

# Journal of Visualized Experiments

## Cell surface receptor identification using genome-scale CRISPR/Cas9 genetic screens. --Manuscript Draft--

Article Type:	Invited Methods Article - JoVE Produced Video
Manuscript Number:	JoVE60803R1
Full Title:	Cell surface receptor identification using genome-scale CRISPR/Cas9 genetic screens.
Section/Category:	JoVE Biology
Keywords:	CRISPR/Cas9, recombinant proteins, cell adhesion, cell surface receptor, genome-scale screening, extracellular protein interactions
Corresponding Author:	Sumana Sharma UNITED KINGDOM
Corresponding Author's Institution:	
Corresponding Author E-Mail:	sumana@ebi.ac.uk
Order of Authors:	Sumana Sharma Gavin James Wright
Additional Information:	
Question	Response
Please indicate whether this article will be Standard Access or Open Access.	Open Access (US\$4,200)
Please indicate the <b>city, state/province, and country</b> where this article will be <b>filmed</b> . Please do not use abbreviations.	Hinxton, Cambridgeshire, UK

20<sup>th</sup> September 2019

Dear Dr Stephanie,

I enclose a manuscript entitled **“Cell surface receptor identification using genome-scale CRISPR/Cas9 genetic screens”** to be considered for publication in Journal of Visualised Experiments.

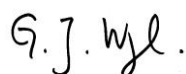
The research in my laboratory over the last decade has been centred on identifying and characterising extracellular interactions made by membrane-embedded receptors that are important for cellular recognition processes. This class of protein interactions are difficult to detect because membrane proteins are difficult to solubilise in their native conformation, and the interactions that they mediate are often extremely transient - typically having half-lives of just fractions of a second. The approach that we and others have adopted in the past have usually focussed on detecting direct interactions between receptor ectodomains expressed as soluble recombinant proteins, which, while broadly successful, has several limitations:

- It is generally only suitable for receptors whose binding functions are retained as soluble recombinant molecules – primarily single-transmembrane-spanning cell surface proteins - and so the role of the many multi-transmembrane-spanning receptors and those receptors that contain multiple polypeptide chains has been largely ignored;
- Interactions that depend on specific posttranslational modifications required that these were added in heterologous recombinant expression systems which is difficult to control;
- Receptor binding domains are necessarily taken out of their native membrane context and so contributions made by the local environment such as the surface glycocalyx are not considered.

To circumvent these technical difficulties, we developed an approach based on genome-wide gene editing using the CRISPR/Cas9 system that is able to elucidate the molecular nature of cell surface receptor-ligand interactions – see Sharma et al. *Genome Res*, 2018. Vol 28: p1372. We have applied this technology to identify novel receptors for a range of biological contexts including interactions important for internalisation of neuronal receptors and cell surface recognitions that regulate platelet functions. The method we have developed, unlike the majority of the existing methods, does not require libraries of recombinant protein and cDNAs, which makes it ideally suited to researchers with focussed interests. We have already had several interests from other groups to apply this method to identify receptors for their specific protein of interest. Therefore, it would be very valuable to have this method presented in a video-format to afford the possibility of successfully transferring this approach to other laboratories.

We believe that the application of this method will make valuable contributions to our mechanistic understanding of how cells communicate with each other. We hope you share our enthusiasm for our work and please don't hesitate to contact me if you require any additional information.

Yours sincerely,



Gavin Wright

**TITLE:****Cell Surface Receptor Identification Using Genome-Scale CRISPR/Cas9 Genetic Screens****AUTHORS AND AFFILIATIONS:**

Sumana Sharma<sup>1,2</sup> and Gavin J Wright<sup>1</sup>

<sup>1</sup>Cell Surface Signalling Laboratory, Wellcome Trust Sanger Institute, Cambridge, UK

<sup>2</sup>EMBL-EBI, Wellcome Genome Campus, Hinxton, Cambridgeshire, UK

**Corresponding Author:**

Sumana Sharma (sumana@ebi.ac.uk)

**Email Addresses of Co-author:**

Gavin J Wright (gw2@sanger.ac.uk)

**KEYWORDS:**

CRISPR/Cas9, recombinant proteins, cell adhesion, cell surface receptor, genome-scale screening, extracellular protein interactions

**SUMMARY:**

This manuscript describes a genome-scale cell-based screening approach to identify extracellular receptor-ligand interactions.

**ABSTRACT:**

Intercellular communication mediated by direct interactions between membrane-embedded cell surface receptors is crucial for the normal development and functioning of multicellular organisms. Detecting these interactions remains technically challenging, however. This manuscript describes a systematic genome-scale CRISPR/Cas9 knockout genetic screening approach that reveals cellular pathways required for specific cell surface recognition events. This assay utilizes recombinant proteins produced in a mammalian protein expression system as avid binding probes to identify interaction partners in a cell-based genetic screen. This method can be used to identify the genes necessary for cell surface interactions detected by recombinant binding probes corresponding to the ectodomains of membrane-embedded receptors. Importantly, given the genome-scale nature of this approach, it also has the advantage of not only identifying the direct receptor but also the cellular components that are required for the presentation of the receptor at the cell surface, thereby providing valuable insights into the biology of the receptor.

**INTRODUCTION:**

Extracellular interactions by cell surface receptor proteins direct important biological processes such as tissue organization, host-pathogen recognition, and immune regulation. Investigating these interactions is of interest to the wider biomedical community, because membrane receptors are actionable targets of systematically delivered therapeutics such as monoclonal antibodies. Despite their importance, studying these interactions remains technically

challenging. This is mainly because membrane-embedded receptors are amphipathic, making them difficult to isolate from biological membranes for biochemical manipulation, and their interactions are typified by the weak interaction affinities ( $K_D$ s in the  $\mu\text{M}$ – $\text{mM}$  range)<sup>1</sup>. Consequently, many commonly used methods are unsuitable to detect this class of protein interactions<sup>1,2</sup>.

A range of methods has been developed to specifically investigate extracellular receptor-ligand interactions that take their unique biochemical properties into consideration<sup>3</sup>. A number of these approaches involve expressing the entire ectodomain of a receptor as a soluble recombinant protein in mammalian or insect cell-based systems to ensure that these proteins contain posttranslational modifications that are structurally important, such as glycans and disulfide bonds. To overcome the low-affinity binding, the ectodomains are often oligomerized to increase their binding avidity. Avid protein ectodomains have been successfully used as binding probes to identify interaction partners in direct recombinant protein-protein interaction screens<sup>4–7</sup>. While broadly successful, recombinant protein-based methods require that the ectodomain of a membrane receptor be produced as a soluble protein. Therefore, it is only generally applicable to proteins that contain a contiguous extracellular region (e.g., single-pass type I, type II, or GPI-anchored) and is not generally suitable for receptor complexes and membrane proteins that span the membrane multiple times.

