In vivo selection of nanobodies to target the tumor immune microenvironment for immunotherapy development

Thillai V. Sekar[1]*, Eslam Elghonaimy[1]*, Katy L. Swancutt[1], Michael Whitney[2], and Todd A. Aguilera[1$]

Affiliations:

[1] Department of Radiation Oncology, the University of Texas Southwestern Medical Center, Dallas, TX 75390, USA

[2] Department of Pharmacology, University of California San Diego, La Jolla, CA 92093, USA

* Equal contribution

[$] To whom correspondence should be addressed: todd.aguilera@utsouthwestern.edu

Abstract

Cancer immunotherapy has changed the landscape of treatment for some cancers, but variable responses reflect limitations imposed by the tumor microenvironment (TME)[1-3]. To facilitate the development of novel immunotherapies, we used an unbiased *in vivo* selection approach using a VHH nanobody library to identify immunologic targets in the TME. Using next generation sequencing (NGS), we demonstrate the parallel enrichment of immune cell-binding nanobodies that target immune cell subsets within the TME. We selected from enriched nanobodies – Cl.1 and Cl.2 – that bind to the $CD11b^+$ and $CD11c^+$ populations and we identified the desmosome complex as the target for Cl.1, suggesting the nanobodies can target key immune populations and may involve unvalidated targets. To demonstrate nanobodies could functionally alter target immune cells *in vivo,* we showed by single cell RNA sequencing (scRNAseq) that enriched nanobodies can increase activation transcripts in $CD8^+$ T cells and dendritic cells. This approach is being used to develop novel cancer immunotherapies, but the methods can be broadly applicable to biologic discovery experiments in various physiologic or pathologic contexts.

Results

In vivo selection with phage libraries first demonstrated organ targeting and that identified peptides could have therapeutic activity[4-6]. We later used this approach to select enzymatically cleaved peptides in the TME and to identify nerve-binding peptides for use in fluorescence-guided surgery[7,8]. The traditional approach to identify enriched peptides, antibodies, or proteins by phage display is to sequence clones identified by high dilution using Sanger sequencing, but with NGS, analysis can be deeper and more informative. In this study, we devised an *in vivo* biopanning strategy for an unbiased selection of nanobodies that bind to specific populations of tumor-infiltrating leucocyte (TIL) in triple negative breast cancer tumors **(Fig. 1A-B)**. The Py8119 and Py117 breast tumors have different immune responsiveness to therapy making them complementary for screening[9,10]. A camelid based VHH nanobody library was derived from the peripheral blood mononuclear cells of sixteen non-immunized llamas. VHH nanobodies are ideal because of their small size, single heavy chain, stability, modularity, low immunogenicity due to their homology to human immunoglobin VH3, and 90-130 amino acid length, which provides spatial coverage of unique epitopes[11-14]. With this approach we aimed to identify specific nanobodies for different immune cell populations then validate that enriched nanobodies could alter the immune response in the TME by single cell RNA sequencing (scRNAseq) (**Fig. 1C-D**)

To validate *in vivo* selection we first detected the full-length nanobody sequences with the key variable regions by NGS. We developed a computational pipeline for sequence analysis to detect unique phage clones and their relative abundance. With stringent methods we reconstructed the VHH sequences in spite of the 600 bp average length PCR amplicon because throughput limits the use of Sanger or long-read sequencing. As expected, we observed greater variability in the CDR3 than in the CDR1 and CDR2 (**Fig. 2A**). The length of the CDR3 region varied with a median of 18 amino acids and the distribtuion was consistent over all rounds of biopanning (**Fig. 2B and Sup. Fig 1A**). To highlight the diversity in the initial library, we assembled 20,083 VHH nanobodies from 405,069 reads and calculated

an average of one read per VHH, thereby reducing the concerns about PCR amplification artifacts in the library prior to *in vivo* selection (**Sup. Table 1 and Sup Fig 1B-C**).

We hypothesized that *in vivo* biopanning with NGS assessment could deeply evaluate the selection of a diverse nanobody library in a complex system that targets different immune cell subtypes in the TME. We injected the library of approximately $1.5x10^{12}$ phages intravenously, then dissociated the tumor and draining lymph node (dLN) tissue after two hours[8]. $CD45^+$ cells were separated, and bacteria were infected with phages from cell lysate. The sample was either tittered to single clones and identified by Sanger sequencing, or the pool of each cell subtype was sequenced by NGS with approximately 450,000 reads per sample (**Fig. 1A-B)**. Three tumor-bearing mouse replicates with Py8119 and Py117 tumors were used for each round of biopanning. We included the dLNs as a second tissue abundant with leucocytes that could help evaluate TME findings.

We aimed to identify parallel enrichment, meaning enrichment of nanobodies with different selectivity for each immune cell subtype could be selected simultaneoulsly by using the diverse $CD45^+$ population of reach round. Before the first biopanning (BP1), we sequentially selected bulk $CD45^+$ cells by ficoll gradient separation, then removed dead cells via magnetic beads to focus our highly diverse library towards leukocytes (BP0). For BP1-4 we isolated five different cell subtypes using a magnetic bead sorting scheme. A total of $4x10^6$- $6x10^7$ phages were isolated from each of five different $CD45^+$ cell subtypes from each round of biopanning (**Fig. 1B**). We selected for $CD11b^+$ myeloid cells, $CD11c^+$ dendritic cells from the $CD11b^-$ flow through, then T cells from the $CD11b^-$ and $CD11c^-$ flow through. For T cells, we selected $CD8^+$, $CD4^+CD25^+$ and $CD4^+CD25^-$ populations. Flow cytometry confirmed enrichment of each cell population from the magnetic beads sorting (**Sup. Fig. 2-3**).

After verifying that specific VHH clones could be identified by NGS and that cell subtypes could be sorted efficiently, we evaluated whether *in vivo* biopanning could enrich specific nanobodies that bind to

targets in the TME. We first assessed the number of VHHs and the VHH diversity for each cell sub-population, round of biopanning, tumor type, and tissue type (**Sup. Table 2)**. We combined all VHH clones from each cell sub-population for tumor and dLN to calculate the proportion of sequences each clone occupied for every round of biopanning. The proportion of occupied repertoire space was plotted for all clones based upon their rank of abundance, as top clone proportion, or by how many counts were identified for each clone, as rare clone proportion (**Fig. 2C**, left and right, respectively). Briefly, the library did not identify duplicate sequences for the same clone, the majority of sequences were identified 10 or fewer times in BP0, and in BP1-4 there was enrichment with multiple highly abundant clones over 100 copies (**Fig. 2C**). We also observed this pattern of enrichment in the dLN and cell subpopulation (**Sup. Fig. 4**). The repertoire space of the top ranked clones in each biopanning was expanded through biopanning rounds, and an average of 40% of the total repertoire of BP3-4 was occupied by 10 clones (**Fig. 2C** and **Sup. Fig. 5**).

