Alicanto®

Diverse VHH discovery from serum

Camelids, such as llamas and alpacas, naturally produce both conventional antibodies and heavy-chain only antibodies (HCAbs) as depicted in Figure 1. The antigen-binding domain of the HCAb is called the VHH.

Alicanto® is a custom antibody discovery service that delivers diverse, natural VHHs without laborious panning. Alicanto® 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 diverse, HCAbs from camelids. The key differentiator of the Alicanto® process is its direct analysis of target-specific HCAbs from serum. While phage display methods discard the serum after screening to ensure the immunization was successful, Alicanto® uses the pre-screened polyclonal antibodies to identify a more diverse panel of assay ready VHHs.

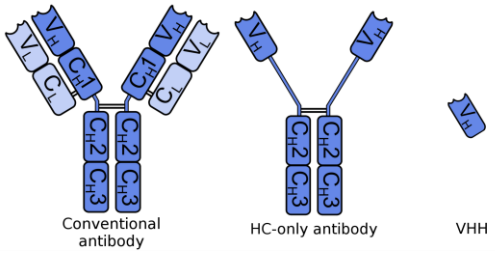

Figure 1: The antibody on the left contains both the heavy chain and light chain and is typical of the IgG1 antibody in llama and conventional antibodies produced by most other animals. The HC-only antibody (HCAb) in the center is typical of IgG2 and IgG3 antibodies in llama. The VHH fragment on the right consists of the variable domain of the HCAbs.

Immunization

We immunized a single adult male llama with a peptide for rabbit B-cell surface marker CD20 conjugated to KLH. The llama was immunized 4 times over 14 weeks, and bleeds were taken 1 week and 2 weeks after each boost. Peripheral blood mononuclear cells (PBMCs) were isolated from each bleed by Ficoll separation. At week 14, serum was collected.

Data generation

B cell selection and next-generation sequencing

Next-generation sequencing libraries were prepared from PBMCs at 8 time points. We extracted RNA from each sample, constructed cDNA, and amplified the HCAb transcripts using primers annealing to the IgG2b (long hinge) and IgG2c (short hinge) as well as the leader peptide. The transcripts were sequenced on our Illumina MiSeq, and we constructed a repertoire containing IgG2b and IgG2c HCAbs. Our primers anneal just outside of the variable region in order to amplify the entire region without masking potential somatic mutations at the primer sites.

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 bleed and isotype of each antibody in the repertoire was retained. Roughly 128,233 unique IgG2b sequences were recovered. IgG2c had lower expression which resulted in a small repertoire for this isotype, with fewer than 2,500 distinct sequences recovered from each time point. Since few IgG2c target-specific antibodies were purified for mass spectrometry analysis, the IgG2c antibodies were omitted from further analysis. 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*. Antibodies with similar CDR3s assumed to have similar specificity.

Mass spectrometry analysis

The serum antibodies were fractionated using affinity chromatography and isotype separation (*7*). Two samples were analyzed by Alicanto®; the anti-CD20 (aCD20) IgG2b fraction and the anti-KLH (aKLH) IgG2b fraction. Each fraction was analyzed using our standard Alicanto® mass spectrometry workflow. Briefly, the fractions are run on SDS-PAGE, the HCAb band was digested in gel with multiple proteases, the digested bands were then analyzed on a nano-HPLC coupled to a ThermoFisher Fusion Lumos Tribrid. In total 110,911 tandem mass spectra were generated for the aCD20 fraction, and 139,795 mass spectra were generated for the aKLH fraction.

Data analysis

MS/MS spectra from each fraction were independently mapped to the IgG2 repertoire. Alicanto® determined the set of CDR3 sequences present in the repertoire (shown in Figure 2). Four CDR3s that were clustered into two related clone clusters were found targeting the CD20 peptide. The limited diversity is not unexpected due to the limited number of epitopes present on the peptide and the general observation that llama prefer to raise IgG1 antibodies over HCAbs to linear epitopes (*8*). From the aKLH fraction, 65 distinct CDR3s in 20 clone clusters were found. None of the aKLH CDR3s overlapped with the aCD20 CDR3s.

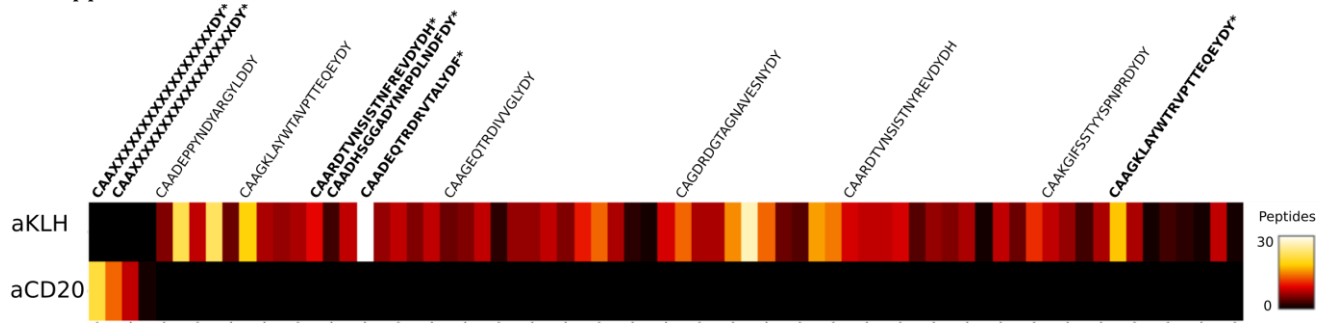

Figure 2: A heatmap of detected CDR3 sequences across the aKLH and aCD20 fractions. The intensity of the cell is determined by the number of peptides found overlapping each CDR3. The labeled CDR3s were expressed and tested via ELISA. The CDRs with an asterisk (*) were positive binders to their intended targets and negative binders to the tested off-target.

Candidate validation

We validated select VHHs by verifying binding specificity to the desired antigen using ELISA. Among the four aCD20 CDR3s, we selected two for production and added a third VHH CDR3 (not shown in heatmap) for validation. Among the 65 aKLH CDR3s, we selected 10 candidates for expression, with the goal of broadly sampling the 20 CDR3 families. The selected candidates are labeled on the clone heatmap above.

The three VHHs from the aCD20 fraction and ten VHHs from the aKLH fraction were synthesized into a pET28a(+) vector and expressed as VHHs in Shuffle Escherichia coli cells. The VHHs were purified and tested via ELISA for binding to the CD20 peptide and KLH. The CDR3s that are bolded and have an asterisk (*) were positive binders to their intended targets and negative binders to the tested off-target. In total two of three aCD20 VHHs and four of ten aKLH VHHs showed affinity and specificity to their intended targets.

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).
7. C. Hamers-Casterman *et al.*, *Nature* **363**, 446–448 (June 1993).
8. E. Pardon *et al.*, *Nature protocols* **9**, 674–693 (2014).