Expression cloning techniques in which a library of complementary DNAs (cDNAs) is transfected into cells and tested for a gain-of-binding phenotype have also been used to identify extracellular protein-protein interactions<sup>8</sup>. The availability of large collections of cloned and sequenced cDNA expression plasmids in recent years has facilitated methods in which cell lines overexpressing cDNAs encoding cell surface receptors are screened for the binding of recombinant proteins to identify interactions<sup>9,10</sup>. The cDNA overexpression-based approaches, unlike recombinant protein-based methods, afford the possibility to identify interactions in the context of the plasma membrane. However, the success of using cDNA expression constructs depends on the cells' ability to overexpress the protein in the correctly folded form, but this often requires cellular accessory factors such as transporters, chaperones, and correct oligomeric assembly. Transfecting a single cDNA might therefore not be enough to achieve cell surface expression.

Screening techniques using cDNA constructs or recombinant protein probes are resource-intensive and require large collections of cDNA or recombinant protein libraries. Specifically designed mass spectrometry-based methods have been utilized recently to identify extracellular interactions that do not require the assembly of large libraries. However, these techniques require chemical manipulation of the cell surface, which can alter the biochemical nature of the molecules present on the surface of the cells and are currently only applicable for interactions mediated by glycosylated proteins<sup>11, 12</sup>. The majority of the currently available methods also heavily focus on the interactions between proteins while largely ignoring the contribution from the membrane microenvironment, including molecules such as glycans, lipids, and cholesterol.

The recent development of highly efficient biallelic targeting using CRISPR-based approaches has enabled genome-scale libraries of cells lacking defined genes in a single pool that can be

screened in a systematic and unbiased way to identify cellular components involved in different contexts, including dissecting cellular signaling processes, identification of perturbations that confer resistance to drugs, toxins, and pathogens, and determining specificity of antibodies<sup>13–16</sup>. Here, we describe a genome-scale CRISPR-based knockout cell screening assay that provides an alternative to the current biochemical approaches to identify extracellular receptor-ligand interactions. This approach of identifying interactions mediated by membrane receptors by genetic screens is particularly suitable for researchers that have a focused interest on individual ligands because it avoids the need to compile large libraries of cDNAs or recombinant proteins.

This assay consists of three major steps: 1) Highly avid recombinant protein binding probes consisting of the extracellular regions of a receptor of interest are produced and used in fluorescence-based flow cytometry-based binding assays; 2) The binding assays are used to identify a cell line that expresses the interaction partner of the recombinant protein probe; 3) A Cas9-expressing version of the cell line that interacts with the protein of interest is produced and a genome-scale CRISPR/Cas9-based knockout screen is performed (**Figure 1**). In this genetic screen, binding of a recombinant protein to cell lines is used as a measurable phenotype in which cells within the knockout library that have lost the ability to bind the probe are sorted using fluorescence-based activated cell sorting (FACS) and the genes that caused the loss of the binding phenotype identified by sequencing. In principle, the genes encoding the receptor responsible for binding the avid probe and those required for its cell surface display are identified.

The first step of this protocol involves the production of avid recombinant protein probes representing the ectodomain of the membrane-bound receptors. These receptors are known to frequently retain their extracellular binding functions when their ectodomains are expressed as a recombinant soluble protein<sup>1</sup>. For a protein of interest, soluble recombinant proteins can be produced in any suitable eukaryotic protein expression system in any format provided that it can be oligomerized for increased binding avidity, and it contains tags that can be used in fluorescence-based flow cytometry-based binding assays (e.g., FLAG-tag, biotin-tag). Detailed protocols for the production of soluble ectodomains of membrane receptors using the HEK293 protein expression system, as well as different multimerization techniques and the protein expression constructs for the production of both pentameric proteins and monomeric proteins have been previously described<sup>1,17</sup>. The protocol here will describe the steps for generating fluorescent avid probes from monomeric biotinylated proteins by conjugating them to streptavidin conjugated to a fluorochrome (e.g., phycoerythrin, or PE), which can be used directly in cell-based binding assays and has the advantage of not requiring a secondary antibody for detection. General protocols for performing genome-scale screens have already been described<sup>20,21</sup>, thus the protocol mainly focus on the specifics of performing flow cytometry-based recombinant protein binding screens using the CRISPR/Cas9 knockout screening system using the Human V1 (“Yusa”) library<sup>18</sup>.

## **PROTOCOL:**

### **1. Production and purification of biotinylated His-tagged proteins**

1.1. Use a mammalian or insect cell-based protein expression system to produce soluble recombinant His-tagged biotinylated proteins (see plasmid constructs in **Table 1**).

NOTE: A detailed protocol for the production of monomeric biotin and His-tagged proteins using the HEK293 cell expression system is described by Kerr et al.<sup>17</sup>. Protein ectodomains expressed using the HEK293 expression system are secreted into the culture medium.

1.2. Collect the soluble proteins by pelleting the cells by centrifugation at 3,000 x *g* for 20 min.

1.3. Filter the supernatant through a 0.22 µm filter and add the Ni<sup>2+</sup>-NTA agarose beads to the filtered protein supernatant in a 1:1,000 ratio (i.e., 50 µL of 50% agarose slurry into 50 mL of supernatant). Incubate overnight or at least 4–5 h at 4 °C on a rotating platform.

1.4. Wash the polypropylene column by adding 5 mL of His-purification wash buffer. Refer to **Table 2** for all buffer compositions.

1.5. Pour the entire bead-protein supernatant mixture into the column. Beads will accumulate at the base.

1.6. Wash the beads 2x with 15 mL of wash buffer. To avoid protein dilution, carefully draw the residual wash buffer from the column with a 5 mL syringe and discard.

1.7. Carefully add 300–500 µL of His-purification elution buffer directly to the beads and incubate for at least 1 h. Collect the eluted protein by again carefully drawing out the liquid using a 1 mL syringe. Exchange the elution buffer to the desired buffer (e.g., normally PBS or HBS) using desalting columns. Store all proteins at 4 °C until further use.

## 2. Quantification and oligomerization of monomeric biotinylated protein

NOTE: To increase the binding avidity, oligomerize biotinylated monomeric proteins on tetrameric streptavidin-PE before using them in binding assays. Achieve optimal conjugation ratios of monomeric proteins and tetrameric streptavidin-PE by testing a dilution series of biotinylated monomers against a fixed concentration of streptavidin and by empirically establishing the minimum dilution at which no excess biotinylated monomers can be detected.

2.1. Make at least eight serial dilutions of biotinylated protein samples using an appropriate dilution buffer (either PBS or HBS with 1% bovine serum albumin [BSA]) in a 96 well plate. Ensure that the final volume of each dilution is at least 200 µL.

2.2. Make a duplicate plate of the samples by removing 100 µL from each well and transferring into a new 96 well plate. Always include a control. In this case controls are tag-only proteins (i.e., biotinylated His-tagged Cd4 domain 3+4 protein). This will be used as a control probe in all binding assays.

2.3. Dilute streptavidin-PE to 0.1 µg/mL in the dilution buffer.

2.4. On just one of the plates, add 100 µL of the diluted streptavidin-PE. The duplicate plate will serve as a control. Add 100 µL of dilution buffer in the control plate to equalize volumes.

2.5. Incubate for 20 min at room temperature (RT). In the meantime, block the wells of a streptavidin-coated plate with the dilution buffer for 15 min.

2.6. Transfer the total volume of the sample from both plates to individual wells of the streptavidin-coated plates and incubate for 1 h at RT.

