Supplementary Information for:

Highly multiplexed and quantitative cell surface protein profiling using genetically barcoded antibodies

Short title: Cell profiling via genetically barcoded antibodies

Samuel B. Pollock[a], Amy Hu[b], Yun Mou[a], Alexander J. Martinko[a], Olivier Julien[a], Michael Hornsby[a], Lynda Ploder[b], Jarrett J. Adams[b], Huimin Geng[c], Markus Müschen[c,1], Sachdev S. Sidhu[b], Jason Moffat[b], James A. Wells[a,2]

[a]Department of Pharmaceutical Chemistry, University of California, San Francisco, CA 94143, USA.

[b]Donnelly Centre, University of Toronto, Toronto, ON M5S 3E1, Canada.

[c]Department of Laboratory Medicine, University of California, San Francisco, CA 94143, USA.

[1]Present address: Department of Systems Biology, Beckman Research Institute and City of Hope Comprehensive Cancer Center, Pasadena, CA 91006.

[2]To whom correspondence should be addressed:
James A. Wells
1700 4th St
San Francisco CA 94158
415-514-4498
jim.wells@ucsf.edu

Classification: Biological Sciences/Applied Biological Sciences

SI Materials and Methods

Cloning. Construction of plasmids expressing the extracellular domains of receptors linked to transmembrane (ECDtm) for generation of over-expression cell lines was carried out using Gibson cloning (36). ECD sequences were taken from Uniprot, codon optimized for mammalian expression (Integrated DNA Technologies, IDT), and obtained as gBlocks (IDT). Primers (obtained from IDT) were used to amplify the gBlocks and append ~60˚C Tm sequence complementary to each insertion site into the vector (LEFT sequence = GCGGGGGGCGGTCATATG, RIGHT sequence = GGATCCGGCGGAAGCGG, see SI Appendix, File S2). The customized vector shown in File S2 (pCDNA5/FRT/TO, ThermoFisher Inc.) has the following construct between the TetO-CMV promoter and the stop codon: Signal sequence – HA tag – GFP – linker – PDGFR transmembrane domain. The vector was opened using the reverse complement of the "LEFT" sequence, and the "RIGHT" sequence as written above. The insert and vector were then gel purified before the Gibson reaction was carried out as previously described (36). The ECDtm plasmids transformed into competent XL-10 cells (Agilent), selected on 100μg/mL carbenicillin plates, and verified by Sanger sequencing (Quintara Biosciences).

Cell culture. The P493-6 cell line was a kind gift from the Ruggero lab (UCSF) and were grown in RPMI-1640 media (GE Healthcare), 10% FBS (Gemini Bio-Products), 1X penicillin-streptomycin (Pen-strep, Gemini Bio-Products). LAX7D and LAX7R cells were generated as described below and in the text (Fig. 2a) and grown in MEM-α media with nucleosides (ThermoFisher Inc.), 10% FBS, 1X Pen-strep. HeLa cells were obtained from ATCC (CCL-2), grown in DMEM/HG media (Life Technologies), 10% FBS, 1X Pen-strep. The MCF10A cell line and transformants were a kind gift from the Bandyopadhyay lab (UCSF) and were grown as described previously (37).

Patient samples. Patient-derived leukemia samples were collected with informed consent from all participants according to NCI/CTEP approved protocol ECOG E2993T5 and studied with approval of the Institutional Review Boards of the University of California San Francisco (UCSF). The samples were taken as bone marrow biopsies, blast content >80%.

Bacterial culture. XL-1 blue cells (Stratagene) were cultured as follows (it is essential that XL-1 cultures are healthy and under proper selection conditions for quantitative phage propagation). A frozen stock at -80˚C was streaked out on a <1-month-old 10µg/mL tetracycline (Tet) agar plate and grown overnight at 37˚C. The following day, an overnight culture was made in a 250mL non-baffled flask (Corning, 70980-250) containing 25mL 2xYT+10µg/mL Tet. Similar flasks with 100µg/mL carbenicillin and 50µg/mL kanamycin were also made up to ensure no resistance to these antibiotics was present. The following morning the OD600 of the overnight culture was measured (OD600 should exceed 3.0). Dilution flasks were made up from the overnight culture in 250mL non-baffled flask containing 25mL 2xYT+10µg/mL Tet (1:1000 should achieve OD600 = 0.5-0.7 within 4-5 hours). Phage were then propagated as described below. In our experience, propagation by flask of 10^6cfu should easily yield an OD600 greater than 1.0 after 16hrs, and even 10^3cfu should yield a culture with visible cloudiness (OD600 > 0.1). If propagation by flask does not achieve these optical densities after 16hrs, a fresh Tet agar plate should be streaked out and the process started again.

ECDtm cell construction via Flp-In. Cell lines were constructed as described in the T-REx Flp-In manual (ThermoFisher Inc., R78007). Briefly, T-REx cells were grown to 70% confluency in a 6-well plate in the presence of 100µg/mL zeocin and 10µg/mL blasticidin. Growth media was replaced with Opti-MEM media (ThermoFisher Inc., 31985062) and cells were co-transfected using 1.5µg pCDNA5/FRT/TO, 1.5µg pOG44 (ThermoFisher, Inc.), and 9uL of 1mg/mL polyethylenimine (PEI, Polysciences, 23966-2) in 200µL Opti-MEM for 6hrs at 37˚C, 5% CO2. Opti-MEM was then replaced with 1:1 Opti-MEM:DMEM/HG, 10% FBS, 1X Pen-strep and placed at 37˚C, 5% CO2 for 24hrs. The 1:1 media was replaced with DMEM/HG, 10% FBS, 1X

Pen-strep, 100µg/mL hygromycin, 10µg/mL blasticidin until the appearance of resistant colonies (generally 1-2 weeks). Cells were expanded into DMEM/HG, 10% FBS, 1X Pen-strep, 50µg/mL hygromycin, 2µg/mL blasticidin and used or made into stocks.

Preparing phage input for cell panning. The PhaNGS input pool, containing approximately 150 individual Fab-phage clones (depending on the exact experiment), was typically constructed as follows: 5µL of each clone was transferred to respective wells of a 96-well round bottom plate (or plates, Corning 3797). Then 100µL log phase XL-1 blue cells (Agilent, OD600 = 0.6-0.7) was added to each well before the plate was covered in a gas permeable film (Diversified Biotech, BEM-1) and placed at 37˚C for 20min. 100uL of infected cells was then transferred to respective wells of a 96-well deep well plate (Corning Axygen, P-2ML-SQ-C) containing 400µL per well 2xYT broth with 100µg/mL carbenicillin and 10^{10} cfu/mL KO7 helper phage (NEB). The deep well plate was covered in a gas permeable film and shaken at 1000rpm and 37˚C for 18-24hrs in an Infors HT shaker. Plates were spun down at 4,000g for 15mins at room temperature and the supernatant was consolidated into 50mL tubes before adding 0.02% sodium azide and storing at 4˚C (this supernatant may be used for up to seven days with only minor loss in signal, but should be used fresh if possible). This method leads to approximately equal quantities of each clone from a propagated supernatant (roughly 10^{11} cfu/mL total). If a fast shaker is not available or a more convenient library creation method is desired, seed stocks of libraries made from the 50mL plate-propagated supernatant may be grown up in flasks (50uL to 50mL propagation volume, see Fig. S4). Seed stocks can be stored at -80˚C in 25% glycerol and grown up directly for greater convenience.

