Single domain antibody discovery in llama using NGS and mass spectrometry

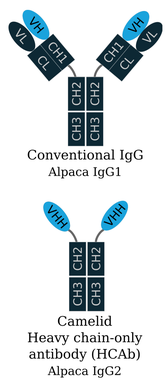
Stefano Bonissone, Anand Patel, Thiago Lima, Natalie Castellana

DIGITAL PROTEOMICS

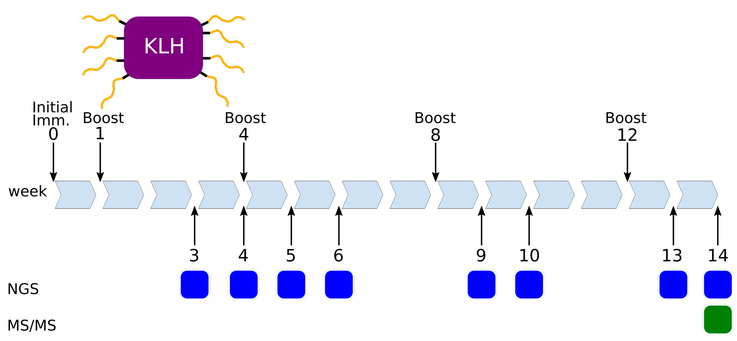
Overview

- Heavy chain-only antibodies (HCAbs) from camelids (llama, alpaca, camel, etc.) have unique stability and developability qualities.
- Immunodominant epitopes can make antibody discovery against other epitopes challenging.

The goal of this study was to discover single chain antibodies from a camelid immune response to both the immunodominant and non-immunodominant epitopes of the immunogen. We use a proteogenomic approach to discover the antibodies.



Experimental design



A llama was immunized with a peptide representing an epitope of rabbit CD20 conjugated to Keyhole Limpet Hemocyanin (KLH) according to a standard 14 week schedule.

Sequencing: HCAb transcripts from multiple time points were amplified and sequenced.

Mass spectrometry: HCAbs from serum isolated at the final time point were enriched for antigen-specificity against KLH or the peptide and subjected to LC-MS/MS.

Data snapshot

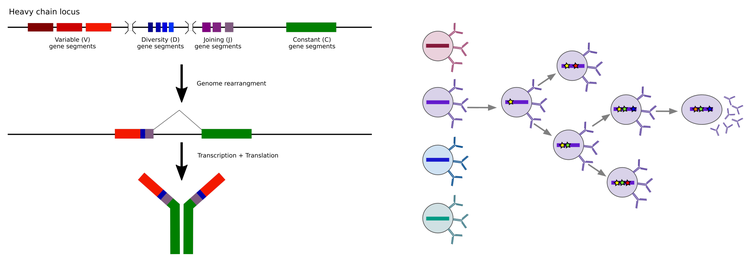
We constructed a repertoire from the reads by performing quality filtering, pair stitching, and error correction.

- **IgG2b (long hinge)** - 128,233 unique IgG2b sequences were recovered.
- **IgG2c (short hinge)** - fewer than 2,500 distinct sequences recovered from each timepoint.

The serum antibodies were fractionated using affinity chromatography and isotype separation into an anti-CD20 (aCD20) IgG2b fraction and an anti-KLH (aKLH) IgG2b fraction. In total 110,911 tandem mass spectra were generated for the aCD20 fraction and 139,795 mass spectra were generated for the aKLH fraction.

Challenges of antibody inference

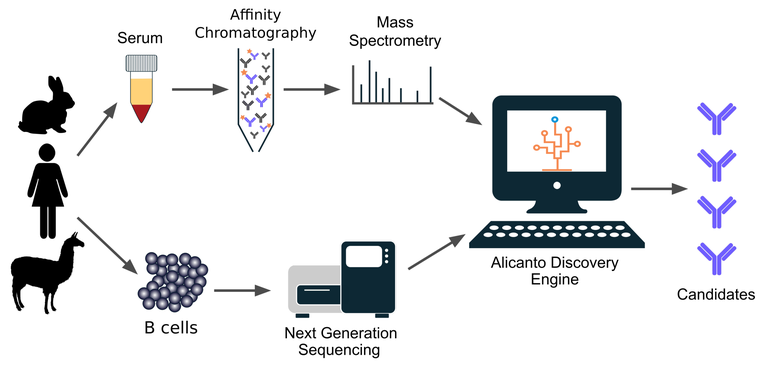
The vast majority of peptides are shared by hundreds or thousands of sequences in the database. The extreme degeneracy is a direct result of B cell development and maturation.



The collection of antibody sequences derived from the same rearrangement event are called a **clonal lineage**.

Alicanto Pipeline

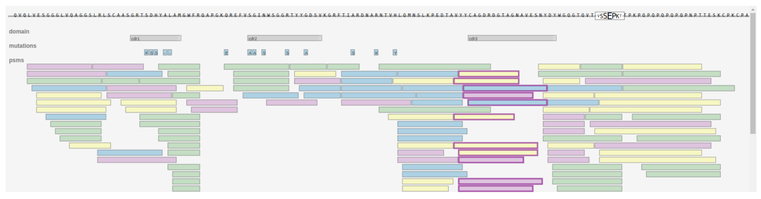
Alicanto combines proteomic analysis of antibodies in serum with sequencing analysis of antibody transcripts from B cells to characterize the expressed antibody repertoire. Additionally, Alicanto identifies serum antibodies that show activity toward the immunogen.



The output of Alicanto is a diverse collection of sequences that are present in both the B-cell repertoire and the target-enriched serum. The diversity of the result varies widely depending on the complexity of the initial immunogen and the stringency of the affinity purification.