2.7. Wash the plate 3x with 200 µL of wash buffer (i.e., either PBS or HBS with 0.1% Tween-20, 2% BSA). Add 100 µL of 2 µg/mL mouse anti-rat Cd4d3+4 IgG (OX68) and incubate for 1 h at RT.

2.8. Wash the plate 3x with the wash buffer. Add 100 µL of an anti-mouse alkaline phosphatase conjugate at 0.2 µg/mL for 1 h at RT.

2.9. Wash the plate 3x with wash buffer and 1x in dilution buffer.

2.10. Prepare p-nitrophenyl phosphate at 1 mg/mL in diethanolamine buffer. Add 100 µL in each well and incubate for 15 min.

2.11. Take absorbance readings at 405 nm. Use the minimum dilution at which there is no signal on the plate as the appropriate dilution factor to create tetramers (**Figure 2**).

2.12. Make a 10x tetramer staining solution for all samples and controls by incubating 4 µg/mL streptavidin-PE and the appropriate biotinylated protein dilution for 30 min at RT. Store conjugated proteins in a light-protected tube at 4 °C until further use.

### 3. Flow cytometry-based cell binding assays

3.1. For adherent cells, remove culture media and wash 1x with PBS without divalent cations. Then add cell detachment solutions (e.g., EDTA). Allow the cells to detach for 5–10 min. Gently tap the flask to release the cells.

NOTE: Avoid using trypsin-based products as they can cleave cell surface proteins.

3.2. Collect detached cells into a tube. For cells growing in suspension (e.g., HEK293 cells), directly collect the cells from culture flasks into a tube.

3.3. Pellet cells at 200 x g for 5 min. Remove the supernatant and resuspend the pellet in wash buffer (i.e., PBS/1% BSA).

3.4. Count the cells using a hemocytometer and adjust the concentration to  $2.5 \times 10^5$ – $1 \times 10^6$  cells/mL. Aliquot 100  $\mu$ L of prepared cell mix on a 96 well U- or V- bottomed plate. Spin the plate for 5 min at 400 x *g*. Remove the supernatant with a multichannel pipette.

3.5. Add 100  $\mu$ L of normalized fluorescently labeled highly avid protein probes and controls into the previously prepared plates with cells and incubate for 1 h at 4 °C. After binding for 1 h, spin the plate at 400 x *g* for 5 min.

3.6. Remove the supernatant and add 200  $\mu$ L of wash buffer (i.e., PBS/1% BSA). Mix well by pipetting up and down.

3.7. Pellet the cells by centrifugation at 400 x *g* for 5 min. Repeat the wash step 1x. After two washes, completely remove the supernatant and resuspend the cell pellet in 100  $\mu$ L of PBS.

3.8. Analyze the cells by flow cytometry. Use the yellow-green laser (i.e., 561 nm) to detect PE fluorescence.

3.8.1. First analyze the cells that have been stained with the control probe. Based on the distribution of PE fluorescence, draw a gate for binding population such that no more than 1% of the control cell falls in this gate.

3.8.2. Analyze the sample and determine the fraction of cells that falls in the binding gate.

NOTE: Cell lines that display a higher binding population are desired for genetic screens, as they have a higher signal-to-noise ratio. Ideally over 80% of the cells should fall within this gate.

#### **4. Determining binding contributions from heat labile epitopes and heparan sulphate sidechains**

NOTE: The activity of many proteins is heat labile, so loss of binding activity following heat treatment is encouraging. It is advised to determine the contribution from negatively charged glycosaminoglycans, mainly heparan sulphate (HS), in mediating binding of the recombinant proteins. This is because the binding by HS in the cell binding assay described here can be additive rather than codependent on other receptors<sup>19</sup>. This means that the observed binding could be entirely mediated by HS side chains of cell surface proteoglycans and not by a specific receptor. Binding to HS on the cell surface is not necessarily nonspecific, but rather a property of a protein, which is useful to know before performing a full genetic screen.

4.1. Prepare heat-treated protein samples to use in binding assays.

4.1.1. Heat the normalized but unconjugated monomeric protein at 80 °C for 10 min.

4.1.2. Conjugate the heat-treated protein to streptavidin-PE assuming the same conjugation ratio as its untreated counterpart as determined by ELISA (refer to section 2).

264  
265 4.2. Prepare heparin-blocked protein samples.

266  
267 4.2.1. Prepare eight 1:3 dilutions of soluble heparin in PBS with a starting concentration of 2  
268 mg/mL and final volume of 100  $\mu$ L.

269  
270 4.2.2. Incubate 100  $\mu$ L of prepared binding probes in the heparin dilutions for at least 30 min.

271  
272 4.3. Use heat-treated protein and the full 200  $\mu$ L of heparin/protein mixture in the binding assays  
273 described in section 3. Representative results are shown in **Figure 3A,B**.

## 274 275 **5. Selection of cell lines stably expressing Cas9**

276  
277 NOTE: Before the cell line that binds the probe of interest can be used in CRISPR screening, it  
278 must first be engineered to express the Cas9 nuclease and a highly active clone selected<sup>19</sup>.

279  
280 5.1. Use the following general lentivirus production protocol to produce lentivirus using the  
281 lentiviral construct for Cas9 expression (refer to **Table 1**).

282  
283 5.1.1. Culture HEK293-FT cells in DMEM/10% FBS media at 37 °C and 5% CO<sub>2</sub>. Seed HEK293-FT  
284 cells 1 day prior to transfection so that they are ~80% confluent on the day of transfection.

285  
286 NOTE: HEK293FT cells are loosely adherent; therefore, when they are used for production of  
287 lentiviruses, consider plating them on culture flasks coated with 0.1% (w/v) gelatin to increase  
288 adherence.

289  
290 5.1.2. Perform transfections in the morning. Add transfer vector, packaging mix, and transfection  
291 reagent into prewarmed transfection compatible media (e.g., Opti-MEM). Mix by inverting the  
292 tube 10–15x. Incubate for 5 min at RT. Refer to **Table 3** for exact volumes.

293  
294 5.1.3. Add the transfection reagent as suggested by the manufacturer. Mix by quick vortexing.  
295 Incubate for 30 min at RT.

296  
297 5.1.4. Very carefully aspirate the spent medium. Add transfection compatible media to plate.

298  
299 5.1.5. Add the transfection reagent/DNA complexes dropwise onto the side of the plate and  
300 slowly spread through the plate by swirling very gently.

301  
302 5.1.6. Incubate at 37 °C for 3–5 h and replace the medium with D10 medium. Incubate overnight.

303  
304 5.1.7. Next day in the morning, replace the medium with fresh D10 medium. Incubate overnight.  
305

5.1.8. Next day late in the afternoon, collect the viral supernatant. Filter with a 0.45 µm filter with low protein binding. Optionally, add fresh D10 medium, incubate overnight and recollect the supernatant the next day.

5.1.9. Virus supernatants are stable at 4 °C for only for a few days. Store at -80 °C for long-term storage.

NOTE: To generate a highly concentrated lentiviral preparation, which could be desirable for the transduction of difficult to transduce cells, supernatants can also be concentrated by centrifugation at 6,000 x *g* overnight at 4 °C. Mark the translucent viral pellet with an ethanol-resistant pen and discard the supernatant. Resuspend the pellet in 1/100th of the original volume for a 100x increase in the concentration.