To create higher titer mixtures where applicable, propagated supernatant was concentrated as follows: ~25mL supernatant was transferred into a JA 25.50 tube (Beckman Coulter, 357003) containing 7.5mL of 20% PEG/2.5M NaCl, and this solution was mixed by inversion and then placed on ice for two hours. The precipitated phage was then pelleted at 12,000g for 15 minutes at 4˚C, the supernatant was decanted, and the phage pellet re-suspended in 1mL PBS. This

solution was transferred to a 2mL Eppendorf tube, an additional 1mL PBS was added, and the tube was mixed end-over-end for 15 minutes at room temperature. This mixture was spun down at 12,000g for 5 minutes at room temperature to remove debris, before transferring the 2mL of the supernatant to a 15mL Falcon tube containing 1mL of PBS.

Panning phage on cells, elution, propagation, and storage. For profiling a population of cells, we typically use ~1-10x10^6 cells so as to create a convenient pellet to visualize when washing (although in our experience the absence of a pellet does not necessarily imply an absence of cells, as pellet visibility disappears below ~10^6 cells). Cells of interest were transferred into 2mL Eppendorf tubes on ice for all experiments described in the manuscript (samples were kept cool throughout), although similar results may be obtained using a 96-well deep well plate (Corning Axygen, P-2ML-SQ-C). Cells were washed once (to remove media, DMSO) by spinning the cells down at 300g for 5 minutes at 4˚C, pouring off the supernatant into liquid waste, re-suspending in 1mL cold PBS, spinning down, and decanting again. The final drops during decanting were removed by inverting and dabbing the tube on a paper towel. The washed cell pellet was then re-suspended in 1mL of the input phage mixture prepared above. The tube was end-over-end rotated for 20 minutes at 4˚C (or placed at 4˚C if using a plate) before spinning down at 300g for 5 minutes at 4˚C and decanting as above. Cells were then washed four times with PBS as above, transferring to fresh 2mL Eppendorf tubes and inverting to coat the walls each time (it is critical that PBS without added protein such as BSA is used for washing along with plastic transfer. Use of PBS+BSA without plastic transfer leads to very low PhaNGS signal.) To elute the remaining cell-bound phage, the final cell pellet was re-suspended in 900µL of 0.1M acetic acid pH 2-3, allowed to sit for 5 minutes, then spun down as above and 800µL of the acid eluent transferred to a 1.5mL Eppendorf tube containing 100µL 1M Tris pH 7.5 to neutralize. The Fab-phage in the neutralized solution were propagated by plate as described above for the input pool or by flask as follows: 1mL of log phase XL-1 blue culture was added to 100µL of neutralized eluent to infect. This solution was shaken at 37˚C for 20 minutes before

adding 10^{10} cfu/mL M13K07 helper phage and shaking at 37°C for another 20 minutes. The entire mixture was then transferred to a 250mL non-baffled flask containing 25mL of 2XYT broth with 100µg/mL carbenicillin and 50µg/mL kanamycin. This flask was shaken at 250rpm, 37˚C for approximately 16-18 hours (OD600 should be greater than 0.1 after ~16hrs), before spinning down the culture at 4000g for 15 minutes at room temperature.

Whether by plate or flask, 50uL of propagated phage were then stored in a 96-well PCR plate that was then covered in foil, heated at 98˚C for 5 minutes to release phagemid DNA and sterilize the solution, and cooled to 12˚C for 5 minutes. This plate was then used immediately or stored for up to two weeks at 4˚C before continuing. Multiple boil cycles may be carried out without affecting the final sequencing results.

For single cell experiments, the same steps as above were performed except, instead of elution, single cells were sorted by forward and side scatter (or fluorescence) on a Sony SH800 cell sorter into each well of a 96-well round bottom plate containing 50µL of 2xYT broth. The plate was then propagated via addition of 50uL log phase XL-1 culture and stored as described above.

Amplification and purification. To amplify the H3 "barcode" of each phagemid for sequencing, 3µL of boiled propagate and 4µL of 5uM each forward and reverse primers were combined into 43µL of Phusion master mix containing 0.5 units of Phusion polymerase (New England Biolabs, M0530S) per well of a 96-well PCR plate (Biorad, HSP9621). Primer design is as follows (shown 5' to 3' on both primers, order as shown, see Supplementary Fig. 2): AATGATACGGCGACCACCGAGATCTACACNNNNNNNNTGAGGACACTGCCGTCTAT TATTGTGCTCGC (forward), CAAGCAGAAGACGGCATACGAGATNNNNNNNNGTGACTGGAGTTCAGACGTGTGCT CTTCCGATCTTGACCAGGGTTCCTTGACCCCAGTAGTC (reverse), where each "NNNNNNNN" region is substituted with the appropriate dual indices (Nextera Dual Index, or see SI Appendix, Dataset 4). The plate was sealed with PCR foil (Excel Scientific, F-96-100) and

the complete mix was run on a thermocycler (2˚C/second ramp rate) (11) as follows: step 1_98˚C, 5min; steps 2 to 4_98-60-72˚C, 20-15-15sec; repeat steps 2-4 11 more times; step 5_72˚C,5min; step 6_hold 12˚C.

Samples were run out on 96-well 2% agarose gels to check for the presence of the amplicon band (~180-220bp). Those samples showing bands were combined (generally 3µL per input well and 6uL per sample well), gel purified and quantified using a Nanodrop (Thermo Fisher Scientific) and/or Bioanalyzer (Agilent Technologies). The library was then transferred to a 1.7mL low-bind tube (Corning Axygen, MCT-175-L-C) and submitted to a sequencing facility for analysis on a HiSeq4000 (Illumina), along with a custom sequencing primer (order as shown): TGAGGACACTGCCGTCTATTATTGTGCTCGC (T_m = 67˚C, GC % = 52).

RNAseq. mRNA-seq was performed with Personalis® (Menlo Park, CA). Total RNA was extracted from patient samples (pair of Diagnosis and Relapse samples) using RNeasy isolation kit (Qiagen). RNA size, concentration and integrity were verified using Agilent 2100 Bioanalyzer (Agilent Technologies). 100 - 200 ng of high quality total RNA from each sample was subjected to sequencing library preparation following the protocols of Illumina TruSeq Stranded Total RNA with Ribo-Zero kit. Libraries were sequenced on Illumina HiSeq 2500 as 100-bp paired-end runs (Stranded Total RNA). Raw image data were converted into base calls and fastq files via the Illumina pipeline CASAVA version 1.8 with default parameters. All 100-bp-long paired-end reads were mapped to the reference human (GRCh38) genome sequence using STAR aligner (38) and gene based expression was calculated using the htseq-count package (39).

