Distinct repertoire expansion patterns during affinity maturation of antigen responding camelid-derived single-domain VHH antibody

Tomonari Matsuda[1]†, Yoko Akazawa-Ogawa[2]†, Lilian-Kaede Komaba[2], Norihiko Kiyose[3], Nobuo Miyazaki[3], Yusaku Mizuguchi[4], Tetsuo Fukuta[4], Yuji Ito[5], & Yoshihisa Hagihara[2]*

[1]Research Center for Environmental Quality Management, Kyoto University, 1-2 Yumiyama, Otsu, Shiga 520-0811

[2]Biomedical Research Institute, National Institute of Advanced Industrial Science and Technology (AIST), 1-8-31 Midorigaoka, Ikeda, Osaka 563-8577

[3]ARK Resource. Co., Ltd., 383-2 Nakahara-machi, Nishi-ku, Kumamoto 861-5271

[4]JSR Corporation, 25 Miyukigaoka, Tsukuba, Ibaraki 305-0841

[5]Graduate School of Science and Engineering, Kagoshima University, 1-21-35 Korimoto, Kagoshima 890-0065

*Corresponding author. Email: hagihara-kappael@aist.go.jp

†These two authors contributed equally to this work.

Abstract

It is difficult to link antibody repertoire expansion with changes in the physical characteristics of antigen-responding antibodies as correct light-chain and heavy-chain matching and conventional antibody production are laborious. Utilization of single-domain antibody solved these problems. A blood of immunized alpaca was collected weekly for three months for antibody repertoire analysis. The sequences processed to cluster and recombinant antibodies were generated to determine whether the clusters respond to antigens. The repertoire expansion in most of the antigen-responding clusters exhibited distinct patterns as compared to that of antigen irrelevant clusters. In addition, the sequences at the tips and root in the molecular phylogenetic cluster tree have strong and weak antigen affinity, respectively. These features may be utilized to predict clusters of antigen-binding antibodies and physical characteristics of antibodies.

One-Sentence Summary

The clusters of antigen responding single-domain antibodies exhibited distinct repertoire expansion patterns.

Antibodies accumulate somatic hypermutations and undergo affinity maturation upon exposure to antigens (*1*). Immunization exploits this mechanism to produce antibodies against the target antigens. Repetitive antigen injections introduce random mutations and increase the antigen-binding affinity of the antibodies. They also ensure the survival of B-cells expressing high-affinity antibody against the antigens. The history of the mutational changes that occur in antibodies during immunization directly reflects the enhancement of the adaptive humoral immune response. It also helps elucidate the process of antibody maturation in response to specific antigens.

High-throughput next-generation sequencing (NGS) of vast immune repertories provides useful information for immunological system research and its practical applications (*2, 3*). Unlike conventional sequencing techniques, NGS enable us to draw a comprehensive picture of immune repertories that respond to antigens. Immunization-induced antibody development can be precisely analyzed by performing NGS on samples collected during immunization. It also provides a time-resolved overview of antibody maturation.

Various empirical and informatics technologies have been developed to facilitate nucleotide sequencing. However, correct light-chain and heavy-chain matching of antibody obtained by high-throughput sequencing is labor-intensive. The production of full-length antibody may also require time-consuming mammalian cell culture. Small antibody formats such as single-chain F_v fragment (scF_v) and F_{ab} may be produced by bacterial hosts. However, this approach may result in aggregation, defective folding, and loss of activity. The V_H domain of camelid heavy-chain antibody (VHH) binds the antigen in single-domain format (*4, 5*) and thus is irrelevant to the problem of correct matching of light- and heavy-chains. In addition, VHH antibody can usually be produced rapidly, conveniently, and inexpensively by an *Escherichia coli* (*E. coli*) expression system (*6*). VHH is a suitable antibody format for the examination of many sequences and exploration of the physical effects of mutational changes induced by affinity maturation.

Here, we report the repertoire development of VHH antibody using pools of peripheral lymphocytes repeatedly collected from alpaca blood during specific antigen immunization. The sequences were clustered by length and similarity and analyzed for time-dependent mutational changes. We also compared the antigen binding activity of various VHHs emerging from the same clusters at different times.

Results

Alpaca polyclonal antibody immune responses against injected antigens

We immunized an alpaca against ranibizumab, F_{ab} from trastuzumab, and a human κ C_L fragment produced by *E. coli* (Fig. 1). The alpaca had already been immunized with the same adjuvants prior to this experiment. Blood samples were collected immediately before immunization and every week thereafter for 14 weeks. We checked the increases in purified immunoglobulin G (IgG) subclasses (IgG1, IgG2, and IgG3) titer and completed immunization after six antigen injections (Fig. 1). The VHH cDNAs were synthesized by reverse transcription of mRNA extracted from a pool of lymphocytes. Short- and long-hinge VHH-specific primers were used and the cDNAs were amplified for sequencing.

Clustering VHH sequences from immunized alpaca

After merging the overlaps, we obtained averages 169k and 161k full-length VHH sequences from IgG2 (short-hinge between VHH and C_{H1}) and IgG3 (long-hinge between VHH and C_{H1}), respectively (Table S1). The same sequences were gathered into "unique sequences" and cleaned of any sequence errors. We then obtained averages 3k and 25k unique sequences from IgG2 and IgG3 at each time point, respectively. The sequences were grouped by their germ-line V and J combinations and their lengths (Fig. S1). Here, we referred to a set of DNA sequences as a "group" that might consist of various antibody families. Information for the D region was not used for grouping as it was too short and introduced ambiguity in sequence matching. It was assumed that in most cases, the lengths of the sequences were the same for all members of each antibody family propagated from a single ancestral sequence to adapt a specific antigen.

After excluding lone sequences, DNA sequences within a group were divided into "clusters" by phylogenetic tree analysis. We hypothesized that the clusters had the same properties as antibody families, namely, sequences binding to the same antigens and sharing the same ancestors. However, it was difficult to conclude that each cluster covers entire sequences evolving from the initial sequence. Moreover, we could not rule out the possibility that each cluster contains sets of entirely different antibodies recognizing different antigens. The alpaca analyzed here had already been immunized with the same adjuvants prior to this experiment. Therefore, we discarded the clusters including the sequences expressed before immunization and assumed that the clusters interacting with the adjuvants were removed at this step. We finally obtained 321 clusters comprising 923 and 3,546 sequences derived from IgG2 and IgG3, respectively. Many of the identical sequences were observed in IgG2 and IgG3. Hence, the total number of unique sequences was, in fact, less than the sum of unique sequences derived from IgG2 and IgG3 (4,469).

Propagation of VHH sequences through immunization

We focused on the 16 predominant clusters to elucidate their characteristics. They showed the highest maximum percentage appearance. The percentage appearance is a sum of the percentage occupancy of the IgG2 and IgG3 sequences in the cluster compared with all IgG2 and IgG3 sequences in each week. The maximum percentage appearance is the largest percentage appearance of the cluster during immunization. The sequences for clusters 7 and 15 were identified by bio-panning of the M13 phage library of the cDNA of VHHs collected at week 9. We examined

three clusters identified by bio-panning (clusters 69, 99 and 210) and one cluster identified by comparing sequence propensities before and after bio-panning (cluster 33) (*7, 8*). We then prepared VHH proteins for the 20 clusters (Table S2) and used enzyme-linked immunosorbent assay (ELISA) and surface resonance spectroscopy (SPR) to assess their affinity against the antigens (Fig. 2).