To gain deeper perspective on the *in vivo* biopanning, we assessed the enrichment of VHHs through diversity extrapolation **Fig. 2D**. This approach considers the effect of the differences in sequence depth and the number of VHHs retrieved from each biopanning. The resulting rarefication curves showed an immediate diversity reduction after BP1, which is consistent with the representation in **Fig. 2C** for tumor and lymph node **(Fig. 2D and Sup. Fig. 6A)**. Diversity reduction was also confirmed by true diversity estimations (**Sup. Fig. 6B**). By focusing on the unique VHHs, we also observed a reduction in the percentage of unique VHHs retrieved with each round of biopanning (**Fig. 2E**). In summary, we observed enrichment after BP1, which was consistent across different tumors, tissues, and cell sub-populations. This suggests that multiple rounds of biopanning are not needed to achieve desirable enrichment, which was independent of tumor, tissue, and cell type.

The most abundant clones in BP0 each disappeared in subsequent rounds of biopanning, which suggests that nonselective amplification of overrepresented clones did not occur (**Fig. 2F, Sup. Fig.**

7A-C). The opposite was true for clones that showed specific enrichment. The most abundant clones at the end of the biopanning had increased their fractional proportion of all sequences, which supports that the greatest enrichment occurred between BP1 and BP2 (**Fig. 2G, Sup. Fig. 7D-F**). These observations suggest that enrichment is achieved through a biological selection process.

The traditional approach to identifying enriched peptides, antibodies, or proteins by phage display is to sequence clones identified at high dilution via Sanger sequencing, which captures the most abundant clones, such as a selected upper proportion depicted in **Fig. 3A**. We hypothesized that NGS would capture a broader distribution of clones that have biologic meaning and that a selective nanobody may not have the highest enrichment. To confirm that NGS could outperform Sanger sequencing and accurately identify enriched sequences, we performed Sanger sequencing on multiple clones from high dilution. Using BP3 and BP4 samples, we isolated 293 clones from high dilution titer plates by Sanger sequencing, and of these, 156 unique CDR3 regions demonstrated enrichment at the level of the CDR3. Overall, 81.6% of all VHH clones were identified by NGS from BP3 or BP4, which confirms that enriched clones could be identified by both methods (**Sup. Fig. 8**). A histogram plotting the Log2 frequency of counts on the x-axis by the proportional density of clones at that frequency on the y-axis shows that the Sanger clones were enriched but covered the broad range of enriched NGS clones in BP3-4 (**Fig. 3B**). NGS detected many enriched clones within the abundance range of Sanger sequencing, which supports a deeper assessment of enriched clones. This is critical for parallel enrichment, given that important antigen binding VHHs may show a broad range of abundance when selected in a complex system via *in vivo* biopanning (**Fig. 3C**).

Before biologically evaluating the VHH nanobodies, we created criteria to shorten the list of NGS-VHHs that have meaningful properties that we can infer from NGS. We collapsed the files from each tumor type, cell subpopulation, tumor, and dLN tissue. We then identified clones that were enriched in BP3 and BP4 over BP1 and BP2. To determine whether there was cell type selectivity, we analyzed the filtered clones

for higher enrichment in later BP rounds than in early rounds for each cell subtype. We generated a heatmap that displays the relative enrichment of all files in a given cell subtype compared to each of the other subtypes, which revealed a list of clones that are specific for each cell subtype and confirmed that parallel enrichment is possible **(Fig. 3D)**.

From the identified list, we selected two VHHs that overlapped with Sanger-VHHs to test their biological properties: Cl.1 and Cl.2. The sequences were cloned into an expression system, transduced, and purified. We cloned two controls, a dendritic cell binding VHH (DC2.1) and a negative control VHH (BCII10)[15]. The VHH proteins were purified, then stained with Coomassie and a separate western blot against the His-tag to confirm protein purity and identity (**Fig. 3E**). Harvested cells were incubated with the nanobody, anti-His antibody, and immunophenotyping antibodies to determine the selectivity of VHH Cl.1 and VHH Cl.2 binding to isolated splenic cells. We observed that the VHH Cl.1 and VHH Cl.2 bind selectively to the $CD11b^+$ and $CD11c^+$ cell populations in the spleen of naïve mice (**Sup. Fig. 9A**). We implanted Py117 tumors into mice, then isolated leukocytes from the draining LNs and tumors. The VHH Cl.1 and VHH Cl.2 bound to the $CD11b^+$ and to a lesser extent $CD11c^+$ and $CD4^+$ populations (**Fig. 3F, Sup. Fig. 9B**). The BCII10 negative control nanobody had minimal signal in different subtypes as with secondary antibody alone. The DC2.1 VHH had strong binding to all cell subtypes, which suggests different specificity of the target or non-specific signal compared to the clones from biopanning (**Fig. 3F and Sup. Fig. 9**). To show antigens can be identified, Cl.1 protein was incubated with membrane extract from splenocytes, immunoprecipitated using both His-tag VHH and an FC fusion. One large band was identified on both gels, extracted and analyzed by tandem mass spectrometry, with results showing peptide sequences from multiple proteins that are part of the desmosome complex (**Sup. Fig 10 A-B**).

We aim to develop therapeutically active nanobodies so after demonstrating unique antigen targets can be identified we turned our attention to show nanobodies can functionally alter immune cells in the

TME. We performed scRNAseq after injecting library pools to demonstrate transcriptional activity that can be attributed to enriched phage. We used treatment unresponsive Py8119 tumors to demonstrate the impact of enriched phage on TIL phenotype with the goal of altering the tumors to be more immunologically responsive. To develop selective pools of enriched phage, library from BP1 was injected into mice bearing Py8119 tumors. After 2 hours, mice were euthanized, and tumors were collected and dissociated to allow infiltrating $CD45^+$ cells to be isolated via FACS. The $CD45^+$ phages were expanded then injected into a second round of Py8119 tumor-bearing mice and digested cells were sorted for $CD8^+$ T cells and $CD11c^+CD11b^-$ dendritic cells. Enriched phage pools from $CD45^+$, $CD8^+$, $CD11c^+CD11b^-$, PBS, and phage without a nanobody were injected a final time into tumor bearing mice to process cells for scRNAseq. Samples were pooled and annotated, then we identified six $CD8^+$ T cell subclusters (**Fig. 3G-H, Sup. Fig. 11**). A greater proportion of clusters 0, 3, and 5 were occupied by the $CD8^+$ T cell-specific phage injections, suggesting $CD8^+$ phages altered the transcriptomic phenotype in a unique way (**Fig. 3I and Sup. Table 2**). The expression profile revealed greater $CD8^+$ T cell activation markers in these populations suggesting that the $CD8^+$-specific phages could lead to greater T cell activation. Similarly, $CD11c^+$ and $CD8^+$ phage-injected mice demonstrated increased DC activation pathway genes relative to controls (**Fig. 3K and Sup. Table 3**). GO term analysis in the DC population revealed enrichment of multiple pathways involving RNA processing and immune response, supporting the induction of an immune response from $CD11c^+$ phages two hours after injection into mice (**Fig. 3L and Sup. Table 4**).