5.2. Transduce the cells with lentiviruses.

5.2.1. Plate 1 x 10<sup>6</sup> cells per well in a 6 well plate with 3 mL of appropriate culture media. Some cells are more readily transduced than others. For easy to transduce cells (e.g., HEK cells), directly add the lentivirus to the cells. For difficult to transduce cells, it might be necessary to follow a spinoculation protocol as described below.

5.2.1.1. Aliquot 2 mL of 2–5 x 10<sup>6</sup> cells/mL in a 15 mL conical tube.

5.2.1.2. Add lentivirus together with 8 µg/mL hexadimethrine bromide and incubate at RT for 30 min.

5.2.1.3. Centrifuge for 100 min at 800 x *g* at 32 °C. Then resuspended the cells in the same media and add the cell suspension into appropriate culture flasks with appropriate media.

5.2.2. Allow transductions for at least 24 h. Afterwards remove the media containing the virus and add fresh medium.

5.2.3. After another 24 h, change the media to one that is supplemented with the appropriate antibiotics. The Cas9 construct contains a blasticidin resistance cassette for selection.

NOTE: The amount of blasticidin must be optimized for each cell line by performing a dose response kill curve. A blasticidin concentration between 2.5–50 µg/mL should kill most untransduced cell lines within 10 days of transduction.

5.3. Perform selection until all cells in the control plate (i.e., nontransduced cells that have been treated with the same concentration of selection antibiotics) are killed.

## 6. Selecting high Cas9-activity clones

NOTE: Polyclonal Cas9 can be used to successfully perform genetic screens; however, selecting a clone with high Cas9 activity improves the screening results<sup>18</sup>.

6.1. Use limiting dilution or single-cell sort individual blasticidin-resistant cells into wells of three 96 well plates containing culture media supplemented with blasticidin. Clones will start to emerge between 2–4 weeks. Select 10–20 clones and expand in 6 well plates.

6.2. Assay the clones for Cas9 activity using the quick-to-assess GFP-BFP (green fluorescent protein-blue fluorescent protein) system, which uses an exogenous gene knockout system in which cells are transduced either with a construct expressing GFP with a gRNA-targeting GFP or an empty gRNA as a control<sup>18</sup>.

6.2.1. Order reporter plasmids: GFP-BFP plasmid, Control-BFP plasmid (**Table 1**).

6.2.2. Produce lentivirus for both the GFP-BFP plasmid and Control-BFP plasmid using the lentivirus production protocol described in section 5.1.

6.2.3. Transduce each Cas9-expressing cell line clone with the lentivirus encoding the GFP-BFP system and Control-BFP separately. Follow the protocol in section 5.2.

6.2.4. After 3 days of transduction, examine the GFP-BFP fluorescence of each clone using flow cytometry. Use 488 nm laser and 405 nm laser to detect GFP and BFP respectively.

6.2.5. Quantitate the Cas9 activity in each clone by examining the ratio of BFP only to GFP-BFP-double positive cells. High activity Cas9 cells should ideally have >95% GFP knockout efficiency (**Figure 4**).

## 7. Generation of genome-wide CRISPR-Cas9 screening knockout library

7.1. For the genome-wide screening using the Human V1 library<sup>18</sup>, order the genome-wide library (refer to **Table 1**) and prepare the plasmid library from the bacterial stab using the protocol provided under "Protocols for Library Replication" in the manufacturer's manual.

7.2. Use the genome-wide library plasmid preparation to produce a lentiviral library encoding gRNAs for targeted disruption of human genes using the lentivirus production protocol described in section 5.1.

NOTE: A good practice is to produce a single batch of lentiviral preparation that is optimized for transduction to improve experimental consistency.

7.3. Use the transduction protocol in section 5.2 to perform small scale test transductions to determine the required amount of virus for each cell line to achieve 30% transduction. Use flow cytometry to assess BFP fluorescence as a proxy for transduction efficiency.

7.4. To transduce HEK293 cells, simply add the predetermined lentiviral preparation to 30–50 x 10<sup>6</sup> cells cultured in normal growth media for ~4 h. Then remove the media with lentivirus and replace with fresh growth media.

7.5. For other cell lines, use the spinoculation protocol in section 5.2.1 but at a larger scale, such that a total of 30–50 x 10<sup>6</sup> cells are transduced. For this, aliquot 2 mL of 5 x 10<sup>6</sup> cells/mL in a 15 mL conical tube and proceed as indicated.

7.6. For adherent cell lines, select transduced cells by adding puromycin 24 h after transduction.

NOTE: Optimize puromycin concentrations by performing a dose response kill curve. Normally concentrations between 1–10 µg/mL should kill nontransduced cells within 3–5 days. Avoid using higher concentrations of puromycin because this may increase the chances of selecting cells that have been transduced by more than one single guide RNA (sgRNA).

7.7. For suspension cells, harvest transduced (i.e., BFP positive) cells 3 days posttransduction using a cell sorter and generate libraries that contain at least 10 x 10<sup>6</sup> cells. Once selected using BFP, grow the cells in media supplemented with appropriate amount of puromycin.

NOTE: Avoid selections only with puromycin for suspension cell lines, because it is difficult to remove dead cells and debris from suspension cell cultures that can interfere with cell sorting.

7.8. Culture mutant library for 9–16 days posttransduction with regular passage every 2–3 days.

## **8. Genetic screening for cell surface binding**

8.1. Pellet the mutant cell library at 200 x *g* for 5 min and resuspend the cells in PBS.

8.2. Divide the cells into two 15 mL conical tubes with at least 50 x 10<sup>6</sup> cells in each tube.

8.3. Spin one conical tube at 200 x *g* for 5 min, remove the supernatant, and freeze the cell pellet at -20 °C. This is the control population and will be processed later.

8.4. Resuspend the pellet in the other tube in 10 mL of PBS/1% BSA. Set aside 100 µL of cells as a negative control on a 96 well plate.

8.5. Add the appropriate preconjugated recombinant protein to the cell suspension in the conical tube and negative control proteins to the 96 well plate.

8.6. Perform cell staining for at least 1 h at 4 °C on a benchtop rotor with gentle rotation (6 rpm).

8.7. Pellet the cells at 200 x *g* for 5 min, remove the supernatant. Perform two wash steps, then resuspend the cells in 5 mL of PBS.

8.8. Strain the cells through a 30  $\mu\text{m}$  cell strainer to remove cell clusters. Analyze using a flow sorter.

8.9. Use the negative control sample to gate for BFP+/PE- cells.

8.10. Sort the sample and collect the BFP+/PE- cells. The sort gates will depend on the binding of the cells to the protein but is normally 1–5% of the PE negative samples are collected. An example of a sorting gate is provided in **Supplementary Figure 1**.

8.11. Collect 500,000–1,000,000 cells from the selected gate. Given the low number of cells, consider collecting the samples in a 1.5 mL centrifugation tube to minimize losses.

8.12. Pellet the sorted cells by centrifuging at 500  $\times g$  for 5 min. Carefully remove the supernatant and discard. It is possible to store the pellet at  $-20\text{ }^{\circ}\text{C}$  for up to 6 months.