Only the VHH clones in clusters 2, 5, 7, 10, 15, and 16 exhibited antigen binding ("hit clusters"). No antigen binding was detected for the clones in clusters 1, 3, 6, 8, 9, 11, 12, 13, or 14 ("miss clusters"). In SPR, clone S38 in cluster 4 bound aberrantly to the C_L fragment. Thus, it could not be identified as an antigen binding clone and cluster 4 was excluded from further analysis. The clones had the same nomenclature as the sequence ID. The first S and L indicate sequences derived from short-hinge antibody (IgG2) and long-hinge antibody (IgG3), respectively. The numbers following the S and L of the sequence ID refer to the descending order of maximum frequency of the sequence. According to both ELISA and SPR, clones S11 (cluster 2), L926 (cluster 15), L792 (cluster 16), and L252126 (cluster 99) exhibited affinity against ranibizumab and F_{ab} from trastuzumab. Thus, the epitopes of those clones were the framework region of the F_v or C_{H1} domain. Clones S1139 (cluster 7) and L54 (cluster 33) had profiles similar to those of the aforementioned clones. At high VHH concentrations, however, the ELISA signal against C_L domain made it difficult to determine the epitope location. The epitope of clone L2477 (cluster 69) occurred on the C_L domain as the ELISA and SPR signals were positive against the C_L fragment and F_{ab}s. Clone L19 (cluster 10) only produced an ELISA signal against F_{ab}s. Its epitope may have been buried in the antigen and exposed by the interaction between the antigen and the plastic surface, of which interaction can induce protein unfolding. There were distinct ELISA signals in clones L38 (cluster 5) and L15235 (cluster 210) against all antigens. However, the SPR signal detected was against the C_L fragment alone. We hypothesized that these epitopes occurred on the C_L and C_{H1} binding surfaces and were accessible from the outside via the protein unfolding induced by the plastic surface.

To visualize cluster propagation in each independent sequence, we examined the transition and timing of sequence appearance/disappearance in the hit-clusters and miss-clusters (Fig. 3A & B). The sequence transition in the cluster was evaluated with the bit score parameter in the Basic Local Alignment Search Tool (BLAST) to plot the distance between the V gene region of each sequence and ancestral germinal V gene sequence. The bit score value increases with similarity of the query sequences to the reference. All hit-clusters had negative slopes of bit score vs. time of sequence appearance. Thus, the sequences in the clusters interacting with the antigen continually changed and became more remote from the ancestral sequence during immunization. In contrast, five of miss-clusters (1, 3, 8, 11, and 14) showed positive slopes of bit score vs. time of sequence appearance. Four others (6, 9, 12, and 13) exhibited slightly negative slopes of bit score vs. time of sequence appearance (Fig. 3B). Sequence evolution was indicated by negative bit score slopes and was a good indicator of clusters including clones binding the antigen.

Most of the hit-clusters displayed continuous sequence turnover. This behavior may be judged by the change in symbol color indicating the timing of first appearance of each sequence in Fig. 3. This effect was clearly seen in clusters 7, 10, 15, and 16. Cluster 2 disappeared at the late stage of immunization despite the fact that it predominated at the early stage. In cluster 5, an early

sequence occurred at week 1, reappeared at weeks 7 and 9, and then vanished (yellowish triangles). After 12 weeks, the sequences that had predominated between weeks 7 and 10 remained. However, new sequences also emerged. Nevertheless, sequence turnover was not obvious for any miss-cluster except 9 and 11. In those cases, new sequences replaced old ones especially after week 12. This behavior resembled that of hit-cluster 5. Distinct sequence turnover may be a hallmark of clusters affected by the immune response. However, this feature was difficult to identify in certain clusters such as 5, 9, and 11.

The value of the Y-axis intercept at the onset of the experiment (initial bit score) may be relevant to immune response progression against a new antigen. For the hit-clusters and miss-clusters, the initial bit scores were 446 ± 90 and 319 ± 58, respectively. The alpaca had not been exposed to the antigens used in this experiment before the onset of the study. Therefore, the hit-cluster sequences should not have been optimized before immunization and have resembled the ancestral germinal V gene sequences. In contrast, clusters with low initial bit scores were unresponsive to the immunized antigens because they might have consisted of sequences that had already matured before immunization.

Four clusters including the empirically identified clones were analyzed by the same plot (Fig. 3C). Negative bit score slopes, sequence turnover, and high initial bit scores were clearly observed in all cases. These findings underscore the relationship between hit-clusters and the three aforementioned criteria.

We searched for clusters matching these criteria from those with low maximum appearances and manually selected clusters 93, 103, 126, 139, 143, 175, 245, and 275 (maximum appearances were 0.2–0.03%). Their bit score slopes were negative and their initial bit scores were very high (> 400) (Fig. 4 and S2). Newly appearing sequences predominated weekly and there was sequence turnover. For these reasons, the aforementioned clusters were putative hit-clusters. The sequences at the tips of the phylogenetic trees constructed for these clusters were prepared as antibody proteins (Table S2) and antigen binding was tested by ELISA and SRP. The VHH clones of clusters 103, 126, 139, 143, 175, 245, and 275 bound the antigen. According to ELISA and SPR, there was no antigen interaction with the clones from cluster 93.

Comparison of VHH clones in the same cluster

To elucidate the effects of accumulating somatic hypermutation during immunization, we analyzed the antigen binding affinity of the clones in cluster 7. As clusters 33 and 120 had nearly identical sequences (one amino acid length difference), they were combined as cluster 33+120 and their clones were also examined. We selected various sequences appearing at different times and positioned at different branches in the phylogenetic tree and produced them as antibody proteins. Fourteen and six sequences (Table S2) were examined for clusters 7 and 33+120, respectively (Fig. 5). Antibody-antigen interactions were evaluated by SPR using trastuzumab F_{ab}. As immunization progressed, the sequences increasingly differed from the root sequence. The latter was closest to the ancestral germline V sequence and exhibited the lowest bit score in the cluster. Antigen-binding affinity increased with the number of mutations. After the last antigen injection at week 10, there was wide variation in clone affinity.

In cluster 7, clone S8492 first appeared at week 3 followed by clone S1139 (L27626; experimentally identified by phage display). Clone S506 first appeared at week 5. These clones were extinct after week 10. Clones L1511 (S6937) and L15236 appeared after 6 and 7 weeks, respectively. S709 (L25474), L2487 (S2363), S6584, S768, L2796 (S10289), S3320 (L9619), S537, L2754, and S675 (L2848) appeared after week 10.

Affinity maturation is driven by both association and dissociation kinetics but the predominant process may vary with each case. S1139 (L27626), S506, L1511 (S6937), and L15236 rapidly dissociated and most of VHH bound to the sensorchip surface disappeared after 12 min of buffering. S8492 appeared first and displayed slower binding and dissociation than these four clones. Eight of nine clones appearing after week 10 (S709 (L25474), L2487 (S2363), S6584, S768, S3320 (L9619), S537, L2754, and S675 (L2848)) dissociated far more slowly than the early clones. L2796 (S10289) appeared at weeks 13 and 14 and bound and dissociated faster than S768. However, both clones had similar K_D values. The K_D of the clones appearing after week 10 were in the range of 10^{-7}–10^{-10}. There may be a certain degree of tolerance for affinity against antigen and there may be no strict affinity cutoff even after substantial immunization progress.

In cluster 33+120, the empirically identified clone L54 appeared between weeks 5 and 9 and then disappeared. L1473 appeared at weeks 6 and 7 and then disappeared. Both clones were characterized by rapid binding and dissociation. Clones L109, L345 L9139, and L13326 appeared later and had fast association but significantly slower dissociation than the early clones. Hence the later clones had stronger antigen binding affinity than the early clones. Unlike the later clones in cluster 7, the clones in cluster 33+120 displayed rapid binding and dissociation.

Discussion

With certain exceptions, repertoire expansion of the sequences in immune-responsive antibody clusters has distinct patterns. Sequences continuously develop and accumulate diversity throughout immunization lasting longer than ten weeks. The alpaca immune system can apparently recognize the potency of many sequences and clusters. Here, humoral immunity was robust when antibodies were formed against the antigens used. Furthermore, the sequences showed intensive turnover. In the hit-clusters, the older sequences became extinct and were superseded by newly emerged sequences. It has been unclear whether affinity maturation during immunization causes antibody sequences to converge. It is possible that sequence divergence increased at the early stage and a few appropriate sequences were selected at the late stage. However, the present study showed that the sequences in the hit-clusters kept changing during immunization and continued to do so even four weeks after the last antigen injection. Prolonged sequence turnover suggests that antibody affinity maturation was ongoing and did not converge hit cluster sequences within our experimental period (14 weeks).