In this study, we demonstrate that phage display can identify cell binding VHH nanobodies through *in vivo* selection designed to identify nanobodies and targets that bind to TILs. By using NGS to evaluate biopanning of selected immune cell subpopulations, we identified the robust enrichment of nanobodies in the TME. Not only did our screening identify the enrichment of the broad $CD45^+$ population, but it also identified parallel enrichment in which multiple nanobodies selective for different cell subtypes could be identified. The power of this approach was amplified by the use of NGS assessment, which

brought us to the conclusion that fewer rounds of biopanning are necessary to achieve enriched phage, thereby allowing for screens in complex experimental settings. We believe these methods could be broadly applicable to multiple biologic discovery experiments in different pathologic contexts. We hypothesize that these methods can be extended to study complex *in vivo* microenvironments and identify nanobodies that are useful for therapeutic and diagnostic applications, as suggested by the scRNAseq experiment. We next aim to demonstrate that this technologic platform can be used to develop novel cancer immunotherapies.

Methods

Chemicals and Biological Materials

The nanobody-encoding M13 phage display library was purchased from QoolAbs Inc. (San Diego, CA). TG1 electrocompetent cells were procured from Lucigen (Middleton, WI). M13KO7 helper phage particle, restriction and modification enzymes, and PCR amplification materials were from New England Biolabs (Ipswich, MA). Antibiotics and bacterial culture media such as 2-YT broth, Terrific broth, and Agar were from ThermoFisher Scientific. All fine chemicals, IPTG, PEG8000, and Imidazole were from Millipore Sigma (St. Louis, MO). HisTrap columns and Ni sepharose columns were from GE Healthcare Life Sciences (Pittsburgh, PA). Cell culture materials are from Mediatech Inc. (Manassas, VA) and CytivaHyClone Fetal Clone II was purchased from GE Healthcare Life Sciences.

Cell culture

Py8119 and Py117 breast cancer cell lines were derived from spontaneous tumors of transgenic MMTV-PyMT mice congenic in the CL57BL/6 background[9,16]. Cells were cultured and maintained in Hams F12 media (Corning) supplemented with 5% CytivaHyClone Fetal Clone II (GE Healthcare Lifesciences), MITO+ (1:1,000 dilution, BD Biosciences), 50 mg ml^{-1} gentamicin and 2.5 mg ml^{-1} amphotericin B. For tumor implantation, cells were seeded in complete medium 48 hours before implantation at 60 to 70% confluence and fresh medium was changed 24 hours prior to harvesting for

implantation. Py8119 and Py117 cells with a passage number ranging from 15 to 25 were used for all tumor implantations. Cells were tested for mycoplasma contamination periodically using ATCC PCR Mycoplasma Detection Kit (Manassas, VA).

Animals

All biopanning experiments were carried out using C57BL/6 wild-type mice with age group ranging from 6 to 8 weeks. All animals were purchased from Jackson Laboratories (Bar Harbor, ME), and maintained in UT Southwestern Medical Center animal facilities by following guidelines according to UT Southwestern Medical Center Institutional Animal Care and Use Committee (2017-102240). Tumors were implanted subcutaneously in the mammary fat pad with 1×10^6 cells of Py8119 or Py117 breast cancer cells suspended in 50 µl PBS and mixed with equal volume of growth-factor reduced Matrigel (Corning-356231). Experiments with Py8119 or Py117 tumor-bearing mice were conducted two or three weeks after inoculation, respectively. The phage library was injected intravenously in 150 µl of PBS pH 7.2, and animals were monitored for two hours post-injection period. We noted that injection of undiluted, high titer (10^{14}) phages reduced the activity of the mice. However, with a 100-fold dilution, there was no clear physiologic effect on the animals.

Isolation of Immune Cells

Mice were euthanized two hours post-IV injection, LN and tumor were excised, and tissue was dissociated into a single cell suspension. The LN was processed by gentle homogenization and 40 µM filtration in 10% serum-supplemented medium. Tumor tissue samples were digested with 4 ml of serum-free medium containing 5 µM liberase and 100 µM DNAse for 40 min at 37°C with shaking at 70 RPM. The tumor digestion protocol #4 was performed by using the gentleMACS Octo tissue dissociator (Miltenyi Biotec) with three 10 min rounds of processing over 40 min. Serum-supplemented medium was added to the digestion and cells were filtered through a 40 µM filter. The cell suspension was pelleted, then RBCs were lysed for 5 min at 4°C in ACK buffer, then cells were resuspended in fresh

medium. TILS were isolated via ficoll gradient, suspended in fresh medium, and washed. Cells were resuspended in MACS staining buffer and passed through LS columns with dead cell removal magnetic bead kit (Miltenyi Biotec). $CD11b^+$-positive selection magnetic beads were first used to isolate $CD11b^+$ cells. The flow through was then stained with CD11c+ positive selection magnetic beads to isolate $CD11c^+CD11b^-$ cells. The flowthrough cells were then divided, 1/3 for $CD8^+$ selection and 2/3 for $CD4^+$ subsets (both positive selection for $CD4^+CD25^+$cells and negative selection for $CD4^+CD25^-$ cells). A small portion of each cell fraction was analyzed by flow cytometry to confirm enrichment of the desired cell subtype. It is important to note that this method does not completely eliminate all other cell types.

In vivo Biopanning

For *in vivo* biopanning, we intravenously injected $1.5x10^{12}$ phages in 150 µl of PBS pH7.2 into Py8119 and Py117 tumor bearing CL57BL/6 mice. Tumor and LN samples were processed for immune cell isolation two hours post-injection of phage library. Three tumor-bearing animals were used for every tumor type in each round of biopanning. For the preliminary round of $CD45^+$ cell-specific phage selection (BP0), we isolated immune cells from tumor tissue as described above, and selectively isolated $CD45^+$ cells by using $CD45^+$-positive selection magnetic beads after dead cell removal. Phage particles enriched in the preliminary round of biopanning were amplified and prepared for BP1.
For BP 1 through 4, we isolated immune cells following ficoll gradient protocol, performed dead cell removal, then selectively isolated six different cell types: (Total TIL ($CD45^+$), $CD11b^+$, $CD11c^+CD11b^-$, $CD8^+$, $CD4^+CD25^-$, and $CD4^+CD25^+$). We used the phage library amplified from the total TIL population for subsequent rounds of biopanning.

