A third set of informative peptides map to only sequences within the same *Clone Cluster* and are called *Clone-unique* peptides. Using this peptidic information, each antibody sequence was transformed into a feature vector and input to the Alicanto candidate selection model. In total 13 heavy chain sequences and 7 light chain sequences were selected for validation.

In the figure below, a screen shot from the Alicanto visualization web application is shown. A single antibody in the repertoire is displayed with the peptides (small rectangles) mapped to the sequence. Purple bordered peptides are CDR-FAM-Unique peptides, while the non-bordered peptides are shared between thousands of antibodies.



Candidate validation

In bulk B-cell sequencing, native pairing information is lost. In order to validate full-size antibody candidates, we evaluated all combinations of the 13 heavy chain and 7 light chain sequences. Absolute Antibody, Ltd. performed small-scale transient expression of all 91 antibodies yielding 3 candidates that were target-reactive by ELISA. Absolute Antibody scaled up production of the 3 antibodies for kinetic analysis. Two of the three antibodies demonstrated picomolar binding to β-galatosidase, which was 2 orders of magnitude better binding than the positive control antibody.

1. N. Bandeira, V. Pham, P. Pevzner, D. Arnott, J. R. Lill, *Nature biotechnology* **26**, 1336 (2008).
2. N. E. Castellana, V. Pham, D. Arnott, J. R. Lill, V. Bafna, *Mol. Cell Proteomics* **9**, PMCID: PMC2877985, 1260–1270 (June 2010).
3. S. Woo *et al.*, *Proteomics* **14**, 2719–2730 (Dec. 2014).
4. Y. Safonova *et al.*, *Bioinformatics* **31**, 53–61 (June 2015).
5. S. R. Bonissone, P. A. Pevzner, *Research in Computational Molecular Biology*, 44–59 (2015).
6. S. W. Cha, S. Bonissone, S. Na, P. A. Pevzner, V. Bafna, *Molecular & Cellular Proteomics* **16**, 2111–2124 (2017).