The bit score plot is an excellent tool to identify hit-clusters. It contains information such as the frequency of appearance, timing of the first appearance, and bit score of each sequence within the same cluster. Typical patterns were observed for clusters 7, 15, and 16 (Fig. 3A). These clusters had high maximum bit scores. Thus, they started from the sequence nearest the germ line sequence with little or no affinity maturation. There was prominent alternation of sequence generation and decreasing bit score over time. Hence, affinity maturation progression occurred. Based on the bit score slopes, sequence turnover, and initial bit scores, we selected eight clusters with relatively low abundance. Of these, seven included clones binding immunized antigen. Thus, they were designated hit-clusters. Although the overall ratio of hit-clusters to all clusters is unknown, the ratio for the top 16 clusters (except cluster 4) was 6/15. The immune response increases antigen-responding antibody expression. Therefore, the hit-cluster ratio may increase with cluster appearance frequency. For these reasons, the proportion of 7/8 of the hit-clusters with relatively low frequency was significant in hit-cluster prediction.

When a hit-cluster was identified by bio-panning or bit score plot analysis, the VHH clones with high affinity against the target antigen could be selected using the phylogenetic tree of the identified cluster. The negative bit score slope predicted that the sequences at the tips of phylogenetic tree appeared at the late stages of immunization. Our data suggested that the binding affinity increased with the delay in clone appearance. Therefore, the clones at the tips of the phylogenetic tree have relatively stronger antigen binding ability. This conclusion is consistent with the basic concept of affinity maturation. Clone S1139 (L27626) in cluster 7 and clone L54 in cluster 33 were identified either by bio-panning or by a combination of bio-panning and NGS analysis (*7, 8*). Both empirically identified clones had low antigen affinity and were localized to the root region of the phylogenetic tree. However, the clones at the tips of the phylogenetic tree exhibited relatively higher antigen affinity than those identified empirically. Nevertheless, not all experimental clones necessarily have low antigen affinity. Turner *et al* (*9*) showed that bio-panning-derived clones within the same clusters shared similar antigen affinity. The bit scores for the other empirically identified clones in clusters 15, 69, 99, and 210 were 366, 470, 483, and 353, respectively, where their largest bit scores were 433, 486, 499, and 387. These empirically identified clones were remote from the root of the phylogenetic tree.

In the present study, we showed that VHH sequences responding to injected antigen were continuously generated and superseded throughout the 3-month course of immunization. Clone development was associated with increasing affinity against the target antigen. These findings were expected and consistent with the concept of immune affinity maturation. To the best of our knowledge, this work is the first to clarify the precise antibody affinity maturation process and empirically assess various antibodies appearing at different times. Based on the foregoing results, it would be feasible to predict VHH hit-clusters and high-affinity VHH clones within those clusters based on their bit score plots and their positions on the phylogenetic tree. A combination of bio-panning and NGS analysis successfully identified binding VHH clones not detected by conventional bio-panning (*7, 9, 10*). In stark contrast to previous research, the prediction method proposed here requires no bio-panning. In this study, we utilized the immune system of alpaca single-domain VHH to avoid the technical difficulties of producing conventional antibody and its fragment and correctly matching light-chains and heavy-chains. However, it is unknown whether the findings of this work are applicable to conventional antibody and immune systems of other animals. Hence, future research should be conducted using different camels and antigens as well as other animal species to validate the mechanism of affinity maturation in humoral immunity.

References

1. C. Berek, C. Milstein, Mutation drift and repertoire shift in the maturation of the immune response. *Immunol Rev* **96**, 23-41 (1987).
2. C. Parola, D. Neumeier, S. T. Reddy, Integrating high-throughput screening and sequencing for monoclonal antibody discovery and engineering. *Immunology* **153**, 31-41 (2018).
3. G. Georgiou *et al.*, The promise and challenge of high-throughput sequencing of the antibody repertoire. *Nat Biotechnol* **32**, 158-168 (2014).
4. C. Hamers-Casterman *et al.*, Naturally occurring antibodies devoid of light chains. *Nature* **363**, 446-448 (1993).
5. S. Muyldermans, A guide to: generation and design of nanobodies. *FEBS J*, (2020).
6. Y. Liu, H. Huang, Expression of single-domain antibody in different systems. *Appl Microbiol Biotechnol* **102**, 539-551 (2018).
7. N. Miyazaki *et al.*, Isolation and characterization of antigen-specific alpaca (Lama pacos) VHH antibodies by biopanning followed by high-throughput sequencing. *J Biochem* **158**, 205-215 (2015).
8. A. Rafique *et al.*, Efficient screening and design of variable domain of heavy chain antibody ligands through high throughput sequencing for affinity chromatography to purify fab fragments. *Monoclonal Antibodies in Immunodiagnosis and Immunotherapy* **38**, 190-200 (2019).
9. K. B. Turner *et al.*, Next-generation sequencing of a single domain antibody repertoire reveals quality of phage display selected candidates. *PLoS One* **11**, e0149393 (2016).
10. P. Deschaght *et al.*, Large diversity of functional nanobodies from a camelid immune library revealed by an alternative analysis of next-generation sequencing data. *Frontiers in Immunology* **8**, 420 (2017).
11. T. H. Jukes, C. R. Cantor, in *Mammalian Protein Metabolism,* H. N. Munro, Ed. (Academic Press, 1969), pp. 21-132.
12. D. S. Doering, P. Matsudaira, Cysteine scanning mutagenesis at 40 of 76 positions in villin headpiece maps the F-actin binding site and structural features of the domain. *Biochemistry* **35**, 12677-12685 (1996).
13. Y. Hagihara, T. Matsuda, N. Yumoto, Cellular quality control screening to identify amino acid pairs for substituting the disulfide bonds in immunoglobulin fold domains. *J Biol Chem* **280**, 24752-24758 (2005).
14. S. F. Altschul, W. Gish, W. Miller, E. W. Myers, D. J. Lipman, Basic local alignment search tool. *J Mol Biol* **215**, 403-410 (1990).
15. H. Edelhoch, Spectroscopic determination of tryptophan and tyrosine in proteins. *Biochemistry* **6**, 1948-1954. (1967).

Acknowledgments

Funding:

JSPS KAKENHI grant 20K07009 (YAO)

JSPS KAKENHI grant 20K20599 (TM)

JSR Corporation (TM, YAO, NM, YI, YH)

Author contributions:

Examples: TM, YAO, LKK, NK, NM, YM, TF, YI, YH

Conceptualization: TM, YAO, NM, YI, YM, TF, YH

Formal analysis: TM

Methodology: TM, YAO, YH

Investigation: TM, YAO, LKK, NK, NM, YM, TF, YI, YH

Visualization: TM, YAO, NK, NM, YH

Funding acquisition: TM, YAO, NM, YI, YH

Project administration: TM, YAO, YH

Writing – original draft: TM, YAO, NK, NM, YH

Writing – review & editing: TM, YAO, NK, NM, YM, TF, YI, YH

Competing interests:

YM and TF of JSR Corporation participated in this research by conceptualization and investigation. No other funding sources participated in study design, data collection, analysis, or interpretation, or report writing.

Kyoto Univ., AIST, ARK Resource. Co. Ltd. and Kagoshima Univ. (TM, YAO, NM, TF, YI, and YH are listed as inventors) applied for patent (JP/2019/080434). Kyoto Univ. (TM, YAO, NM, TF, YI, and YH are listed as inventors) applied for patent (WO/2020/213730).

Data and materials availability:

Antibody nucleotide sequence data reported are available in the DDBJ Sequenced Read Archive under the accession numbers DRR305325-DRR305354. SAMD00387596-00387610 correspond to the data from blood collected at week 0-14. All other data are available in the main text or the supplementary materials. All data used in the analysis are available for purposes of reproducing or extending the analysis. The materials used in this work are available under materials transfer agreements (MTAs).