Amplification of M13 Phages

We produced the M13 phage library by transforming 2 µg of recombinant phagemid DNA in TG1 electrocompetent *E.coli* cells and by adding M13KO7 helper phage to a final titer of 10^9/ml. We used ampicillin to a final concentration of 50 µg/ml for selection and phage particles were precipitated with

0.3 volume of PEG8000 after centrifuging out the bacterial pellet. Phages were titered by serially diluting the phage library and plating them after infecting them with TG1 logarithmic culture for 30 min at 37° C. For the amplification of phages in between biopanning, 50 µl of ficoll-separated immune cells or $CD45^+$ sorted cells were mixed with 200 µl of logarithmic TG1 bacterial culture and incubated at 37° C for a minimum of 30 min, then 1 ml of 2-YT medium supplemented with 50 µg/ml of ampicillin was added. M13KO7 helper phage was added to a final titer of 10^9/ml, and the phage-infected bacterial culture was transferred to sterile 250 ml 2-YT broth with 50 µg/ml of ampicillin and incubated overnight at 37°C with shaking at 200 RPM. To harvest phage particles, we centrifuged bacterial culture at 4500 RPM for 10 min at 4° C and the culture-free medium was precipitated with 0.3 volume of PEG8000 and incubated at 4°C for two hours. Phage precipitate was concentrated by centrifugation at 4500 RPM for 20 min at 4°C. The phage pellet was suspended in 5 ml of sterile PBS pH7.2 and filtered through 0.22 µM low-protein binding filter. Phages were tittered as previously described.

Selection of Phage clones and Sanger Sequencing

Phage clones with nanobody inserts were screened by colony-PCR during phage tittering of immune cells ($CD45^+$, $CD11b^+$, $CD11c^+CD11b^-$, $CD8^+$, $CD4^+CD25^-$, $CD4^+CD25^+$) isolated from biopanning 3 and 4. Phage clones were selected from plates of higher dilutions to target highly enriched clones. Nanobody-positive clones were inoculated in 2-YT medium supplemented with 50 µg/ml of ampicillin and recombinant phagemid DNA extracted after overnight incubation at 37° C. Phagemid-extracted DNA was labeled and stored at -20°C for future use. Phage clones were Sanger sequenced at the UT Southwestern Medical Center DNA sequencing core facility using phiS2 5'-ATGAAATACCTATTGCCTACGG forward and psiR2 5'-CGTTAGTAAATGAATTTTCTGTATGAGG reverse primer sequences.

Next-Generation Sequencing (NGS)

For NGS sample preparation, 25 µl of each immune cell subtype isolated either by MACS magnetic beads or flow cytometry sorting were mixed with 200 µl of logarithmic TG1 bacterial culture and incubated two hours at 37° C. After incubation, phagemid DNA was extracted from bacterial culture by QIAprep spin mini prep plasmid extraction kit (Qiagen) and OD260 value was measured by NanoDrop™2000 (ThermoFisher). Phagemid DNA was amplified for five cycles using the phiS2 forward and psiR2 reverse primers. The amplicons were further amplified for 20 cycles and the resulting amplicons were resolved in 1.5% agarose gel to cut out band sizes ranging from 550 to 850 bp. Amplicons were concentrated by QIAquick gel extraction kit and estimated by Bioanalyzer (Agilent). The NGS sequence library was created by ligating illumina universal sequence adapter with barcode for MISeq paired 300 bp reads (Illumina). NGS sequencing was performed at the UT Southwestern Medical Center NGS core facility.

NGS Data Analysis

NGS analysis began with removal of remaining adapters or primers that carried over from PCR amplification or sequencing. Poor quality reads were subsequently removed. To construct the VHH antibodies from NGS sequencing, we first used MiXCR V3 to align the reads to the IMGT Alpaca IG reference library with default settings. Then, we fully assembled all regions by changing the assembling features to include not just CDR3 but all regions from FR1 to FR4, including CDR1 and CDR2. This method ensures a full reconstruction of the VHHs. For any further analysis, we exported the fully assembled VHHs, disregarding partially assembled VHHs, to VDJtools and/or Immunarch for exploratory visualization, diversity calculation, overlapping, antibody tracking and enrichment. Diversity was calculated via commonly used indices such as Simpson and Shannon. True diversity or the effective number of types was defined as the number of equally abundant types needed for the average proportional abundance of the types to equal that observed in the dataset of interest, where all types may not be equally abundant.

Expression and Purification of Nanobody

We subcloned each nanobody clone sequence in the pET21d+ vector in between NheI and XhoI sites. Subcloned nanobody-expressing vectors were transformed in BL21 (DE3)-pLysS-competent cells and a single colony was cultured in 50 ml of 2YT broth supplemented with 50 µg ml^{-1} ampicillin in an orbital shaker at 37°C with rotation at 200 RPM. After overnight incubation, the culture was transferred to 500 ml of 2-YT broth in 1L conical flask and incubated further until the OD600 reached ~0.6, then induction was performed using IPTG at a final concentration of 500 µM. Bacterial cells were pelleted by centrifugation at 5000 RPM for 10 min at 4° C. The culture pellet was suspended with 5 ml of bacterial lysis buffer (10 mM imidazole, 300 mM NaCl, 0.1% Triton X-100 in phosphate buffer pH 8.0) freshly supplemented with lysozyme and kept in ice for 30 min to ensure bacterial disruption, then sonicated three times with a 5 sec on, 5 sec off cycle at 60% amplitude. The lysate was centrifuged at 5000 RPM for 20 min at 4° C, and the supernatant was transferred to a 15 ml conical tube for AKTA pure purification. His-tag purification was performed by using 1 ml HisTrap HP column in AKTA pure 25 (Cytiva, MA). The protocol included binding for 5 min in the presence of phosphate buffer with 300 mM NaCl and 10 mM imidazole, washing for 15 column volumes with a flow rate of 1 ml/min in the presence of phosphate buffer with 300 mM NaCl and 20 mM imidazole. Finally, a linear gradient of imidazole ranging from 50 to 300 mM was used for elution. Fractions were resolved in 15% SDS-PAGE and the pure fractions were pooled and concentrated by passing through 30 kDa and 3 kDa cut off Amicon centrifugal filter unit. Protein estimation was performed by nanodrop OD280 and BCA assay (ThermoFisher, NY). LPS level was measured via Pierce LAL chromogenic endotoxin quantification kit to ensure the LPS limit fo less than 0.5 EU/mL (ThermoFisher, NY). Nanobody proteins were confirmed by immunoblotting using anti-His-tag (2365S, Cell Signaling Technology, MA) and anti-VHH (A01860, GenScript, NJ) antibodies.

Protein Binding by Flow Cytometry

Target binding of nanobody proteins was assessed by flow cytometry analysis. Immune cells from tumor and draining lymph nodes were isolated from Py8119 tumor-bearing mice as described previously and suspended in PBS supplemented with 5% BSA. Two million cells were mixed with 3.0μl of Fc block and 0.25 μl of Zombie NIR, incubated at 4°C for 15 min, and washed once with PBS. Nanobody protein (2 to 4 μg) was added to the Fc block-treated cells with a final reaction volume of 100 μl and incubated for 30 min at 4° C. After washing once with PBS, APC-conjugated His-tag antibody was added, and the mixture was incubated further for 15 min at 4° C. For a negative control, APC-conjugated His-tag antibody was added to the cells treated with no nanobody proteins. Cells were washed twice and suspended in PBS before flow cytometry analysis. The following panel of antibodies was used to analyze the nanobody binding in specific subtypes of immune cells: BV421-CD45+, BV605-CD11b^{+}, AF488-CD11c^{+}, BV510-CD8^{+}, BV785-CD4^{+}, PE/Cy7-CD25^{+}, PerCP/Cy5.5-CD19^{+}, and PE-I-A/I-E. Ultracomp eBeads (ThermoFisher, NY) were used for compensation.