Flow cytometry. Cell samples were transferred to a 96-well V-bottom plate (Axygen, P-96-450-V-C), and washed once with PBS+0.5% BSA+0.02% sodium azide. Cells were re-suspended in the appropriate 1µM Fab solution or control and allowed to sit on ice for 20 minutes. Pellets were washed three times and protein A, Alexafluor 647 conjugate (Thermo Fisher Scientific, P-

21462) was added as a secondary. Cells sat on ice for 20 minutes in secondary solution in the dark, then were washed an additional two times, filtered through a 40μm filter and run on a Beckman-Coulter Cytoflex Cytometer. Data was analyzed using FlowJo. Median values were used to compare protein abundance between conditions. Uncertainty is described using confidence values (CV).

SILAC mass spectrometry. Cell samples were generated by Stable Isotope Labeling by Amino Acids in Cell Culture (SILAC) as described previously (33). Mass spectrometry work-up was performed as described previously (34). Briefly, ~$2x10^7$ cells of the first condition (i.e. MCF10A-EV) were grown in SILAC media reconstituted with L-Lysine-HCl and L-Arginine-HCl ("light" media, Sigma-Aldrich), while ~$2x10^7$ cells of the second condition (i.e. MCF10A-$KRAS^{G12V}$) were grown in SILAC media reconstituted with 13C6, 15N2 L-Lysine and 13C6, 15N4 L-Arginine ("heavy" media, Cambridge Isotope Laboratories). Cells were washed, lifted, and treated with NaIO4 followed by biocytin hydrazide. These cells were lysed with RIPA and the biotinylated protein enriched on streptavidin beads. Peptides were released by on-bead trypsinization. A separate sample was generated using PNGase F to cleave remaining peptides off the resin. Both samples were desalted and run on a Q-Exactive Mass Spectrometer (ThermoScientific Inc.). Data was analyzed using MaxQuant (35). Median SILAC ratio values were used to determine fold-change values between conditions.

SI Figures and Tables

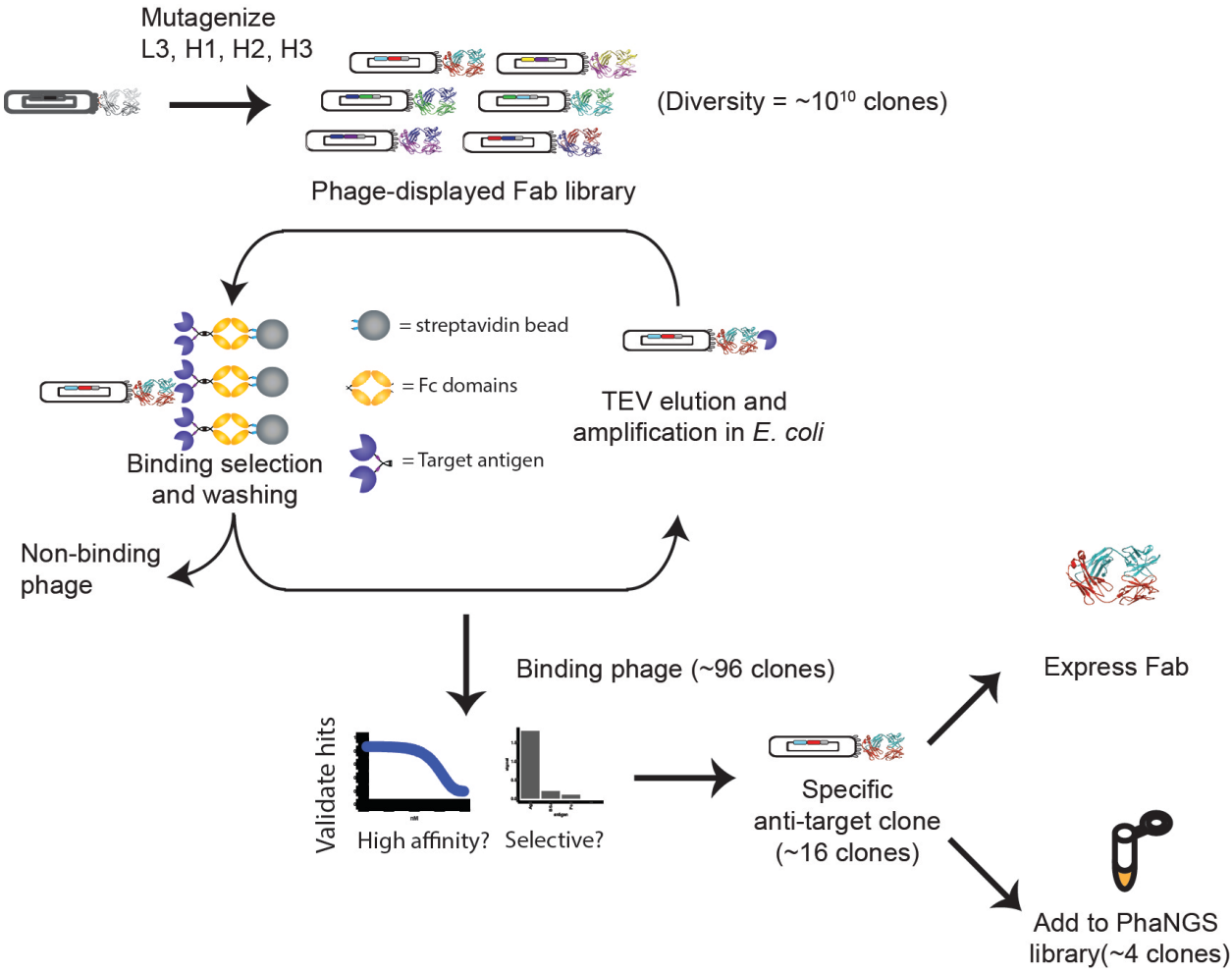


Figure S1

Phage selection scheme.

Phage were pre-selected for the PhaNGS library using a well-established workflow[5]. Briefly, a selection library with high diversity was generated. We used Kunkel mutagenesis to randomize the L3, H1, H2, and H3 regions of a displayed Fab based off of a Trastuzumab-based scaffold as described by Sidhu and colleagues (8). This library was applied to a purified and immobilized antigen, which consisted of target antigen dimerized via fusion to an Fc domain and immobilized on magnetic streptavidin beads. Non-binding phage were washed away in a series of wash steps,

and those phage that remained bound to the beads were eluted using either acid or TEV protease (for those where a cleavage site was inserted between the antigen and the Fc domain), propagated to re-make the “round 2” selection library, and the process was repeated until satisfactory enrichment of the phage pool against the target over a control antigen (such as Fc alone) was achieved. Typically, 96 clones were isolated and tested for selectivity and affinity. The best of these clones were expressed as Fabs for further validation, or, in some cases, applied as phage on over-expression cell lines for characterization using PhaNGS (Supplementary Fig. 6). On average about four passing clones were incorporated into the final PhaNGS library.

Figure S2

Primer scheme for initial PCR amplification prior to NGS.

The 5' complementary sequence is the same sequence used for the custom read1 primer, which matches the conditions required for a typical HiSeq4000 primer (T_m ~ 67˚C, 52% GC content). Nucleotides are written from 5' to 3' so when ordering, the reverse primer needs to be reverse complemented (see SI Appendix, Dataset S4). N = A/T/C/G.