Supplementary Materials

Materials and Methods

Figs. S1 to S3

Tables S1 to S2

Fig. 1. Immunization and blood collection schedules (A) and time course of purified polyclonal alpaca antibody titer (B). (**A**) F_{ab} of trastuzumab, ranibizumab, and human κ C_L fragment were injected as antigens into a single alpaca. Blood was collected once before immunization (week 0) and 14 times after initiating immunization (weeks 1–14). Library for phage display screening was prepared using blood collected at week 9. (**B**) Polyclonal IgG1 (conventional antibody), IgG2 (heavy-chain antibody with short hinge), and IgG3 (heavy-chain antibody with long hinge) were purified from blood collected at weeks 4, 9, 11, and 13. Titers were measured against F_{ab} of trastuzumab, ranibizumab, and human κ C_L fragment.

Fig. 2. Antigen-binding affinities of clones in top 16 clusters and empirically identified clones. Cluster number indicates order of maximum percentage of appearances. The percentage of appearance is the sum of the percentage occupancy of IgG2 and IgG3 sequences in cluster relative to all IgG2 and IgG3 sequences in each week. Maximum percentage of appearance is the highest percentage of appearance of cluster during immunization. Clusters 7, 15, 33, 69, 99, and 210 included empirically identified antigen-binding sequences. Antigen-binding affinities of clones were evaluated by ELISA (left three panels) and SPR (right three panels) against immobilized human κ C_L fragment (left), F_{ab} of trastuzumab (middle), and ranibizumab (right). Values inside SPR panels indicate concentration ranges for VHH clones measured as analytes.

Fig. 3. Plot of sequence bit scores vs. time of appearance of sequences in hit-clusters (A), miss-clusters (B), and empirically identified clusters (C). Symbol colors indicate times of first appearance of each sequence. Colors corresponding to times of first appearance are indicated in color bar at panel bottoms. Circles and triangles indicate sequences observed in short-hinge (IgG2) and long-hinge (IgG3) antibody, respectively.

Fig. 4. Clusters with relatively low maximum percentage of appearances. Clusters 93 (**A**) and 103 (**B**) were selected based on negative bit score slopes, distinct sequence turnover, and high initial bit score as shown by bit score plot (upper left panels). Maximum percentage of appearances of clusters 93 and 103 were 0.23 and 0.20, respectively. Position of selected VHH clone in phylogenetic tree is indicated by red circle (upper right panel). Phylogenetic trees were drawn using DNA sequences including cluster. Bars below phylogenetic trees indicate distance = 0.01 calculated by JC69 (*11*) and corresponding to ~1% nucleotide sequence difference. Antigen-binding affinity of VHH clone against immobilized human κ C_L fragment (left), F_{ab} of trastuzumab (middle), and ranibizumab (right) are shown by ELISA (lower left three panels) and SPR (lower right three panels). Values inside SPR panels indicate concentration ranges for VHH clones measured as analytes.

Fig. 5. Analysis of VHH clones in same cluster. **A & B**, Phylogenetic tree of clusters 7 (**A**) and 33+120 (**B**). S1139 (L27626) in cluster 7 and L54 in cluster 33 are empirically identified clones. **C:** Relationship between bit score V and dissociation constant (K_D) against F_{ab} of trastuzumab. Within same cluster, bit score value decreases with increasing binding strength between VHH and antigen. Closed and open squares indicate clones from clusters 7 and 54, respectively. **D & E:** Time of appearance-dependent K_D change vs. F_{ab} of trastuzumab in clusters 7 (D) and 33+120 (E) . Circle and triangle indicate clones derived from short-hinge (IgG2) and long-hinge (IgG3)

antibody, respectively. K_D decreased over time; however, there was a range of K_D values, especially for later clones from cluster 7.

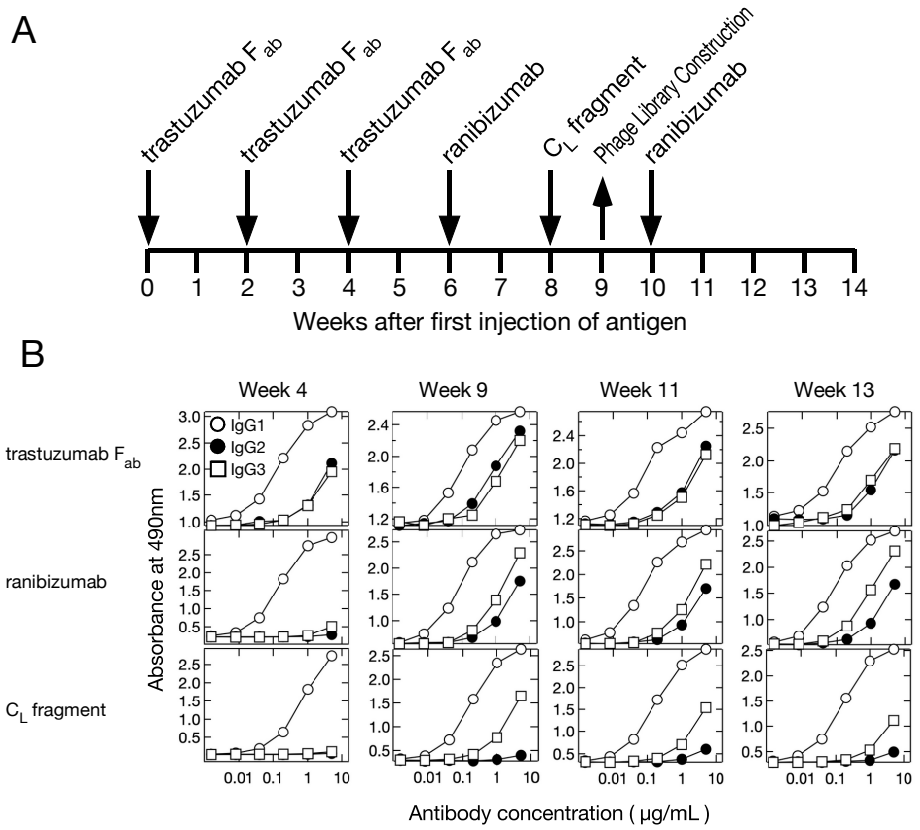

A
trastuzumab F ab
trastuzumab F ab
trastuzumab F ab
ranibizumab
C L fragment
Phage Library Construction
ranibizumab
0 1 2 3 4 5 6 7 8 9 10 11 12 13 14
Weeks after first injection of antigen
B
Week 4
Week 9
Week 11
Week 13
IgG1
IgG2
IgG3
trastuzumab F ab
ranibizumab
C L fragment
Absorbance at 490nm
0.01 0.1 1 10
Antibody concentration (μg/mL)