Immunoprecipitation of Nanobody Targets and Mass Spectrometry

To identify nanobodies specific target proteins, immunoprecipitation (IP) analysis was performed using membrane protein extracts of mouse splenocytes. Membrane protein was isolated with Mem-PER™ plus membrane protein extraction kit (ThermoFisher). For IP, Protein G Dynabead (ThermoFisher) was washed with PBS, mixed with 5μl of anti-his-tag antibody, incubated for 1h at RT, and washed three times with PBS. Nanobody protein at a concentration of 10μg was mixed with 500μg of membrane protein in PBS and mixed with Dynabead-antibody complex after 1h incubation at 4°C. The nanobody-membrane protein mixture was incubated further for 2h at 4°C, washed three times with PBS, pelleted, and resuspended in 15μl of water with 15μl of 4X protein loading dye (BioRad, Hercules, CA). The sample was resolved in 4 to 15 % gradient acrylamide gel and stained with silver stain (BioRad, Hercules, CA). Alternatively, nanobody-IgG-Fc fusion protein was used in place of anti-his-tag antibody. Specific pulled-down protein bands were analyzed by mass spectrometry for protein identification using

Orbitrap Elite (ThermoFisher) and Protein Discoverer 2.4 software for identification and quantification at the Proteomics Core Facility, UT Southwestern Medical Center, Dallas, Texas.

single cell RNA sequencing

Tissue was mechanically minced by gentleMACS dissociator (Miltenyibiotec) and enzymatically digested with 100U/ml Collagenase IV (Worthington), 10 mg/ml DNAse (Sigma Aldrich) and 10.5 mM Y-27632 (Sigma Aldrich) in HBSS (Corning) and RPMI 1640 (Mediatech, Inc., Corning) complemented with 3% FetalClone II (HyClone). It was incubated for 40 minutes at 37°C shaking incubator, then filtered through a 40µm mesh filter (Falcon). Red blood cells (RBC) were removed by ACK lysing buffer (Gibco). cells were washed with 0.04% Bovine Serum Albumin, BSA (Miltenyibiotec) in PBS(Corning). Immune cells were enriched using CD45 magnetic microbeads. Viability and cells count was measured by 0.4% Trypan blue exclusion method. Cells were frozen in 90% FetalClone II and 10% DMSO. At the time of sequencing, thawed cells were counted using a TC20 automated cell counter (BioRad) and adjusted to 1000 cells/ul in 0.04% BSA/PBS and 1ml/sample were submitted to the Next Generation Sequencing Core, McDermott Center, UT Southwestern Medical Center. Cells were loaded according to standard protocol of the Chromium single cell 3′ kit capturing 10,000 cells (V3 chemistry). All subsequent procedures including library construction were performed according to the standard manufacturer's protocol. Single cell libraries were sequenced on NovaSeq (Illumina) to a depth of 50,000 reads per cell.

Single cell RNA-seq data processing

Raw data and FASTQs were generated by the Next Generation Sequencing Core at UT Southwestern. Using Cell Ranger version 5.0.1 pipeline, FASTQs files were aligned to the

mouse reference genome (mm10) using STAR version 2.7.2a and counting using Cell Ranger count with default parameters and recommendations. Counting matrix was imported into Seurat (3.2.3) via R (4.0.3) for quality assessment and downstream analysis. Cells were filtered by excluding cells with less than 200 genes, all genes in less than 3 cells, and genes expressed being greater than 20% mitochondrial genes. Data were then normalized by NormalizeData function with default parameters. Variable genes were detected using the FindVariableFeatures function. Cell cycle scores were calculated by CellCycleScoring function, and a cell cycle difference was calculated by subtracting the S phase score from the G2M score. Data were scaled and centered using linear regression on the counts and the cell cycle score difference. Principle component analysis was run with the RunPCA function using default parameters. Batch effects were corrected, and samples were integrated by matching mutual nearest neighbors (fastMNN) using SeuratWrappers library with default parameters. Cell clusters were identified via the FindNeighbors and FindClusters functions, with 0.5 resolution and UMAP clustering algorithms. A FindAllMarkers table was created, and clusters were defined by using SingleR and celldex with ImmGenData from (https://www.immgen.org/)[17] and finally with canonical markers (Sup. Fig. 11). Gene Ontology (GO) enrichment for each cluster and between samples was performed using topGO version 2.42.0 and differentially expressed genes. GO annotations were obtained from the Bioconductor database org.Mm.eg.db version 3.12.0. Graphs and plots were generated by dittoseq version 1.2.5 or Seurat.

Acknowledgements

We thank the Moody Foundation Flow Cytometry Core at Children's Research Institute, the McDermott Center Next Generation Sequencing Core, and the Proteomics Core facility at UT Southwestern. This

work was supported by the CPRIT Firs-Time Tenure-Track Recruitment award RR170051 (T.A.A.), the Presidents Research Council Distinguished Researcher Award (T.A.A.), the Carroll Shelby Family Foundation (T.A.A.), and Cancer Biology Training Grant T32 CA 124334 (K.S.).

Disclosures

T.A.A. reports receiving research support from Galera Therapeutics (sponsored trial and research), Apexigen Inc (investigator-initiated trial support), has ownership interest including patents from Stanford University (licensed to AKSO Biopharmaceutical) and University of California San Diego (licensed to Avelas Biosciences), and has stock in Avelas Bioscience.

Figures

Figure 1: ***In vivo* biopanning for selection of VHHs that target the immune TME and assessment.** **A.** The analysis pipeline for each round of *in vivo* biopanning using VHH-nanobody phage display library. **B.** Immune cell subtype sorting strategy for phage extraction from digested tumor and lymph node processed for sequencing of selected VHH nanobodies with NGS. **C.** NGS pipeline for deep assessment of nanobody enrichment. **D.** scRNAseq experimental outline to determine functional changes in target immune cell populations.