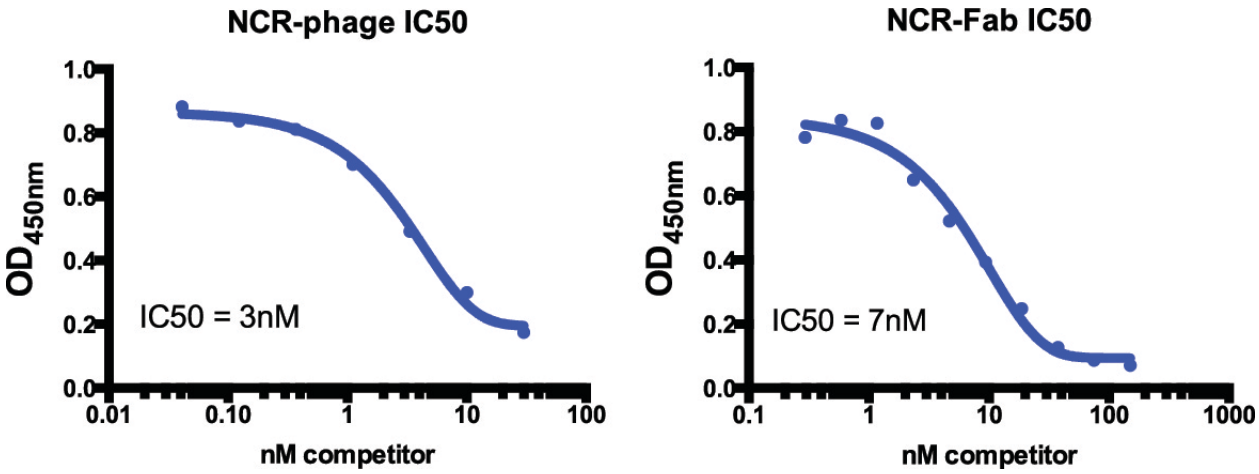


Figure S3

Fab-phage and Fab show similar affinity against the same target *in vitro*.

Expressed Fab and phage antibody affinity for NCR3LG1.01 are compared in side-by-side ELISA experiments. ELISA assays were carried out as described (5).

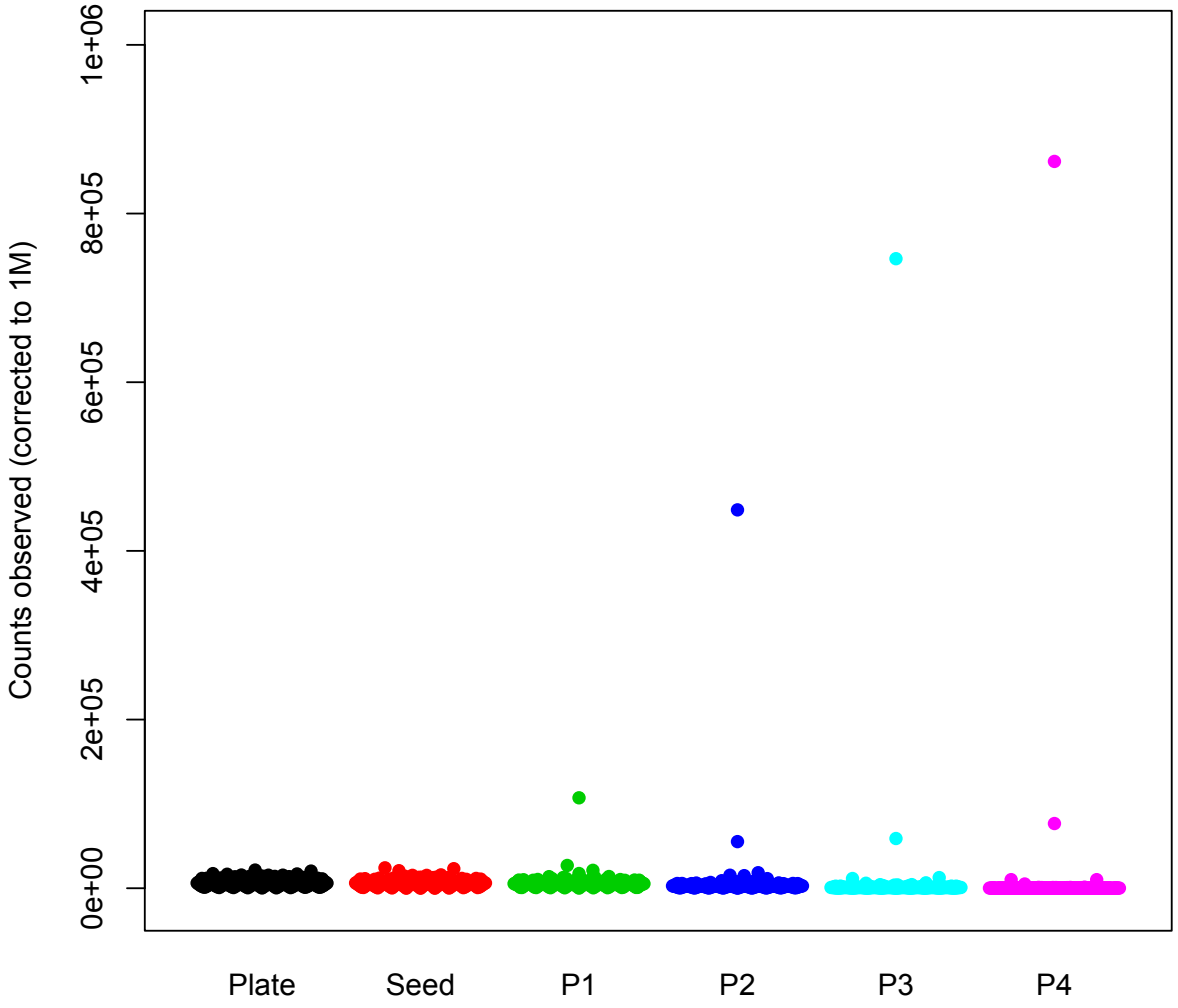


Figure S4

Most Fab-phage do not show a significant growth advantage over the others.

Input phage mixtures can be assembled by growing up each phage clone in individual wells of a deep well plate (Plate) or using a smaller seed stock that is propagated into greater volume (Seed). To assess whether the composition of a plate-derived phage mixture containing 155 clones changes significantly when grown up as a seed stock, we propagated 50µL of phage

mixture into 50mL by flask (see SI Materials and Methods), then sequenced the mixture by NGS. A significant change in relative mixture composition is not observed. If the same phage mixture (Plate) is diluted 10^6-fold and grown back to 10^{12}cfu/mL multiple times (P1-P4), only two clones (EPHA4.01 and, to a lesser extent, CSF1R.03) eventually take over the library population and can reduce counts for all other clones. This was only significant after several passages for a couple of the clones, and only observed in extreme circumstances not present in a typical PhaNGS experiment.