Fig. 1

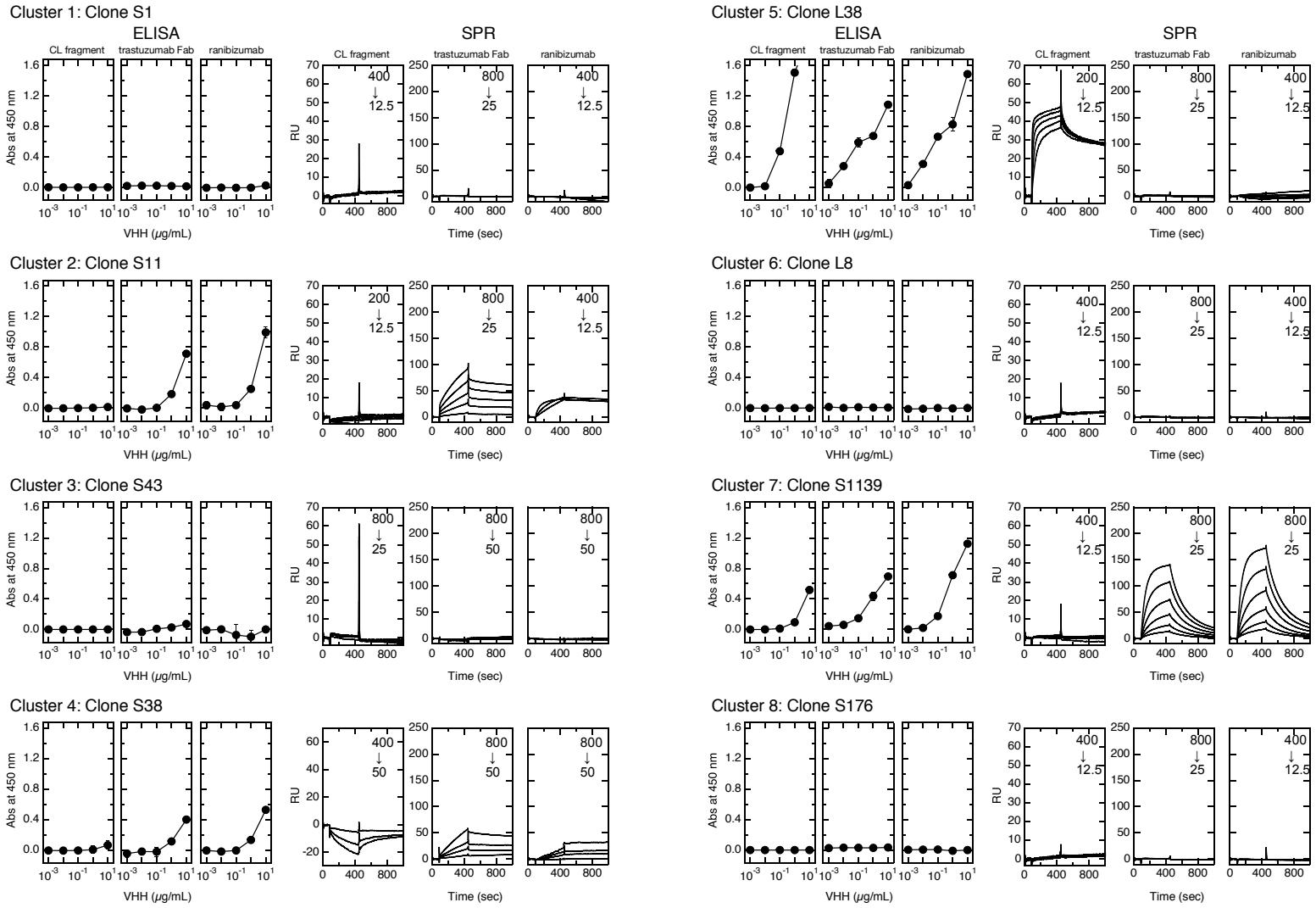

Cluster 1: Clone S1
ELISA
SPR
CL fragment
trastuzumab Fab
ranibizumab
Abs at 450 nm
VHH (μg/mL)
RU
Time (sec)
400 ↓ 12.5
800 ↓ 25
400 ↓ 12.5
Cluster 2: Clone S11
200 ↓ 12.5
800 ↓ 25
400 ↓ 12.5
Cluster 3: Clone S43
800 ↓ 25
800 ↓ 50
800 ↓ 50
Cluster 4: Clone S38
400 ↓ 50
800 ↓ 50
800 ↓ 50
Cluster 5: Clone L38
200 ↓ 12.5
800 ↓ 25
400 ↓ 12.5
Cluster 6: Clone L8
400 ↓ 12.5
800 ↓ 25
400 ↓ 12.5
Cluster 7: Clone S1139
400 ↓ 12.5
800 ↓ 25
800 ↓ 25
Cluster 8: Clone S176
400 ↓ 12.5
800 ↓ 25
400 ↓ 12.5

Fig. 2

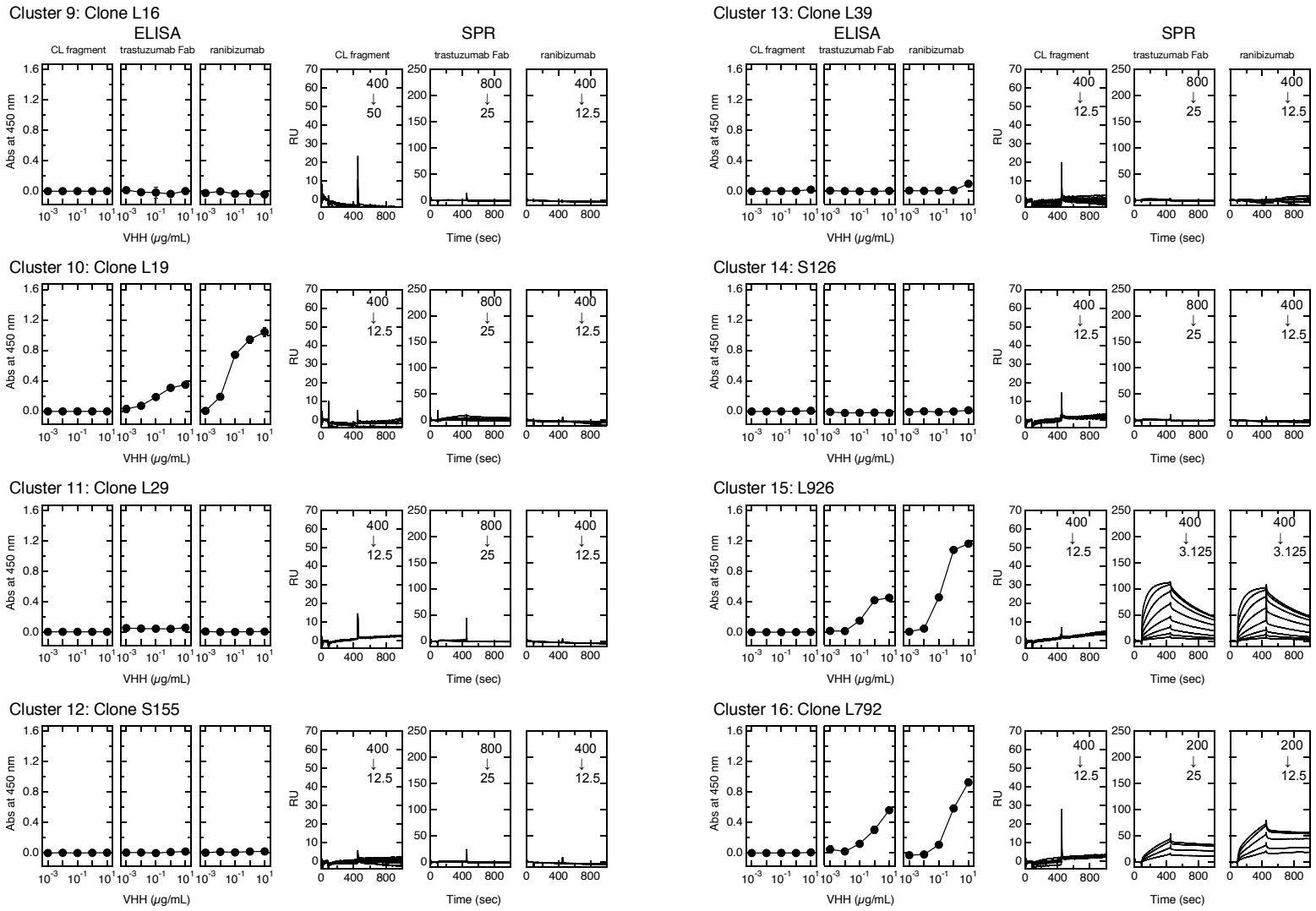

Cluster 9: Clone L16
ELISA
SPR
CL fragment
trastuzumab Fab
ranibizumab
Abs at 450 nm
RU
VHH (μg/mL)
Time (sec)
400 ↓ 50
800 ↓ 25
400 ↓ 12.5
Cluster 10: Clone L19
400 ↓ 12.5
800 ↓ 25
400 ↓ 12.5
Cluster 11: Clone L29
400 ↓ 12.5
800 ↓ 25
400 ↓ 12.5
Cluster 12: Clone S155
400 ↓ 12.5
800 ↓ 25
400 ↓ 12.5
Cluster 13: Clone L39
400 ↓ 12.5
800 ↓ 25
400 ↓ 12.5
Cluster 14: S126
400 ↓ 12.5
800 ↓ 25
400 ↓ 12.5
Cluster 15: L926
400 ↓ 12.5
400 ↓ 3.125
400 ↓ 3.125
Cluster 16: Clone L792
400 ↓ 12.5
200 ↓ 25
200 ↓ 12.5