Figure 2: **Diversity and enrichment across *in vivo* biopanning validates the selection approach.** **A.** Protein alignment of NGS VHHs showing the diversity in CDR1, CDR2 and CDR3 regions. **B.** Frequencies of CDR3 lengths by amino acid shown as histograms in the library, Py117, and Py8119 samples demonstrate an equal distribution across the samples with median length of 18 aa. **C.** Top clone proportion (left panel) and rare clonal proportion (right panel) showing summary proportion of VHHs with specific indices and counts, respectively in Py117 and Py8119 tumor samples from BP0 to BP4. **D.** Rarefication curve assessed the diversity of BP1 to BP4 in Py117 and Py8119 tumor samples through extrapolation and subsampling showing the reduction in diversity after BP1. **E.** Percentage of

unique VHH numbers in tumor samples of Py117 and Py8119. Tracking Top VHHs either in BP0 towards BP4 **(F)** or in BP4 backward to BP0 **(G)** demonstrates the biological enrichment of biopanning process *in vivo.*

Figure 3. Selection and evaluation of nanobodies developed by *in vivo* biopanning. A. Distribution of VHHs selected by titration and Sanger sequencing in NGS selected VHHs repertoire, if VHHs selected by titration they will be theoretically captured at the 75th count percentile of enriched VHHs in BP3 and BP4. **B.** Distribution of VHHs selected by titration and Sanger sequencing in NGS-selected VHHs repertoire in combined BP3 and BP4 as a proportional density plot by Log2 relative abundance of each clone detected on NGS. **C.** Distribution of abundance of all enriched clones in BP3 and BP4 identified by NGS only (left) or clones identified by Sanger sequencing and their abundance detected by NGS (right). **D.** Heatmap of relative enrichment VHHs that represent each subpopulation specifically. **E.** Purified protein by ÄKTA pure detected by Coomassie blue staining or using His tag antibody. **F.** VHH CL.1 and VHH CL.2 detected in cell subpopulation of Py117 tumor sample by flow cytometry. **G.** Integrated scRNAseq immune cell annotation for all samples. CD8+ T cell subclusters **(H)**, the relative distribution based upon sample type **(I)**, and activation marker expression **(J)**. **K.** Dendritic cell activation GO term enrichment of identified genes in each scRNAseq sample type. **L.** GO enrichment for DC cluster in DC library injected mice was performed using topGO, significance was determined by Fisher exact test (-log10).

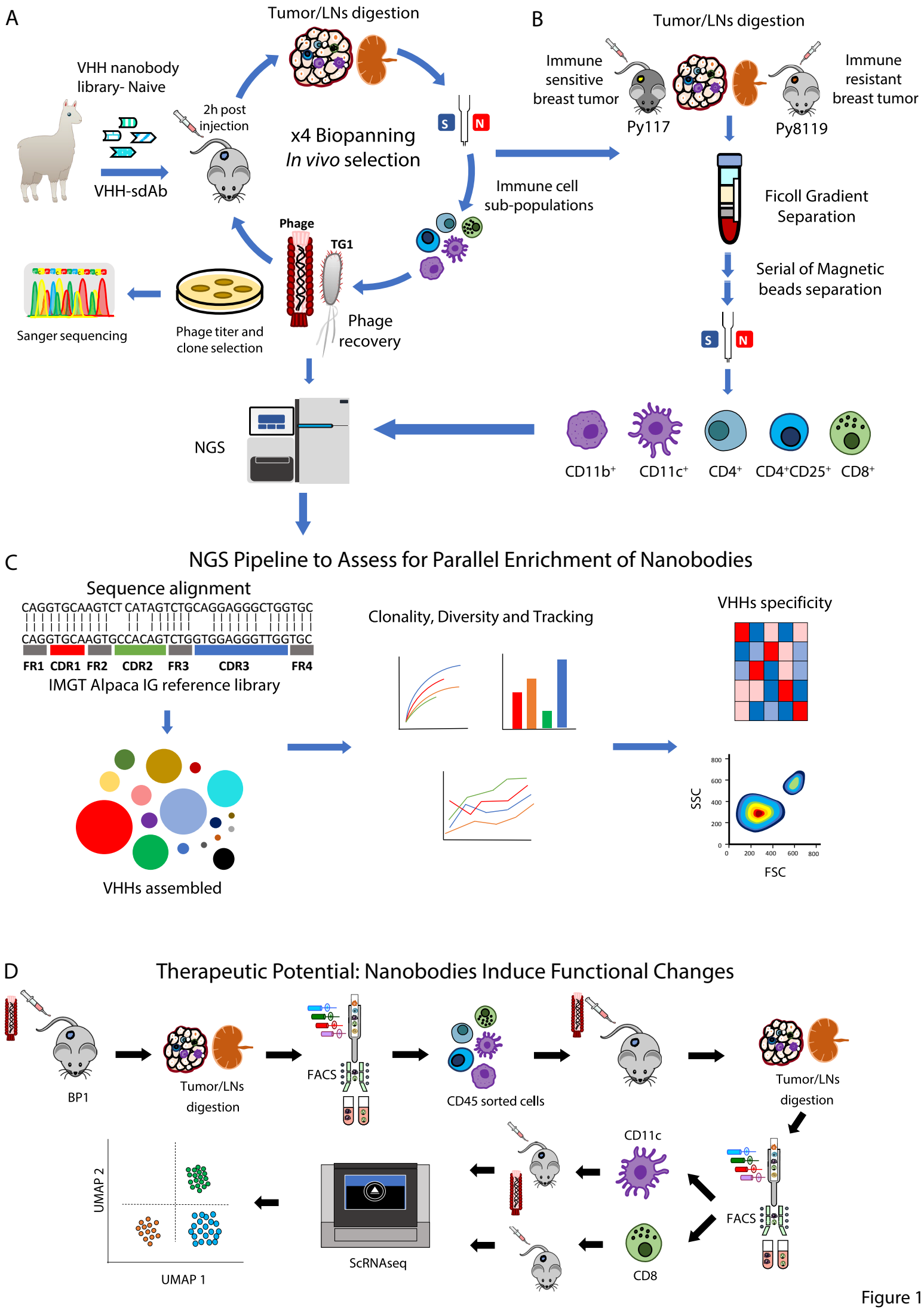
A
Tumor/LNs digestion
VHH nanobody library- Naive
2h post injection
VHH-sdAb
x4 Biopanning
In vivo selection
S
N
Immune cell sub-populations
Phage
TG1
Phage recovery
Phage titer and clone selection
Sanger sequencing
NGS
B
Tumor/LNs digestion
Immune sensitive breast tumor
Py117
Py8119
Immune resistant breast tumor
Ficoll Gradient Separation
Serial of Magnetic beads separation
S
N
CD11b+
CD11c+
CD4+
CD4+CD25+
CD8+
C
NGS Pipeline to Assess for Parallel Enrichment of Nanobodies
Sequence alignment
CAGGTGCAAGTCT CATAGTCTGCAGGAGGGCTGGTGC
CAGGTGCAAGTGCCACAGTCTGGTGGAGGGTTGGTGC
FR1 CDR1 FR2 CDR2 FR3 CDR3 FR4
IMGT Alpaca IG reference library
VHHs assembled
Clonality, Diversity and Tracking
VHHs specificity
SSC
FSC
D
Therapeutic Potential: Nanobodies Induce Functional Changes
BP1
Tumor/LNs digestion
FACS
CD45 sorted cells
Tumor/LNs digestion
CD11c
FACS
CD8
ScRNAseq
UMAP 2
UMAP 1