## 9. Genomic DNA extraction and first PCR for gRNA enrichment

9.1. Extract genomic DNA from high-complexity control population.

9.1.1. If the control population was frozen at  $-20\text{ }^{\circ}\text{C}$ , take out the conical tube and add PBS. Keep on ice to thaw the pellet.

9.1.2. Use a commercial kit (see **Table of Materials**) using the manufacturer's recommendations to extract genomic DNA from  $50 \times 10^6$  cells. Adjust the DNA concentration to 1 mg/mL.

9.1.3. For each sample, set up a master mix for PCR corresponding to 72  $\mu\text{g}$  of DNA. Aliquot 50  $\mu\text{L}$  per well in 36 wells of a 96 well PCR plate. The necessary primer sequences are listed in **Table 4**. Use the guide in **Table 5** and **6**.

9.1.4. Resolve 5  $\mu\text{L}$  of the PCR from 6–12 representative samples on a 2% (w/v) agarose gel. A single clear band at  $\sim 250$  bp should be observed. If the bands are faint, repeat the PCR for an additional 2–3 cycles.

9.1.5. Use a multichannel pipette to collect 5  $\mu\text{L}$  of PCR products from each well (180  $\mu\text{L}$  in total) and pool them in a reservoir with 900  $\mu\text{L}$  of binding buffer from a commercial kit (see **Table of Materials**).

9.1.6. Purify the PCR products using a commercial PCR purification kit. Elute DNA into 50  $\mu\text{L}$  of elution buffer from a commercial kit (see **Table of Materials**) and measure the DNA concentration.

9.2. Samples that have been sorted for the loss of binding phenotype are unlikely to be composed of a large number of independent clones. Therefore, it is not necessary to perform PCR with 72  $\mu\text{g}$  of DNA. Isolate the DNA using an appropriate commercial kit (see **Table of Materials**). Set up

3–4 PCR reactions using the protocol described before (section 9.1.3) with 100 ng/μL DNA. If the sorted cell number is less than 100,000 use cell lysates instead of genomic DNA preparations.

9.2.1. Aliquot approximately 10,000 cells/well in a 96 well PCR plate.

9.2.2. Pellet the cells in the plate and carefully remove most of the supernatant. The pellet will not be visible.

9.2.3. Add 25 μL of water in each well and heat the samples at 95 °C for 10 min.

9.2.4. Add 5 μL of 2 mg/mL freshly diluted proteinase K to each well for 1 h and incubate at 56 °C. Then heat the sample for 10 min at 95 °C to inactivate the proteinase K.

9.2.5. Use 10 μL of cell lysate mixture per PCR reaction. Lysates should be used within 24 h.

## **10. Second round of PCR for index barcoding and sequencing**

10.1. Dilute the products from the first round PCR to 40 pg/μL.

10.2. Set up one PCR per sample (use the guide provided in **Tables 7** and **8**). The use of a high-fidelity polymerase is important to minimize errors introduced by the polymerase during sgRNA amplification.

10.3. Resolve 5 μL of PCR products on a 2% (w/v) agarose gel. A single clear band at ~330 bps should be observed.

10.4. Purify PCR products using paramagnetic beads by adding 31.5 μL of (0.7x total volume) of resuspended beads to the PCR products, mixing well, and incubating for 5 min at RT.

10.5. Place the tube on a magnetic rack for 3 min. The beads should be captured on the side of the plate and the solution should be clear. Carefully remove and discard the supernatant.

10.6. Add 150 μL of 80% freshly prepared ethanol to the tube. Incubate for 30 s, and then carefully remove and discard the supernatant.

10.7. Repeat step 13.6, this time with 180 μL. Then air dry the beads for 5 min.

10.8. Remove the tube from the magnet. Elute DNA target from beads into 35 μL of sterile EB buffer. Incubate for 3 min, then put the tube back in the magnetic rack for 3 min.

10.9. Transfer approximately 30 μL of the supernatant containing the eluted PCR products to a clean tube.

10.10. Sequence the samples on a next-generation sequencing platform. For the HumanV1 gRNA library, use the custom primer listed in **Table 4** to sequence 19 bp.

## **11. Bioinformatics analysis to identify the receptor and related pathways**

11.1. Map sequences from sorted and unsorted population to the reference library using the count function of MAGeCK. The function will yield a raw count file (**Supplementary Table 1**).

NOTE: Detailed instructions on the installation of MAGeCK and the usage of different functions within MAGeCK are described in a previously published protocol by Wang et al.<sup>20</sup>.

11.2. Check the technical standard of the control library used in the screen.

11.2.1. Median normalize the raw counts and use the **ggplot2** package in R<sup>21</sup> or equivalent software to plot an empirical cumulative density function plot of the counts in plasmid and control unsorted samples (**Figure 5A**).

11.2.2. Run the **-test** function of MAGeCK using counts from plasmid population as "control" and the counts from unsorted control samples as the "test" sample. The function will yield a gene summary file (**Supplementary Table 1**).

11.2.3. Open the gene summary file and draw the distribution of log-fold-changes (**neg|lfc** column) for previously categorized essential and nonessential genes<sup>22</sup> (**Figure 5B**).

11.2.4. Select significantly depleted genes (**neg|fdr** < 0.05) and perform pathway enrichment analysis using the **enrichr**<sup>23</sup> package or any equivalent pathway enrichment packages in R (**Figure 5C**).

11.3. Run the **-test** function of MAGeCK with default setting. Use raw counts from unsorted control sample as "**control**" and counts from sorted sample as "**treatment**" when performing the analysis.

11.4. Open the gene summary file generated by MAGeCK and rank the **pos|rank** column in ascending order. Use **FDR (pos|fdr** column) < 0.05 as a cutoff for identification of hits. The receptor is usually ranked highly, often in the first position.

11.5. Plot the Robust-Ranking-Algorithm (RRA) scores for positive selection (**pos|score**) in R or an equivalent software (**Figure 5D**).

11.6. Select the gene hits and perform pathway enrichment analysis using the **enrichr** package or any equivalent pathway enrichment packages in R to identify the enriched pathways.

## **REPRESENTATIVE RESULTS:**

Sequencing data from two representative genome-scale knockout screens for the identification of the binding partner of human TNFSF9 and *P. falciparum* RH5 performed in NCI-SNU-1 and HEK293 cells respectively are provided (**Supplementary Table 1**). The binding behavior of RH5 was affected by both heparan sulphate and its known receptor BSG<sup>24</sup> (**Figure 3C**), whereas TNFSF9 specifically bound to its known receptor TNFSF9 and did not lose binding upon preincubation with soluble heparin. Protein 3 in **Figure 3B** represents TNFSF9.

For both cell lines, the distribution of gRNAs in the control mutant library after 3 days (9, 14, and 16 days posttransduction) are also provided (**Supplementary Table 1**). The gRNA distribution revealed that the library complexity was maintained throughout the course of the experiment (**Figure 5A**). The genetic screen for the identification of the ligand for TNFSF9 was performed on day 14 posttransduction, whereas that for RH5 was performed day 9 posttransduction. The technical quality of the screens was assessed by examining the distribution of observed fold-changes of gRNAs targeting a reference set of nonessential genes compared to the distribution for reference set of essential genes<sup>22</sup> (**Figure 5B**). In addition, pathway-level enrichment also revealed that expected essential pathways were identified and significantly enriched in the "drop-out" population when comparing the control sample to the original plasmid library. An example with day 14 NCI-SNU-1 sample is depicted in **Figure 5C**.