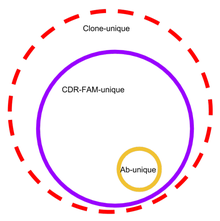
Selecting antibodies from the repertoire

In the figure below, peptides are mapped to a specific sequence in the database. Peptides unique to a clonal lineage of antibodies are highlighted in purple. All other peptides are shared by antibodies in other clonal lineages.



Identified peptides can be classified based on how unique they are to a particular antibody or clonal lineage.

- **Ab-unique** peptides uniquely map to a single antibody sequence in the repertoire.
- **CDR-FAM-unique** peptides uniquely map to only antibodies with the same CDR1 and CDR3.
- **Clone-unique** peptides map to only antibodies from the same clonal lineage.
- **Shared** peptides map to multiple antibodies from different clonal lineages.



Using Alicanto, we selected clonal lineages and antibodies within those lineages using the collection of identified of peptides. Since KLH is a much larger protein, is known to be highly immunogenic, and has more epitopes, we found many more clonal lineages targeting KLH than the peptide. In total we found 67 clonal lineages targeting KLH and 4 targeting the peptide.

Digital Proteomics

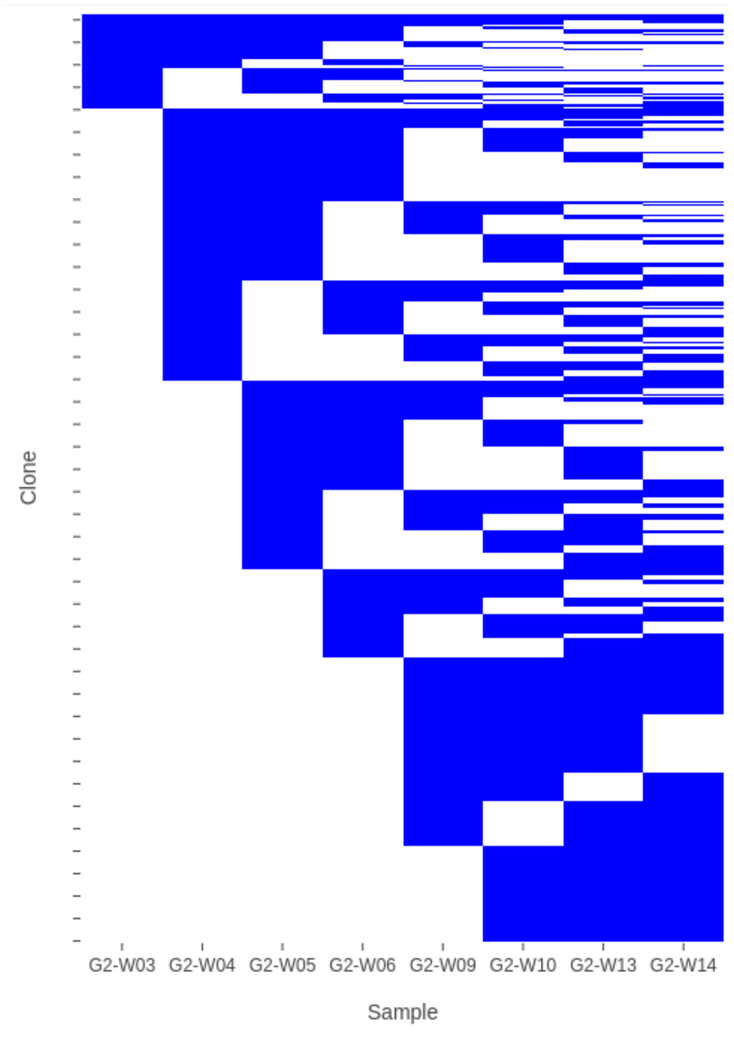
Read
more about all of our antibody discovery and sequencing services on our website

www.digitalproteomics.com

Temporal dynamics of serum antibodies

We plotted the occurrence of each clonal lineage in each of the sampled time points to assess the temporal dynamics of the repertoire in response to immunization.

- Few clones appear in all time points.
- By the final two weeks of the immunization, no new clones appear.
- More new clones are observed 1 week after boost (W5, W9) than 2 weeks after boost (W3, W6).

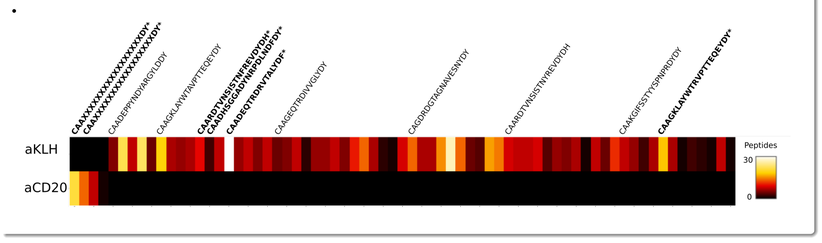


Data analysis

Mass spectra from each fraction were independently mapped to the IgG2b repertoire. Alicanto determined the set of CDR3 sequences present in the repertoire.

- Four CDR3s that were clustered into two related clone clusters were found targeting the CD20 peptide
- 65 distinct CDR3s in 20 clone clusters were found were found targeting KLH.

.



VHH Candidate validation

We validated select VHHs by verifying binding specificity to the desired antigen using ELISA.

- **aCD20**: We selected three VHHs for validation. Two of the three (66%) reacted to the peptide and not KLH.
- **aKLH**: We selected 10 candidates VHHs, and found four (40%) reacted to KLH and not the peptide.

Relevant resources

- Y. Safonova et al.,Bioinformatics 31, 53–61 (June 2015)
- S. R. Bonissone, P. A. Pevzner, Research in Comp. Mol. Biol., 44–59 (2015)
- C. Hamers-Castermanet al., Nature 363, 446–448 (June 1993)