Fig. 2

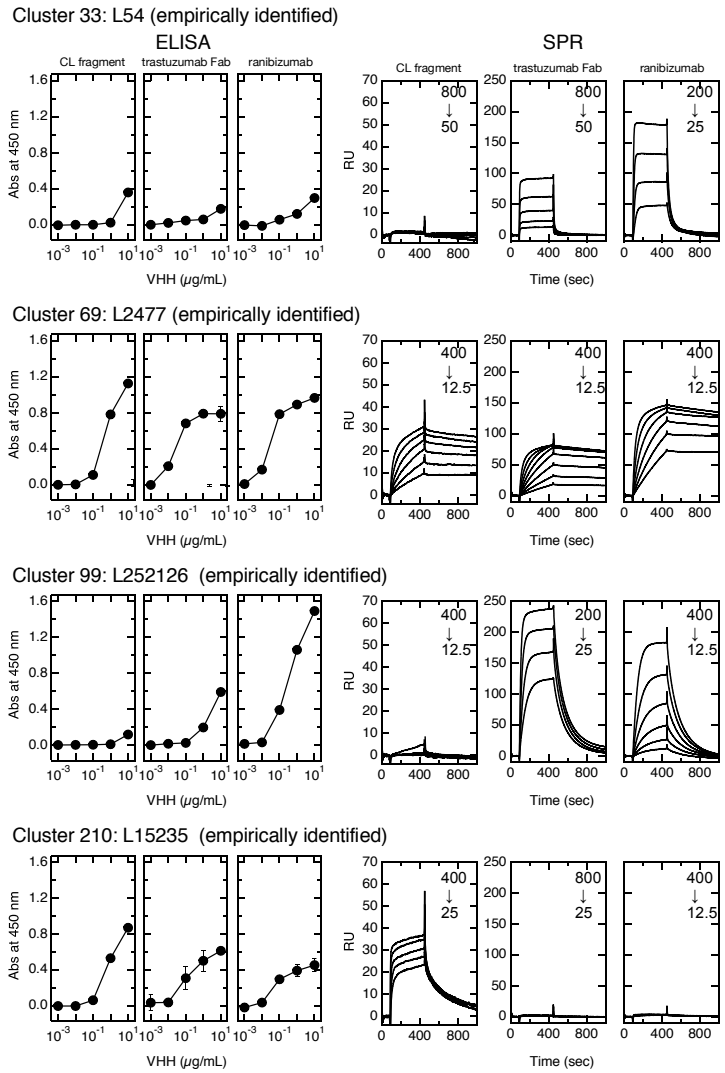

Cluster 33: L54 (empirically identified)
ELISA
SPR
CL fragment
trastuzumab Fab
ranibizumab
Abs at 450 nm
RU
VHH (μg/mL)
Time (sec)
800 ↓ 50
800 ↓ 50
200 ↓ 25
Cluster 69: L2477 (empirically identified)
400 ↓ 12.5
400 ↓ 12.5
400 ↓ 12.5
Cluster 99: L252126 (empirically identified)
400 ↓ 12.5
200 ↓ 25
400 ↓ 12.5
Cluster 210: L15235 (empirically identified)
400 ↓ 25
800 ↓ 25
400 ↓ 12.5

Fig. 2

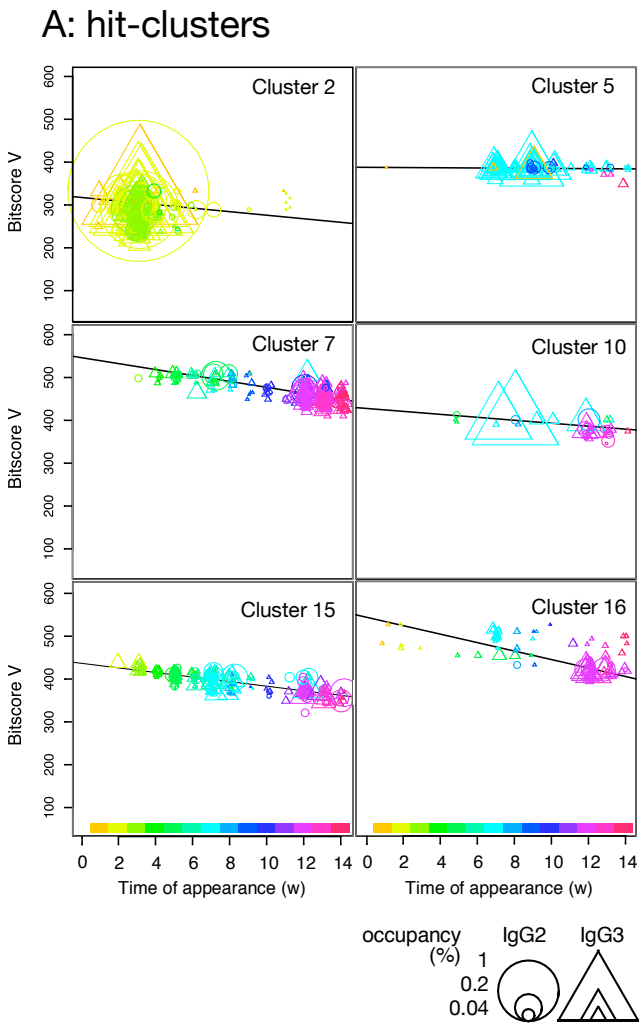
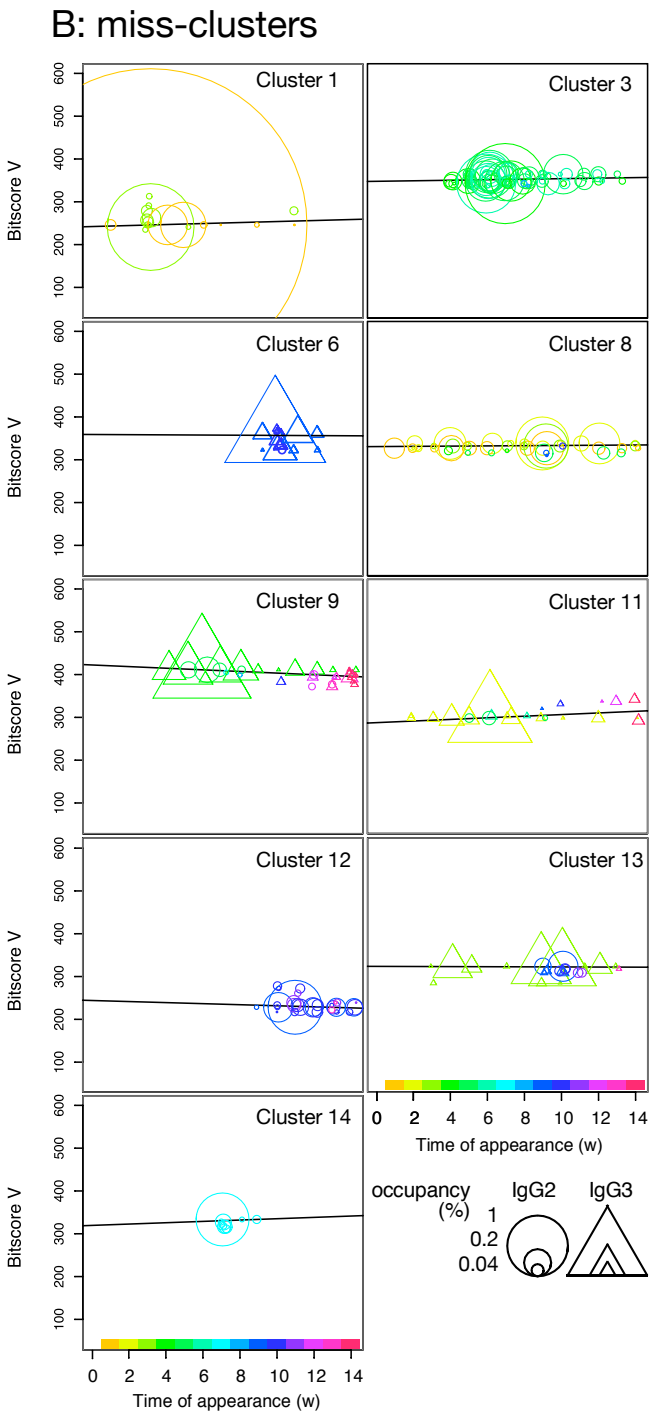