Figure 1

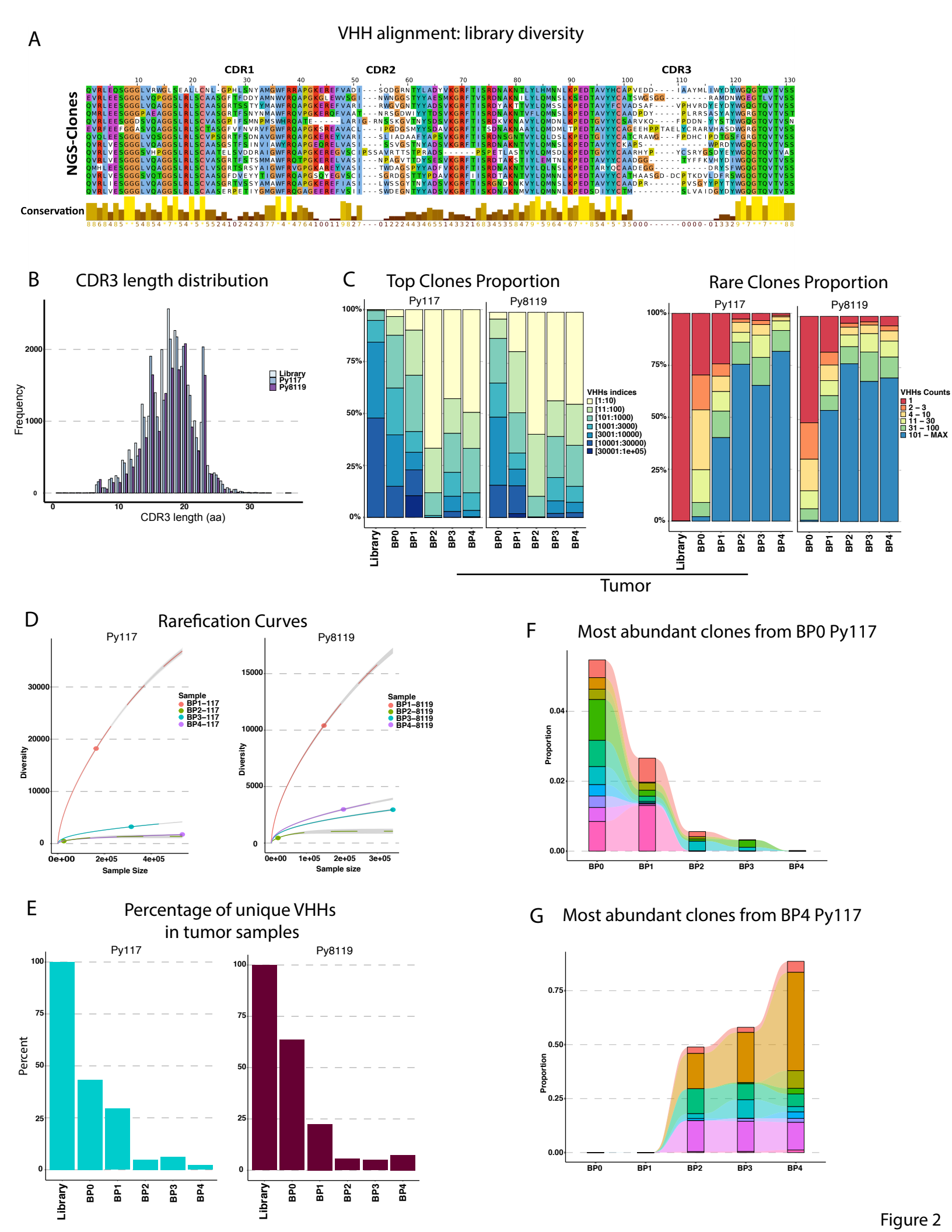
A
VHH alignment: library diversity
CDR1
CDR2
CDR3
NGS-Clones
Conservation
B
CDR3 length distribution
Frequency
CDR3 length (aa)
Library
Py117
Py8119
C
Top Clones Proportion
Py117
Py8119
VHHs indices
[1:10)
[11:100)
[101:1000)
[1001:3000)
[3001:10000)
[10001:30000)
[30001:1e+05)
Library
BP0
BP1
BP2
BP3
BP4
Tumor
Rare Clones Proportion
VHHs Counts
1
2 – 3
4 – 10
11 – 30
31 – 100
101 – MAX
D
Rarefication Curves
Py117
Py8119
Sample
BP1–117
BP2–117
BP3–117
BP4–117
BP1–8119
BP2–8119
BP3–8119
BP4–8119
Diversity
Sample Size
Sample size
E
Percentage of unique VHHs
in tumor samples
Percent
F
Most abundant clones from BP0 Py117
Proportion
G
Most abundant clones from BP4 Py117