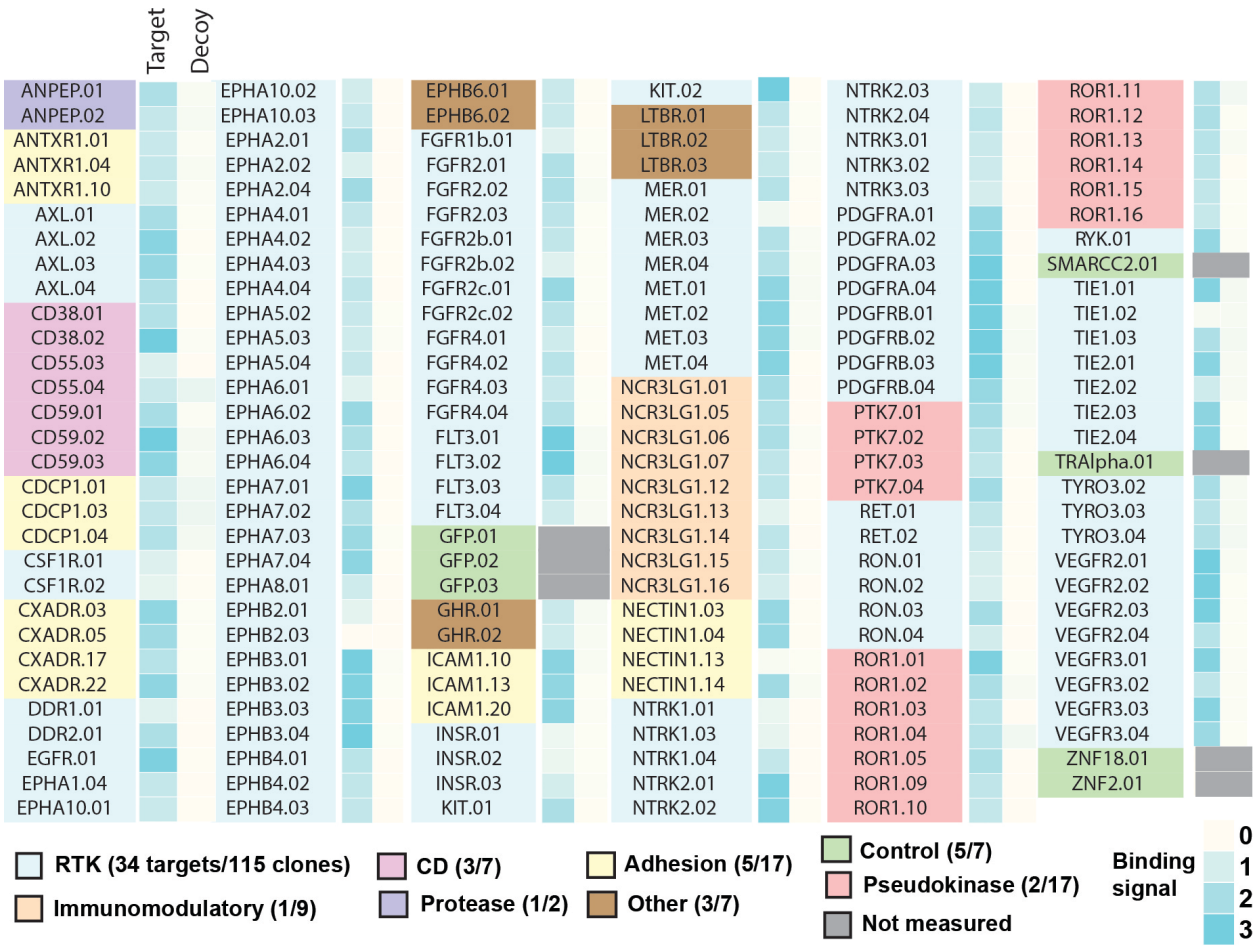


Figure S5

Current complete PhaNGS library with ELISA validation data.

A complete list of phage used in experiments throughout the manuscript are pictured. Targets are color-coded according to their class, which in turn is assigned based on functional descriptions from Uniprot. Names are assigned based on the most common alias according to Uniprot. All common names can be found in the supplementary files. Each name is assigned a number as the clone is added to the library, which constitutes its unique ID. Beside each name are heatmap tiles which display the A450 ELISA signal (also referred to as the "binding signal") against the target ectodomain or a decoy. ELISA assays were carried out as described (5). Decoys consist of Fc

domain alone for all targets except the ANTXR1 clones, CD55.03, and the GHR clones, which were instead assayed against BSA. Control phage were not assayed. See SI Appendix, Dataset 1 for details.

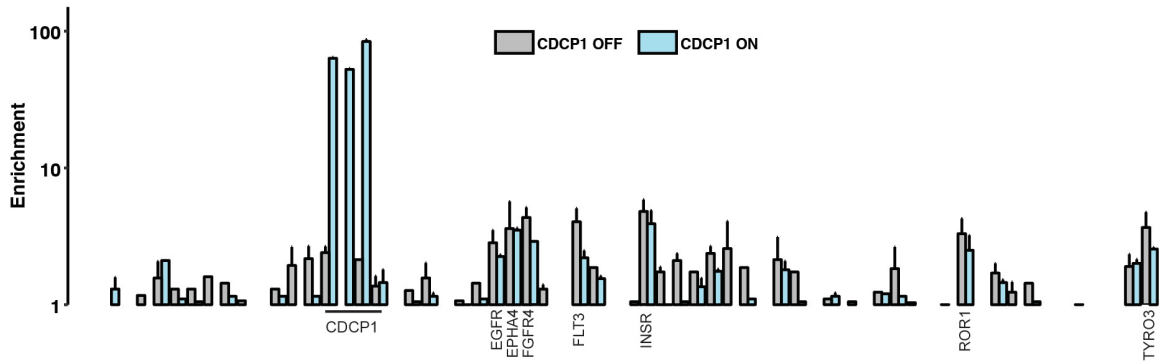
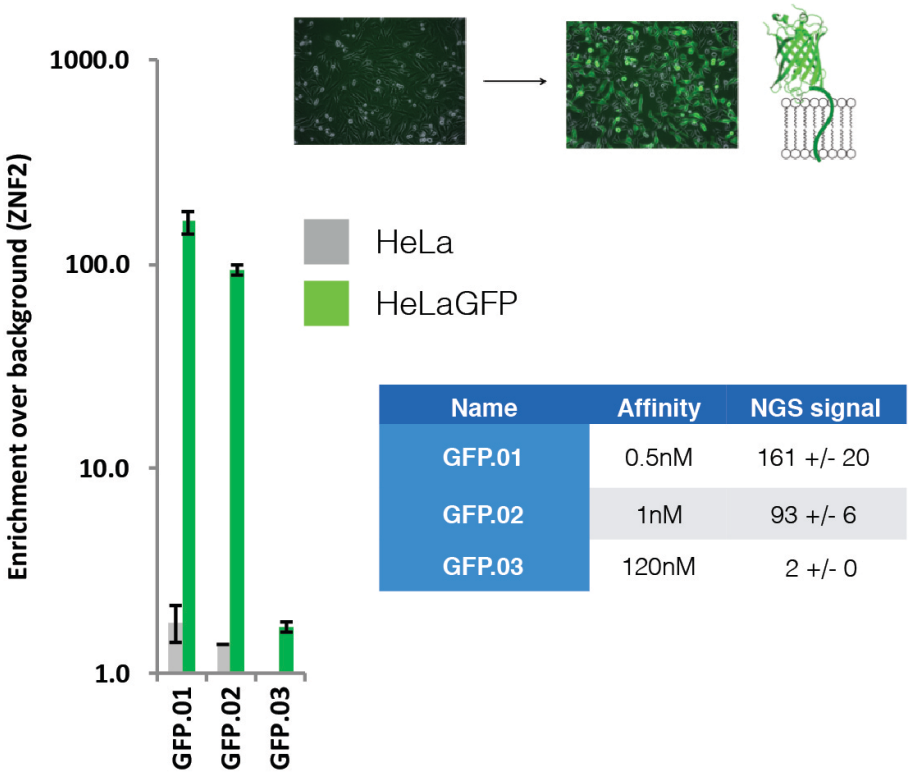


Figure S6

Profile of HEK293T and responsiveness to CDCP1 over-expression.