A: hit-clusters
Cluster 2
Cluster 5
Cluster 7
Cluster 10
Cluster 15
Cluster 16
Bitscore V
Time of appearance (w)
occupancy (%)
IgG2
IgG3
1
0.2
0.04

Fig. 3A


B: miss-clusters
Cluster 1
Cluster 3
Cluster 6
Cluster 8
Cluster 9
Cluster 11
Cluster 12
Cluster 13
Cluster 14
Bitscore V
Time of appearance (w)
occupancy (%)
1
0.2
0.04
IgG2
IgG3

5

Fig. 3B

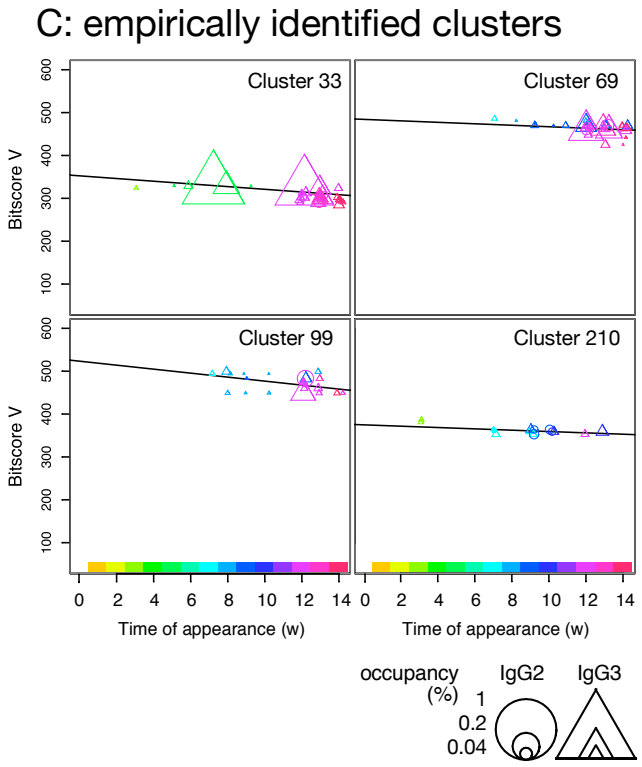

C: empirically identified clusters
Cluster 33
Cluster 69
Cluster 99
Cluster 210
Bitscore V
Time of appearance (w)
occupancy (%)
IgG2
IgG3
1
0.2
0.04

Fig. 3C

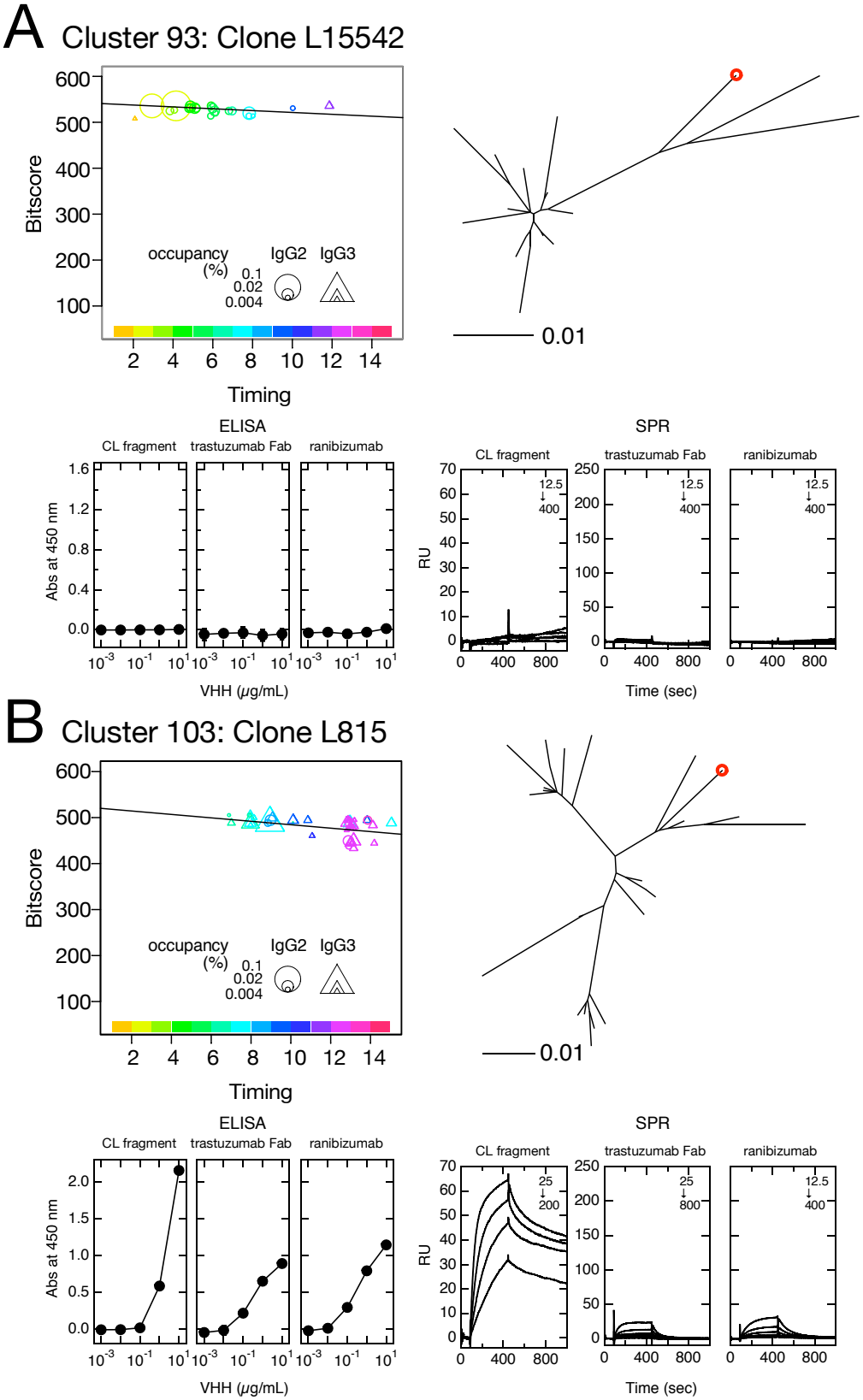
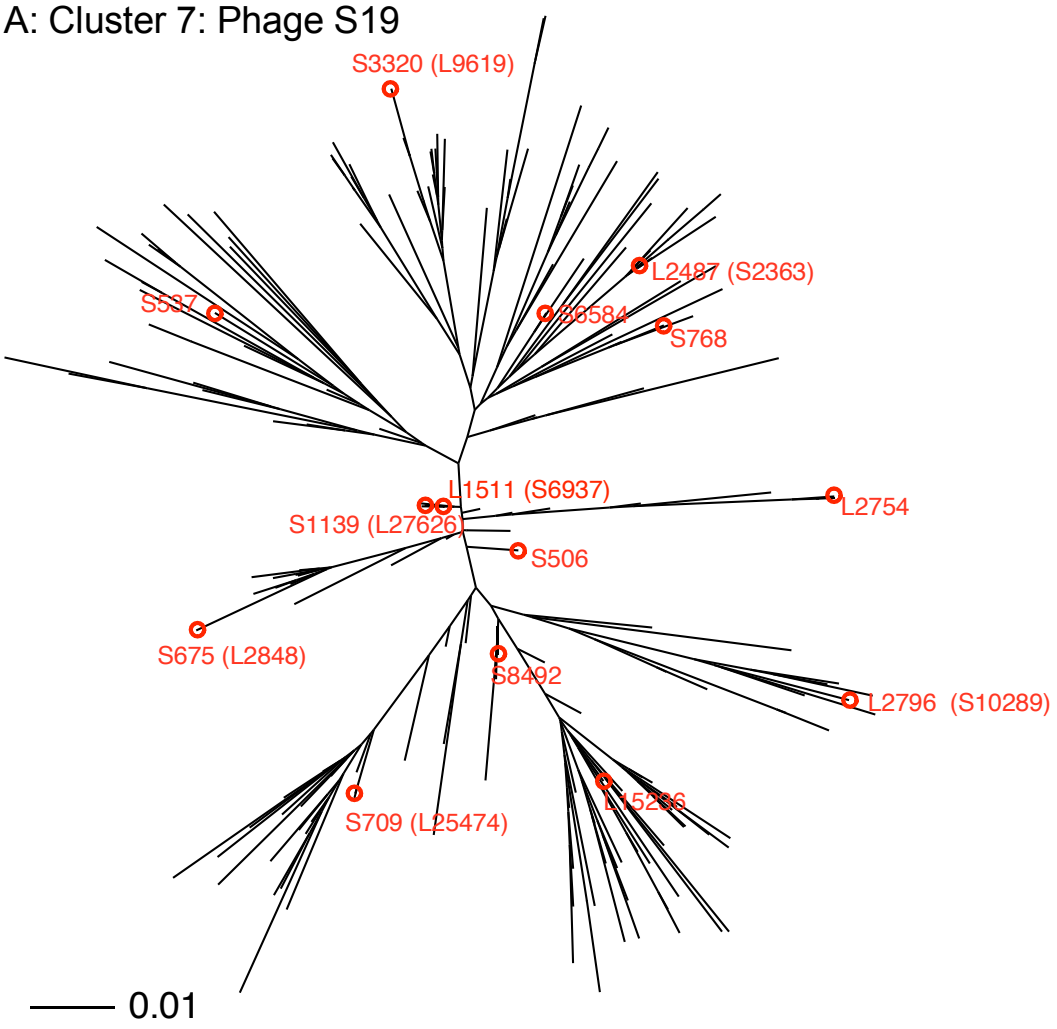