The distribution of the gRNAs in the control versus sorted population using the **-test** function of MAGeCK (see **Supplementary Table 1** for the gene summary output from MAGeCK) was used to identify the receptor from the phenotypic screens. The modified RRA score reported by MAGeCK in the gene-level analysis is plotted against the genes ranked by p values. The RRA score in MAGeCK provides a measure in which gRNAs are ranked consistently higher than expected. In the screen for TNFSF9, the top hit was TNFSF9, which is a known binding partner of TNFSF9 (**Figure 5D**). In addition, a number of genes related to the TP53 pathway were also identified. In the case of RH5, in addition to the known receptor (*BSG*) and the gene required for the production of the sulfated GAGs (*SLC35B2*), an additional gene (*SLC16A1*) was also identified (**Figure 5E**). *SLC16A1* is a chaperone required for trafficking BSG to the surface of cells<sup>25</sup>. Together, these results demonstrate the ability of the screen to identify directly interacting receptors and the cellular components required for that receptor to be expressed on the surface of the cells in a functional form.

#### **FIGURE AND TABLE LEGENDS:**

**Figure 1: Overview of the genetic screening approach to identify cell surface receptors.** This assay consists of three major steps: First, recombinant proteins representing the ectodomain of cell surface receptors are expressed in a cell line that can add structurally critical posttranslational modifications such as HEK293 cells. Monomeric protein ectodomains are oligomerized by conjugating to streptavidin-PE to increase their binding avidity. Second, these avid probes are used in cellular binding assays where bright staining on the cell lines indicated by a prominent shift in PE fluorescence (in green) compared to a negative control protein (in black) demonstrates the presence of a cell surface binding partner. Third, receptor-positive Cas9-expressing cell lines are selected and genome-scale screening using gRNAs targeting the vast majority of protein-coding genes is performed. While generating mutant libraries, it is

common to use 30% transduction efficiency, which is based on the Poisson distribution probability that ensures each cell receives a single gRNA such that the resultant phenotype is attributed to a specific knockout. The BFP marker expressed by the transduced cells is used to select cells containing gRNAs using FACS. Phenotypic screens are performed between 9–16 days posttransduction. On the day of the screen, the total mutant cell population is divided into two. One half is kept as the control population and the other half is selected for recombinant protein binding. The cells from the mutant library that are no longer able to bind the recombinant protein are sorted using FACS and the enrichment of gRNAs in the sorted versus control population is used to identify genes required for cell surface binding of the labeled avid probe. Steps in the protocol that require considerable time are indicated. This figure has been modified from Sharma et al.<sup>19</sup>.

**Figure 2: Establishing the ratios of biotinylated protein to streptavidin-PE using an ELISA-based method.** An example of streptavidin-PE conjugation strategy used to generate an avid probe from a biotinylated monomeric protein. A dilution series of biotinylated monomers was incubated against a fixed concentration of streptavidin. The minimum dilution at which no excess biotinylated monomers can be detected was determined by ELISA. ELISA was performed with or without preincubating a range of protein dilutions with 10 ng of streptavidin-PE. In the presence of streptavidin-PE, the minimum dilution at which no signal was identified (circled black) and the amount of protein required for the saturation was calculated to generate a 10x stock solution with 4 µg/mL streptavidin-PE.

**Figure 3: Representative binding of proteins to cell lines.** (A) Protein binding to cell lines had a clear increase in cell-associated fluorescence compared to the control sample. Heat treatment (80 °C for 10 min) of recombinant protein abrogated all binding back to a negative control, demonstrating that the binding behavior was dependent on correctly folded protein. (B) Different classes of protein binding behavior to cell surfaces; dependence on GAGs. From left to right, the proteins can be classified into three types: Protein type 1 only adsorbs to HS. These proteins lose their binding after preincubation with heparin concentrations over 0.2 mg/mL. Protein type 2 binds to HS in addition to a specific receptor. These proteins lose partial binding in the preblocking experiments. Protein type 3 does not bind HS. These proteins do not lose binding compared to parental lines. (C) An example of a protein (i.e., RH5) that binds to HS and a specific receptor in an additive manner. Targeting either the receptor (i.e., BSG) or enzymes required for HS synthesis (e.g., SLC35B2, EXTL3) only partially reduces the binding of RH5 to cells relative to controls. Transduced polyclonal lines can be used in such experiments to establish binding behavior. This figure has been modified from Sharma et al.<sup>19</sup>.

**Figure 4: Selecting clonal cell lines with high Cas9 activity.** Genome-editing efficiency of both polyclonal and cloned lines of NCI-SNU-1 cell lines were assessed using the GFP-BFP reporter system, in which cell lines were transduced with viruses with a gRNA-targeting plasmid encoded GFP or without (i.e., "empty"). A schematic is depicted. Flow cytometry was used to test both BFP and GFP expression after transduction and compared to uninfected control. GFP expression was used as a proxy for Cas9 activity, whereas BFP expression marked transduced cells. The profile for uninfected and empty infected cells looked similar for all clones. Representative

profiles are depicted in the left panel. All five clones of the NCI-SNU-1 cell line showed a higher loss of GFP compared to the polyclonal line (right panel), with clone 4 showing the highest efficiency with the lowest refractory population. This figure has been modified from Sharma et al.<sup>19</sup>.

**Figure 5: Representative results from genetic screens for the identification of the cell surface binding partners.** (A) Cumulative distribution function plots comparing the gRNA abundance in the plasmid library to the mutant libraries of HEK-293-E and NCI-SNU-1 cells on day 9, 14, and 16 days posttransduction. For any given number, cumulative density function reports the percent of datapoints that were below that threshold. The small shift of the mutant cell population compared to the original plasmid population represents the depletion in a subset of gRNAs compared to the plasmid library. (B) Distribution of log-fold changes in genes that have been previously categorized as being essential (red) or nonessential (black) in the HEK293 and NCI-SNU-1 cell lines. The distribution of fold-changes for nonessential genes centered at ~0, whereas that for essential genes shifted to the left towards negative fold changes. (C) Significantly enriched pathways in genes depleted in NCI-SNU-1 mutant control population 14 days posttransduction. Expected known cell-essential pathways were identified. (D) Robust Rank Algorithm (RRA)-score for genes that were enriched in the sorted cells that had lost the ability to bind the TNFRSF9 probe. Genes were ranked according to the RRA-score. The known interaction partner TNFSF9 and genes related to the TP53 pathway (labeled in red) were identified in the screen. (E) Rank-ordered RRA-scores for genes identified from gRNA enrichment analysis required for RH5 binding to HEK293 cells (left panel). *SLC35B2* and *SLC16A1* were identified within a false-discovery-rate (FDR) threshold of 5%. Two additional genes in the HS biosynthesis pathway (i.e., *EXTL3* and *NDST1*) were identified within FDR of 25%. Schematic depicting the general GAG biosynthesis pathway with the relevant genes mapped to the corresponding steps (panel 2). Genes required for the commitment to chondroitin sulphate biogenesis (i.e., *CSGALNACT1/2*) were not identified in the screen. This figure has been modified from Sharma et al.<sup>19</sup>.