HEK293T cells were engineered using the Flip-In system (Thermo, see SI Materials and Methods and SI Appendix, File S2) to induce the expression of the ectodomain of CDCP1 fused to the transmembrane helix of PDGFR upon treatment with tetracycline (Tet) (1ug/mL, 18hrs). Cells were profiled using PhaNGS and those targets showing signal above background have been annotated on the x-axis.



Name	Affinity	NGS signal
GFP.01	0.5nM	161 +/- 20
GFP.02	1nM	93 +/- 6
GFP.03	120nM	2 +/- 0

Figure S7

Phage affinity correlates with observed signal.

To study the relationship between Fab affinity and NGS signal on cells, we created a HeLa line stably over-expressing GFP tethered to the cell surface (HeLaGFP, inset micrographs and diagram). Three anti-GFP antibodies of varying affinities were tested as a pool for binding on cells expressing GFP (n=2, green bar) or control HeLa cells (n=2, grey bar). A background control consisting of a Fab-phage directed to an intracellular zinc-finger transcription factor (ZNF2) was set to 1 to serve as the baseline signal. Error bars shown as standard deviation.

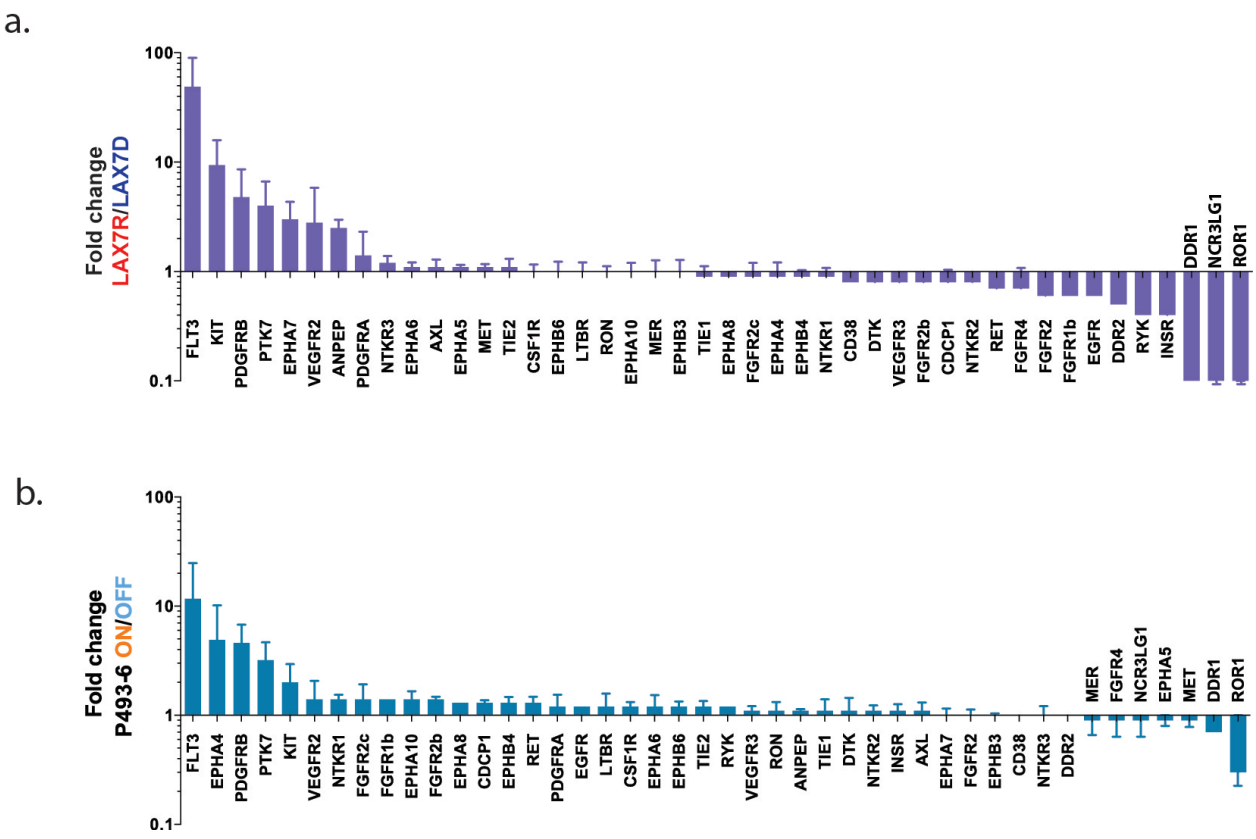


Figure S8

LAX7 and P493-6 fold-change charts.

Charts show average fold-change across similar anti-target clones for all targets from (**a**) LAX7D to LAX7R or (**b**) P493-6 OFF to ON (Fig 2).