Figure 2

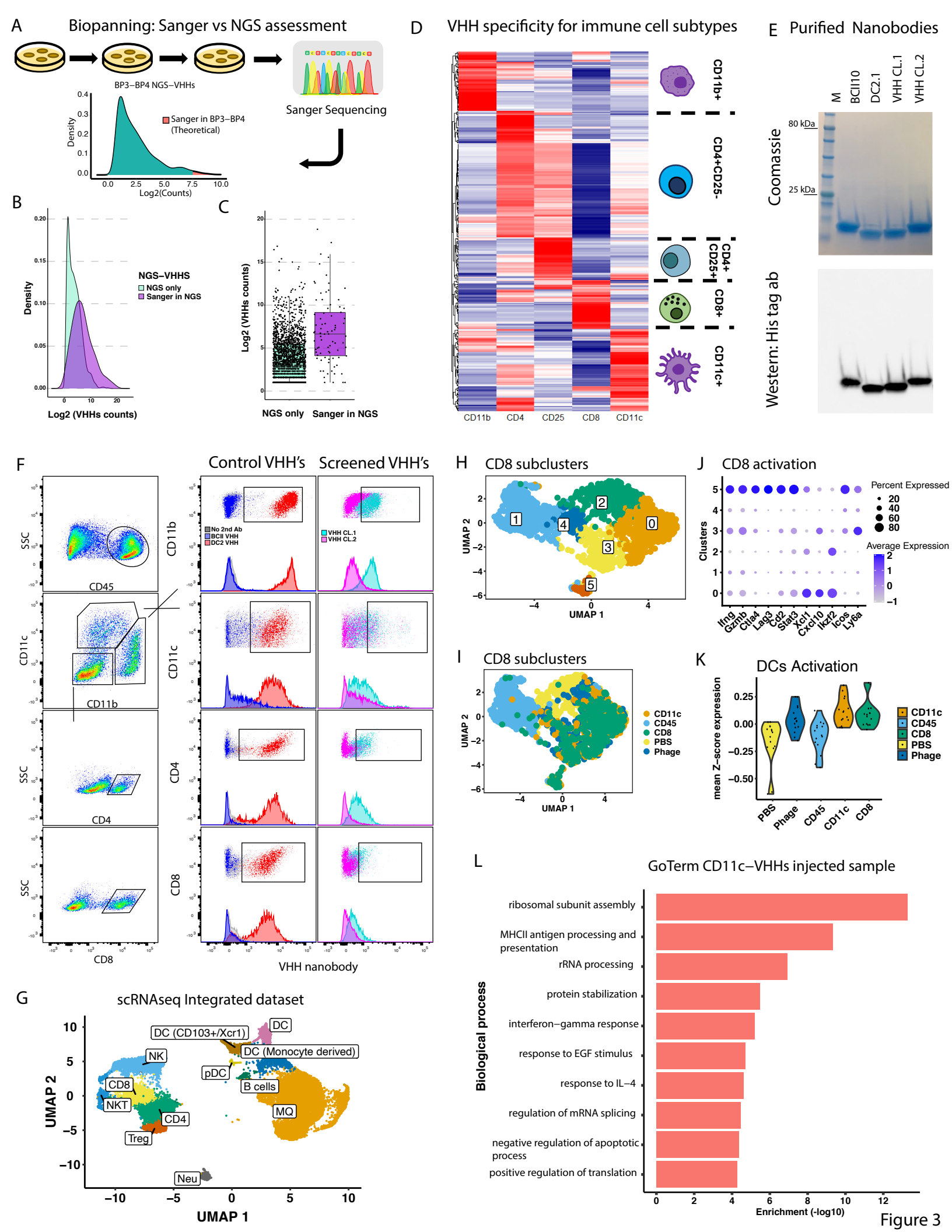
A
Biopanning: Sanger vs NGS assessment
BP3–BP4 NGS–VHHs
Sanger in BP3–BP4 (Theoretical)
Density
Log2(Counts)
Sanger Sequencing
B
NGS–VHHS
NGS only
Sanger in NGS
Density
Log2 (VHHs counts)
C
Log2 (VHHs counts)
NGS only
Sanger in NGS
D
VHH specificity for immune cell subtypes
CD11b+
CD4+CD25-
CD4+ CD25+
CD8+
CD11c+
CD11b
CD4
CD25
CD8
CD11c
E
Purified Nanobodies
M
BCII10
DC2.1
VHH CL.1
VHH CL.2
80 kDa
25 kDa
Coomassie
Western: His tag ab
F
Control VHH's
Screened VHH's
SSC
CD45
CD11b
CD11c
CD4
CD8
No 2nd Ab
BCII VHH
DC2 VHH
VHH CL.1
VHH CL.2
VHH nanobody
G
scRNAseq Integrated dataset
DC (CD103+/Xcr1)
DC
DC (Monocyte derived)
NK
pDC
CD8
B cells
NKT
MQ
CD4
Treg
Neu
UMAP 2
UMAP 1
H
CD8 subclusters
UMAP 2
UMAP 1
I
CD8 subclusters
CD11c
CD45
CD8
PBS
Phage
UMAP 2
UMAP 1
J
CD8 activation
Percent Expressed
Average Expression
Clusters
Ifng
Gzmb
Ctla4
Lag3
Cd2
Stat3
Xcl1
Cxcl10
Ikzf2
Icos
Ly6a
K
DCs Activation
mean Z-score expression
PBS
Phage
CD45
CD11c
CD8
L
GoTerm CD11c–VHHs injected sample
Biological process
ribosomal subunit assembly
MHCII antigen processing and presentation
rRNA processing
protein stabilization
interferon–gamma response
response to EGF stimulus
response to IL–4
regulation of mRNA splicing
negative regulation of apoptotic process
positive regulation of translation
Enrichment (-log10)

Figure 3

References

1. Hegde PS, Chen DS. Top 10 Challenges in Cancer Immunotherapy. Immunity 2020;52:17-35.
2. Murciano-Goroff YR, Warner AB, Wolchok JD. The future of cancer immunotherapy: microenvironment-targeting combinations. Cell Res 2020;30:507-19.
3. Zhang Y, Zhang Z. The history and advances in cancer immunotherapy: understanding the characteristics of tumor-infiltrating immune cells and their therapeutic implications. Cell Mol Immunol 2020;17:807-21.
4. Pasqualini R, Ruoslahti E. Organ targeting in vivo using phage display peptide libraries. Nature 1996;380:364-6.
5. Mi Z, Lu X, Mai JC, et al. Identification of a synovial fibroblast-specific protein transduction domain for delivery of apoptotic agents to hyperplastic synovium. Mol Ther 2003;8:295-305.
6. Laakkonen P, Akerman ME, Biliran H, et al. Antitumor activity of a homing peptide that targets tumor lymphatics and tumor cells. Proceedings of the National Academy of Sciences of the United States of America 2004;101:9381-6.
7. Whitney M, Crisp JL, Olson ES, et al. Parallel in vivo and in vitro selection using phage display identifies protease-dependent tumor-targeting peptides. J Biol Chem 2010;285:22532-41.
8. Whitney MA, Crisp JL, Nguyen LT, et al. Fluorescent peptides highlight peripheral nerves during surgery in mice. Nat Biotechnol 2011;29:352-6.
9. Aguilera TA, Rafat M, Castellini L, et al. Reprogramming the immunological microenvironment through radiation and targeting Axl. Nat Commun 2016;7:13898.
10. Aguilera TA, Elghonaimy E, Shehade H, et al. Induced tumor heterogeneity reveals factors informing radiation and immunotherapy combinations. Clin Cancer Res 2020.
11. Iezzi ME, Policastro L, Werbajh S, Podhajcer O, Canziani GA. Single-Domain Antibodies and the Promise of Modular Targeting in Cancer Imaging and Treatment. Front Immunol 2018;9:273.
12. Gonzalez-Sapienza G, Rossotti MA, Tabares-da Rosa S. Single-Domain Antibodies As Versatile Affinity Reagents for Analytical and Diagnostic Applications. Front Immunol 2017;8:977.
13. Muyldermans S. Nanobodies: natural single-domain antibodies. Annu Rev Biochem 2013;82:775-97.
14. Hamers-Casterman C, Atarhouch T, Muyldermans S, et al. Naturally occurring antibodies devoid of light chains. Nature 1993;363:446-8.
15. De Groeve K, Deschacht N, De Koninck C, et al. Nanobodies as tools for in vivo imaging of specific immune cell types. J Nucl Med 2010;51:782-9.
16. Gibby K, You WK, Kadoya K, et al. Early vascular deficits are correlated with delayed mammary tumorigenesis in the MMTV-PyMT transgenic mouse following genetic ablation of the NG2 proteoglycan. Breast Cancer Res 2012;14:R67.
17. Heng TS, Painter MW, Immunological Genome Project C. The Immunological Genome Project: networks of gene expression in immune cells. Nature immunology 2008;9:1091-4.