**Table 1. Plasmids used in this approach.**

**Table 2: Buffers required for this study.**

**Table 3. Amounts and volumes of reagents for lentivirus packaging mix.**

**Table 4: Primer sequences for amplifying gRNA and NGS.**

**Table 5: PCR for the amplification of gRNAs from high complexity samples.**

**Table 6: PCR conditions for the first PCR.**

**Table 7: PCR for the index tagging of sgRNAs from genetic screens.**

**Table 8: PCR conditions for second PCR.**

**Supplementary Figure S1:** A guide to drawing gates for sorting the nonbinding population. **(A)** An ideal protein candidate for screening should have a clear shift of binding population compared to the control population and the binding should be retained on cells lacking machinery for HS biosynthesis. A heparin blocking experiment can be used in place of testing on *SLC35B2* targeted cell lines. **(B)** Cells lacking the surface staining from the protein ectodomain but expressing BFP fluorescence from lentiviral transduction were collected. The cells displayed are from a screen for the identification of receptor for GABBR2<sup>22</sup>. This figure has been modified from Sharma et al.<sup>19</sup>.

**Supplementary Figure S2: Cell surface glycoprotein transcriptomics based PCA plot using RNA-seq data from over 1,000 cancer cell lines.** Cell lines from Cell Model Passport<sup>27</sup> were clustered using K-means clustering according to the FPKM values of ~1,500 cell surface glycoproteins. Representative cell lines from each cluster are labeled. Cluster 5 was entirely composed of cell lines of hematopoietic origin (also see **Supplementary Table 2**).

**Supplementary Figure S3: Essentiality scores for KEGG-annotation protein export and N-linked glycosylation genes from project scores.** Adjusted Bayes-essentiality scores for ~330 cell lines (columns, not labeled) are plotted for genes of protein export and N-linked glycosylation pathway (X-axis). Scores higher than 0 represent significant depletion in the mutant population compared to the original plasmid library. The genes can be divided into three distinct clusters that represent different levels of essentiality in the cell lines. This clustering can be used to decide the day of sorting. If the screen is performed at a late time point (day 16), it is possible that genes that are known to be essential for cells (clusters 1 and 3) will not be identified.

**Supplementary Table 1: Raw count files for and MAGeCK software generated gene\_summary files related to the representative genetic screens.**

**Supplementary Table 2: Clustering of cell lines according to the expression of cell surface receptors.**

## **DISCUSSION:**

A CRISPR-based screening strategy to identify genes encoding cellular components involved in cellular recognition is described. A similar approach using CRISPR activation also provides a genetic alternative to identify directly interacting receptors of recombinant proteins without the need to generate large protein libraries<sup>26</sup>. However, one major advantage of this approach is that it is applicable to interactions mediated by surface molecules natively displayed on the cell and does not depend on the overexpression of receptors, which can influence the binding avidity of the receptor. Unlike other methods, therefore, this technique makes no assumptions regarding the biochemical nature or cell biology of the receptors and provides an opportunity to study interactions mediated by proteins that are normally difficult to study using biochemical approaches, such as very large proteins, or those that traverse the membrane multiple times or form complexes with other proteins, and molecules other than proteins such as glycans,

glycolipids, and phospholipids. Given the genome-scale nature of the method, this approach also has the advantage of not only identifying the receptor but also additional cellular components that are required for the binding event, thereby providing insights into the cell biology of the receptor.

One of the major limitations of this method when using it to identify the receptor of an orphan protein is the initial requirement to first identify a cell line that binds to the protein. This is not always easy and identifying a cell line that displays a binding phenotype that is also permissive to genetic screens can be the time-limiting step for deploying this assay. Some cell lines tend to bind to more proteins than others. This is especially relevant for proteins that bind HS, because these proteins tend to bind to any cell line that displays HS side chains, irrespective of the native binding context. Additionally, we have observed that upregulation of syndecans (i.e., proteoglycans that contain HS) in cell lines leads to increased binding of HS-binding proteins<sup>26</sup>. This could be a factor to take into consideration when selecting the cell line for screening. However, also important to note is that the additive binding of HS does not interfere with the binding to a specific receptor. This means that if binding is observed, it is possible that it is mediated solely by HS because the binding mediated by HS in this assay is additive rather than codependent<sup>19</sup>. In such a scenario, the heparin blocking approach described can identify such behaviors without having the need to perform a full genetic screen.

A useful resource for choosing cell lines is Cell Model Passport, which contains genomics, transcriptomics, and culture condition information for ~1,000 cancer cell lines<sup>27</sup>. Depending on the biological context, cells can be chosen based on their expression profiles. To aid the selection of cell lines, we clustered ~1,000 cell lines in Cell Model Passport according to the expression of ~1,500 preannotated human cell surface glycoproteins<sup>28</sup> (**Supplementary Figure 2**; cluster information for each cell line together with growth conditions are provided in **Supplementary Table 2**). When testing the binding of a protein with unknown function, it is useful to select a panel of representative cell lines from each cluster to increase the chance of covering a wide range of receptors. Given a choice, it is recommended to choose cell lines that are easy to culture and easy to transduce. As these cell lines will be used in genome-scale screening, it is preferable that they can be grown easily in large quantities and are permissive to lentiviral transduction, because it is the most commonly available method for delivery of sgRNA for CRISPR-based genetic screening in the later steps.

Generally, the phenotype selections are carried out in a single sort. However, this is determined by the brightness of the stained cell population compared to the control. Iterative rounds of selections could be adopted for scenarios in which the signal-to-noise ratio of the desired phenotype is low, or when the aim of the screen is to identify mutants that have strong phenotypes. When using an iterative selection approach for FACS-based screens, it is important to consider that the sorting process can cause cell death, mainly due to the sheer force of the sorter. Thus, not all collected cells will be represented in the next round of sorting.

Library complexity is a very important factor in performing successful genetic screens, especially for negative selection screens because the extent of depletion in these can only be determined

by comparing results to what was present in the starting library. For negative selection screens, it is common to maintain libraries of 500–1,000 x complexity. Positive selection screens, however, are more robust to library sizes, because in such screens only a small number of mutants are expected to be selected for a particular phenotype. Therefore, in the positive selection screen described here, the library size can be decreased to 50–100x complexity without compromising the quality of the screen. In addition, in these screens it is also possible to use a control library for a given cell line on a given day as a "general control" for all samples sorted on the day for that given cell line. This will reduce the number of control libraries that need to be produced and sequenced.

Another important consideration for using this approach is the limitations of loss-of-function screens in identifying genes that are essential for in vitro cell growth. The timing of the screens is crucial in this regard, as the longer the mutant cells are kept in culture, the higher the likelihood that cells with mutations in essential genes become nonviable and are no longer represented in the mutant library. The recent genetic screens performed as a part of the Project Score initiative in over 300 cell lines show that multiple genes in the KEGG-annotated protein secretion and N-glycosylation pathway are often identified as being essential for a number of cell lines (**Supplementary Figure 3**)<sup>29</sup>. This can be taken into consideration when designing screens if the effect of genes required for proliferation and viability is to be investigated in the context of cellular recognition process. In this case, carrying out screens at an early timepoint (e.g., day 9 posttransduction) would be generally appropriate. However, if the approach is used to identify a few targets with strong size effects rather than general cellular pathways, it might be appropriate to perform screens at a later time point (e.g., day 15–16 posttransduction).