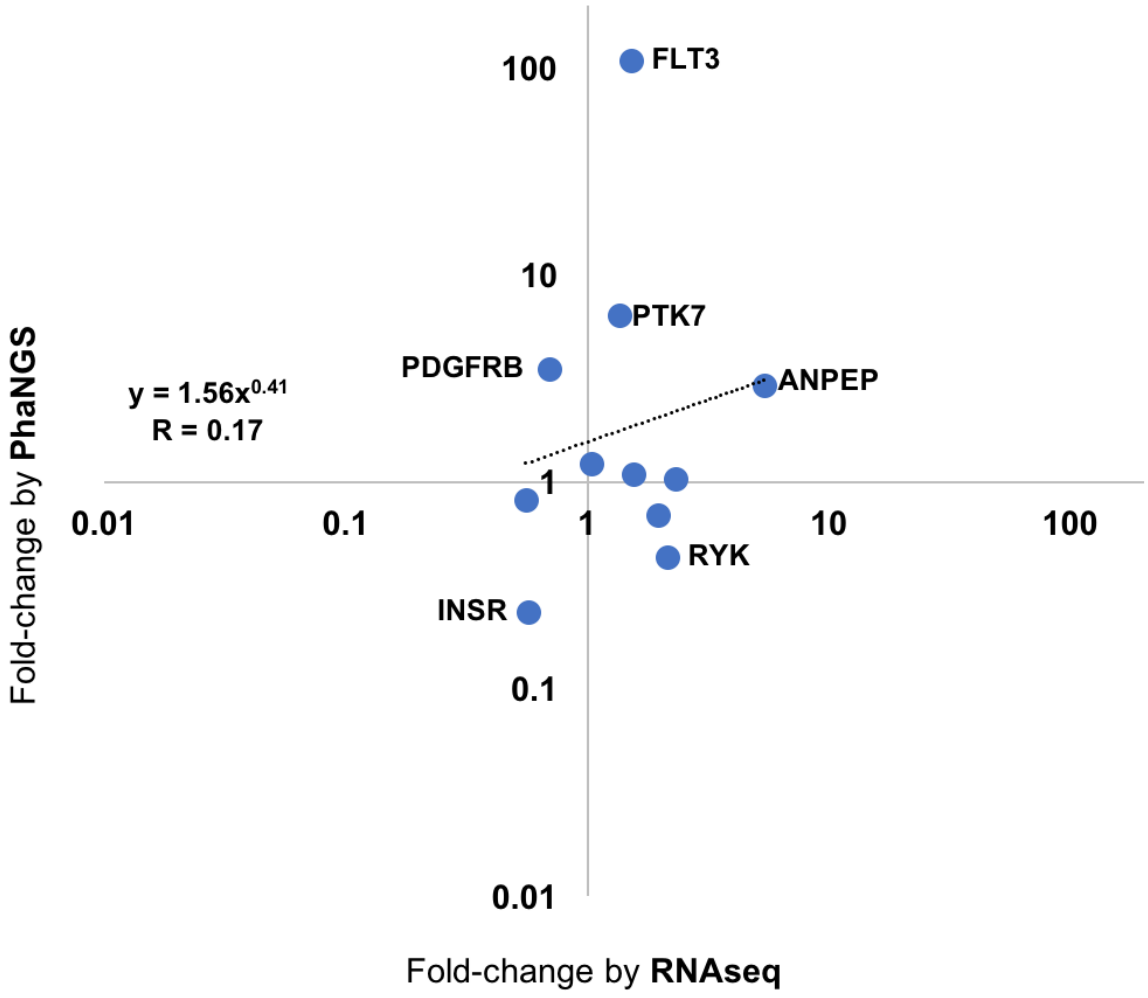


Figure S9

RNA fold-change in abundance compared to protein fold-change abundance by PhaNGS in ALL lines shows little correlation.

A comparison between RNAseq and PhaNGS of target abundance fold-change from LAX7D to LAX7R. Only those targets that showed significant signal (see supplementary files) in both RNAseq and PhaNGS were compared. In ~40% of cases (4/11) the transcriptional change was matched by a similar protein abundance change as assessed by PhaNGS (with the exception of FLT3, which showed very little change by RNAseq but a large change by PhaNGS). In ~30% of

cases (3/11), the direction of transcription change is opposite to the protein abundance change observed by PhaNGS. In the remaining cases, low fold-change values were observed.

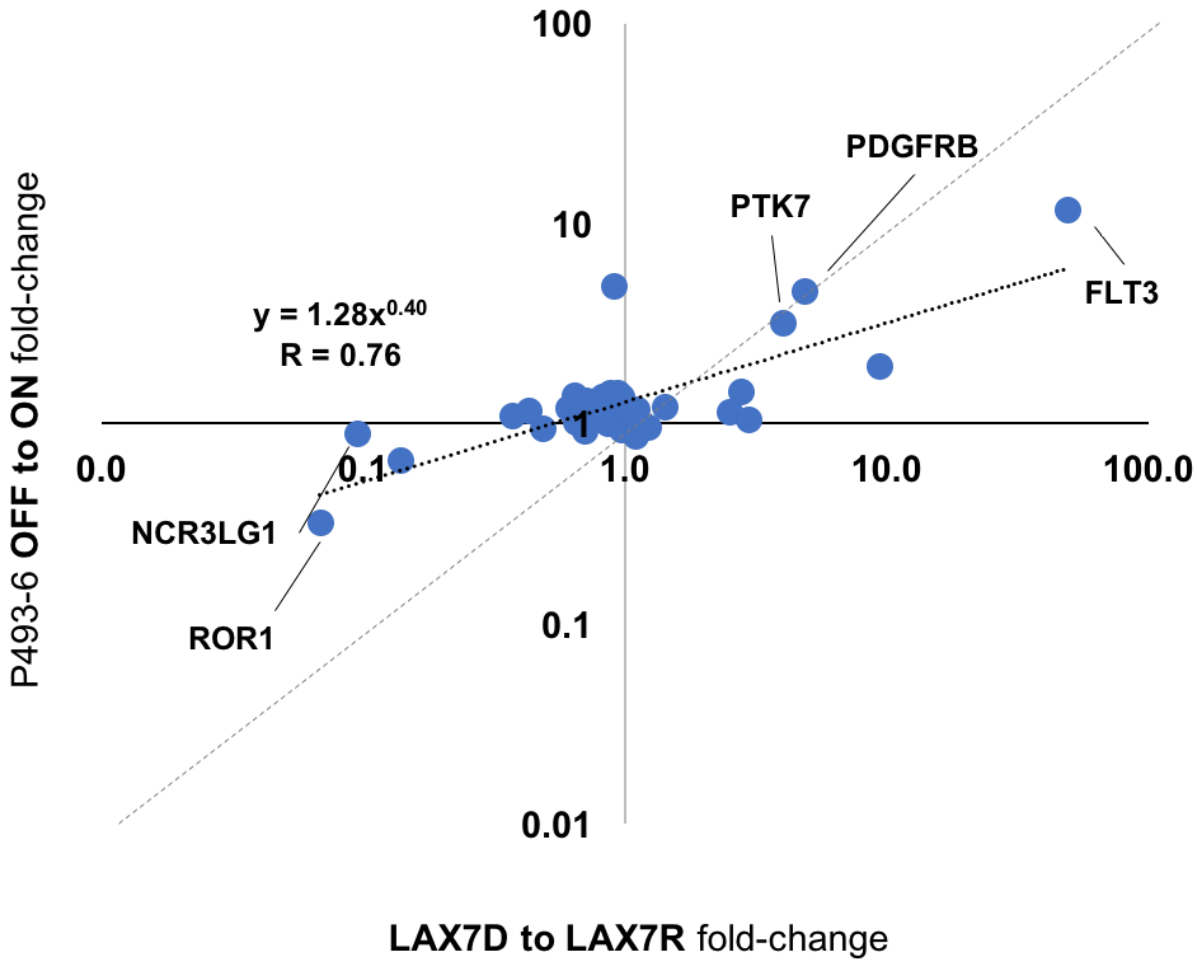


Figure S10

Comparison of protein levels measured by PhaNGS for LAX7 and P493-6 experiments.

To directly compare the changes in LAX7 and P493-6 experiments, we plotted the fold-change in protein abundance for the experiments side-by-side. A correlation was observed (R = 0.76) as targets up- or down-regulated in the LAX7 experiment are similarly modulated in the P493-6 experiment.