A
Cluster 93: Clone L15542
Bitscore
Timing
occupancy (%)
0.1
0.02
0.004
IgG2
IgG3
0.01
ELISA
CL fragment
trastuzumab Fab
ranibizumab
Abs at 450 nm
VHH (μg/mL)
SPR
RU
Time (sec)
B
Cluster 103: Clone L815

Fig. 4

A: Cluster 7: Phage S19



B: Cluster 33+120: Phage LH3

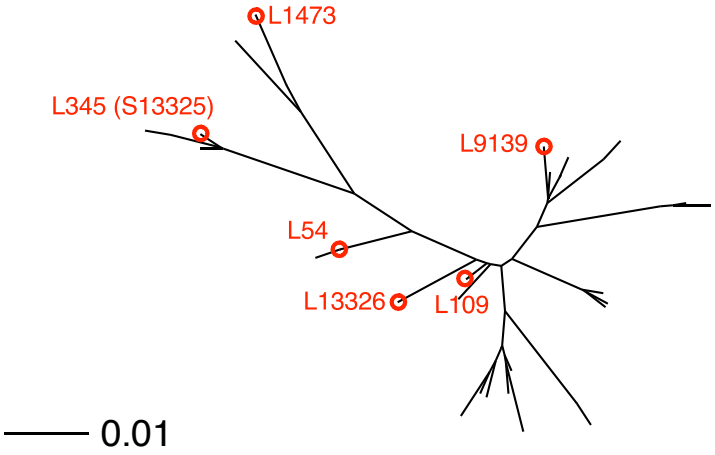


Fig. 5A & B

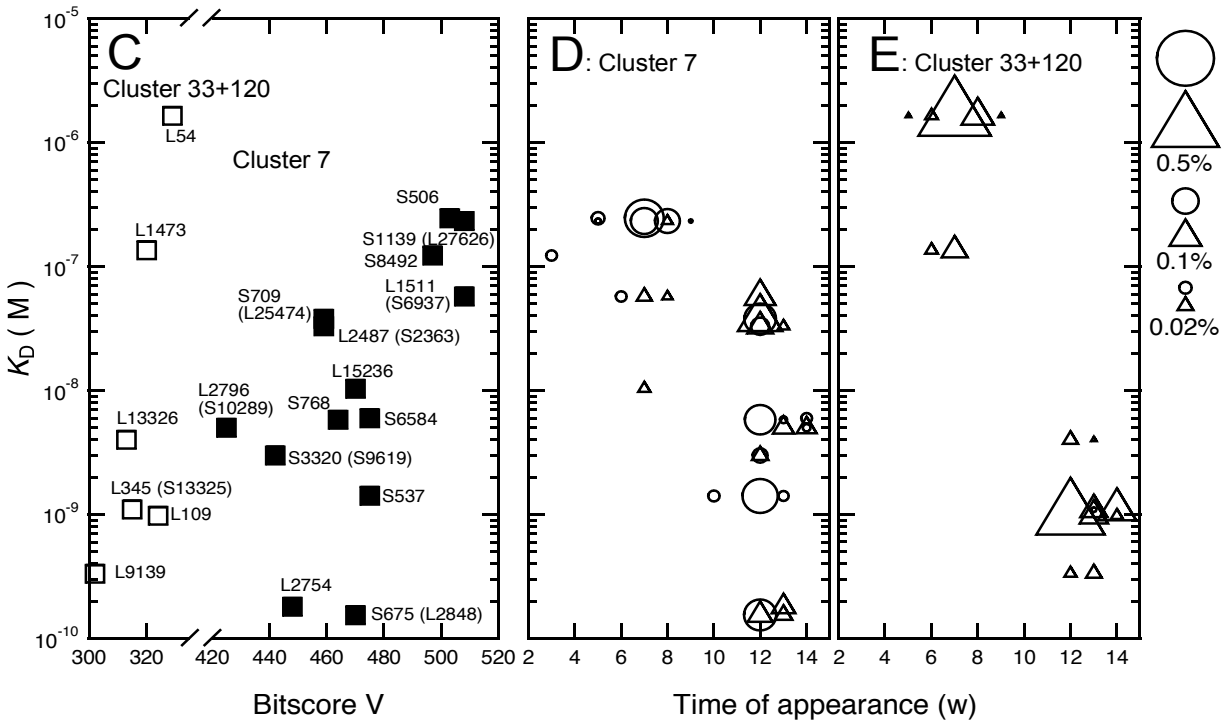

C
Cluster 33+120
L54
Cluster 7
S506
L1473
S1139 (L27626)
S8492
L1511
(S6937)
S709
(L25474)
L2487 (S2363)
L15236
L2796
(S10289)
S768
S6584
L13326
S3320 (S9619)
L345 (S13325)
S537
L109
L9139
L2754
S675 (L2848)
K_D (M)
Bitscore V
D: Cluster 7
E: Cluster 33+120
Time of appearance (w)
0.5%
0.1%
0.02%

Fig. 5 C, D & E