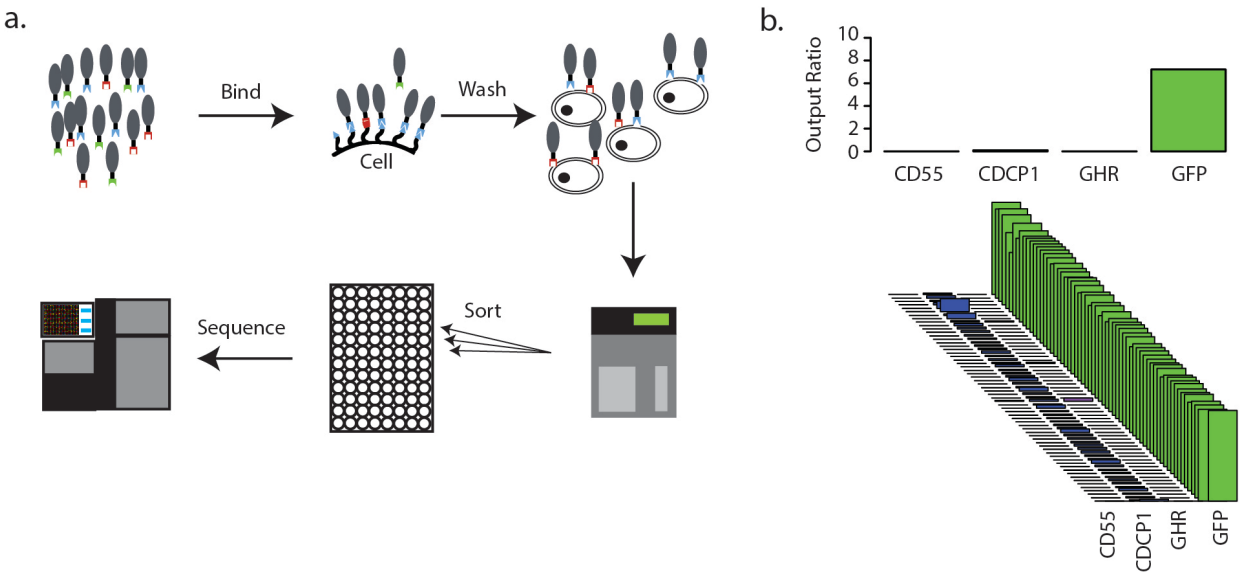


Figure S11

Single cell scheme and HEKGFP single cell PhaNGS data.

(**a**) Schematic diagramming single cell PhaNGS work-up. Most of the work-up is similar to population-scale PhaNGS: a curated input library is added to a cell mixture and non-binders are washed away. In single cell mode, the next step is not to elute the phage off the cells but to sort the cells with Fab-phage on their surfaces into single wells of a 96-well plate containing 2xYT broth. The phage are then propagated overnight by addition of *E.coli* liquid culture, amplified using sequencing primers, and submitted for sequencing. (**b**) In the top panel, a representative graph from a single cell over-expressing GFP (HEKGFP). Four Fab-phage were present in the input library: CD55.01, CDCP1.01, GHR.01, and GFP.01. Signal is shown as the ratio of the counts of phage recovered to the counts of phage added to cells. In the bottom panel, compiled graphs from 62 individual HEKGFP cells which showed adequate counts for data processing, out of a total of 72 wells attempted (86% success rate).

Step	6-well	96-well	576-well
Prepare Library	-	-	-
Bind Phage & Wash	$5.40	$16.20	$113.30
Propagate Eluents	-	-	-
Amplify Propagates	$2.20	$33.10	$198.60
Sequence	$27.50	$140.00	$740.00
Total Cost	$35.10	$189.30	$1,051.90
Cost per Sample	$5.90	$2.00	$1.80

Table S1

PhaNGS cost estimates.

Cost estimate for three different PhaNGS experiments of differing sizes. Cost estimates do not include minor plastic ware or other inexpensive reagents like dNTPs. Estimates assume the use of minimal read count necessary for achieving adequate data quality (~500,000 reads per well). “Prepare Library” and “Propagate Eluents” include primarily phage propagation, for which the costs are insignificant. The “Bind Phage & Wash” step includes the cost of deep-well plastic ware (see SI Materials and Methods for catalog numbers). The “Amplify Propagates” step includes the cost of Phusion polymerase (NEB), master mix, and primers. The “Sequence” step includes the cost of shipping to a sequencing core facility, plus the fraction of a Hiseq4000 lane required for the sample (given a requirement of 500,000 reads per well, assuming $1,000 per lane). Should the cost of sequencing 400 million reads decrease, or should a greater number of reads become capable of being sequenced for the same price, this cost would decrease.

SI Note S1

Fab-phage are present below saturating conditions

The Fab-phage cannot be concentrated to such a level to allow complete saturation of receptors on cells. The upper solubility limit of M13K07 Fab-phage is ~$5x10^{12}$ cfu/mL. This is equivalent to ~8nM. However, the library described here has ~150-members, so each Fab-phage is present at ~50pM. The fact that in phagemid format only ~10% of phage particles display a Fab (12) brings that concentration of Fab phage to 5pM in a 150-member mixture. This is more than 10-fold below the K_d value for the highest affinity Fab-phage (typically K_d ~1nM). Thus, the PhaNGS experiment is in an under-saturation regime relative to most receptors we anticipate detecting on the cell surface.

References Cited

5. Hornsby, M. *et al.* A High Through-put Platform for Recombinant Antibodies to Folded Proteins. *Mol Cell Proteomics* **14**, 2833-2847, doi:10.1074/mcp.O115.052209 (2015).
8. Persson, H. *et al.* CDR-H3 Diversity is Not Required For Antigen Recognition by Synthetic Antibodies. *Journal of molecular biology* **425**, 803-811, doi:10.1016/j.jmb.2012.11.037 (2013).
10. Bass, S., Greene, R. & Wells, J. A. Hormone phage: an enrichment method for variant proteins with altered binding properties. *Proteins* **8**, 309-314, doi:10.1002/prot.340080405 (1990).
11. Aird D*, et al.* (2011) Analyzing and minimizing PCR amplification bias in Illumina sequencing libraries. *Genome Biology* 12(2):R18.
33. Ong SE*, et al.* (2002) Stable isotope labeling by amino acids in cell culture, SILAC, as a simple and accurate approach to expression proteomics. *Mol Cell Proteomics* 1(5):376-386.
34. Wollscheid B*, et al.* (2009) Mass-spectrometric identification and relative quantification of N-linked cell surface glycoproteins. *Nat Biotechnol* 27(4):378-386.
35. Cox J & Mann M (2008) MaxQuant enables high peptide identification rates, individualized p.p.b.-range mass accuracies and proteome-wide protein quantification. *Nat Biotechnol* 26(12):1367-1372.
36. Gibson DG*, et al.* (2009) Enzymatic assembly of DNA molecules up to several hundred kilobases. *Nat Meth* 6(5):343-345.
37. Qu Y*, et al.* (2015) Evaluation of MCF10A as a Reliable Model for Normal Human Mammary Epithelial Cells. *PLOS ONE* 10(7):e0131285.
38. Dobin A & Gingeras TR (2015) Mapping RNA-seq Reads with STAR. *Current protocols in bioinformatics / editoral board, Andreas D. Baxevanis ... [et al.]* 51:11.14.11-11.14.19.

39. Anders S, Pyl PT, & Huber W (2015) HTSeq—a Python framework to work with high-throughput sequencing data. *Bioinformatics* 31(2):166-169.