The results from the screening are very robust; in eight recombinant protein binding screens performed in the past, the cell surface receptor was the top hit in every case<sup>19</sup>. When using this approach to identify the interaction partner, one should therefore expect the receptor and the factors contributing to its presentation on the surface to be identified with a high statistical confidence. Once the screen is performed and a hit is validated using a single gRNA knockout, further follow-ups can be performed using existing biochemical methods such as AVEXIS<sup>4</sup> and direct saturable binding of purified proteins using surface plasmon resonance. The approach described here is applicable for all proteins for which it is possible to generate a soluble recombinant binding probe.

In summary, this is a genome-scale CRISPR knockout approach to identify interactions mediated by cell surface proteins. This method is generally applicable to identify cellular pathways required for cell surface recognition in a wide range of different biological contexts, including between an organism's own cells (e.g., neural and immunological recognition), as well as between host cells and pathogen proteins. This method provides a genetic alternative to biochemical approaches designed for receptor identification, and because it does not require any prior assumptions regarding the biochemical nature or cell biology of the receptors it has great potential to make completely unexpected discoveries.

## ACKNOWLEDGMENTS:

This work was supported by the Wellcome Trust grant number 206194 awarded to GJW. We thank the Cytometry Core facility: Bee Ling Ng, Jennifer Graham, Sam Thompson, and Christopher Hall for help with FACS.

#### DISCLOSURES:

The authors have nothing to disclose.

#### REFERENCES:

1. Wright, G. J. Signal initiation in biological systems: the properties and detection of transient extracellular protein interactions. *Molecular bioSystems*. **5** (12), 1405–1412 (2009).
2. van der Merwe, P. A., Barclay, A. N. Transient intercellular adhesion: the importance of weak protein-protein interactions. *Trends in Biochemical Sciences*. **19** (9), 354–358 (1994).
3. Wood, L., Wright, G. J. Approaches to identify extracellular receptor-ligand interactions. *Current Opinion in Structural Biology*. **56**, 28–36 (2019).
4. Bushell, K. M., Söllner, C., Schuster-Boeckler, B., Bateman, A., Wright, G. J. Large-scale screening for novel low-affinity extracellular protein interactions. *Genome Research*. **18** (4), 622–630 (2008).
5. Visser, J. J. et al. An extracellular biochemical screen reveals that FLRTs and Unc5s mediate neuronal subtype recognition in the retina. *eLife*. **4**, e08149 (2015).
6. Özkan, E. et al. An extracellular interactome of immunoglobulin and LRR proteins reveals receptor-ligand networks. *Cell*. **154** (1), 228–239 (2013).
7. Martinez-Martin, N. et al. An Unbiased Screen for Human Cytomegalovirus Identifies Neuropilin-2 as a Central Viral Receptor. *Cell*. **174** (5), 1158–1171.e19 (2018).
8. Bianchi, E., Doe, B., Goulding, D., Wright, G. J. Juno is the egg Izumo receptor and is essential for mammalian fertilization. *Nature*. **508** (7497), 483–487 (2014).
9. Mullican, S. E. et al. GFRAL is the receptor for GDF15 and the ligand promotes weight loss in mice and nonhuman primates. *Nature Medicine*. **23** (10), 1150–1157 (2017).
10. Turner, L. et al. Severe malaria is associated with parasite binding to endothelial protein C receptor. *Nature*. **498** (7455), 502–505 (2013).
11. Frei, A. P. et al. Direct identification of ligand-receptor interactions on living cells and tissues. *Nature Biotechnology*. **30** (10), 997–1001 (2012).
12. Sobotzki, N. et al. HATRIC-based identification of receptors for orphan ligands. *Nature Communications*. **9** (1), 1519 (2018).
13. Sharma, S., Petsalaki, E. Application of CRISPR-Cas9 Based Genome-Wide Screening Approaches to Study Cellular Signalling Mechanisms. *International Journal of Molecular Sciences*. **19** (4) (2018).
14. Gebre, M., Nomburg, J. L., Gewurz, B. E. CRISPR-Cas9 Genetic Analysis of Virus-Host Interactions. *Viruses*. **10** (2) (2018).
15. Zotova, A., Zotov, I., Filatov, A., Mazurov, D. Determining antigen specificity of a monoclonal antibody using genome-scale CRISPR-Cas9 knockout library. *Journal of Immunological Methods*. **439**, 8–14 (2016).
16. Puschnik, A. S., Majzoub, K., Ooi, Y. S., Carette, J. E. A CRISPR toolbox to study virus-host interactions. *Nature Reviews. Microbiology*. **15** (6), 351–364 (2017).

17. Kerr, J. S., Wright, G. J. Avidity-based extracellular interaction screening (AVEXIS) for the scalable detection of low-affinity extracellular receptor-ligand interactions. *Journal of Visualized Experiments*. (61), e3881 (2012).
18. Tzelepis, K. et al. A CRISPR Dropout Screen Identifies Genetic Vulnerabilities and Therapeutic Targets in Acute Myeloid Leukemia. *Cell Reports*. **17** (4), 1193–1205 (2016).
19. Sharma, S., Bartholdson, S. J., Couch, A. C. M., Yusa, K., Wright, G. J. Genome-scale identification of cellular pathways required for cell surface recognition. *Genome Research*. **28** (9), 1372–1382 (2018).
20. Wang, B. et al. Integrative analysis of pooled CRISPR genetic screens using MAGeCKFlute. *Nature Protocols*. **14** (3), 756–780 (2019).
21. R Core Team R: A language and environment for statistical computing. at <<http://www.R-project.org/>> (2013).
22. Hart, T. et al. Evaluation and Design of Genome-Wide CRISPR/SpCas9 Knockout Screens. *G3*. **7** (8), 2719–2727 (2017).
23. Kuleshov, M. V. et al. Enrichr: a comprehensive gene set enrichment analysis web server 2016 update. *Nucleic Acids Research*. **44** (W1), W90–W97 (2016).
24. Crosnier, C. et al. Basigin is a receptor essential for erythrocyte invasion by *Plasmodium falciparum*. *Nature*. **480** (7378), 534–537 (2011).
25. Kirk, P. et al. CD147 is tightly associated with lactate transporters MCT1 and MCT4 and facilitates their cell surface expression. *The EMBO Journal*. **19** (15), 3896–3904 (2000).
26. Chong, Z.-S., Ohnishi, S., Yusa, K., Wright, G.J. Pooled extracellular receptor-ligand interaction screening using CRISPR activation. *Genome Biology*. **19** (1), 205 (2018).
27. van der Meer, D. et al. Cell Model Passports—a hub for clinical, genetic and functional datasets of preclinical cancer models. *Nucleic Acids Research*. **47** (D1), D923–D929 (2019).
28. Bausch-Fluck, D. et al. A mass spectrometric-derived cell surface protein atlas. *PLoS One*. **10** (3), e0121314 (2015).
29. Behan, F. M. et al. Prioritization of cancer therapeutic targets using CRISPR-Cas9 screens. *Nature*. **568** (7753), 511–516 (2019).

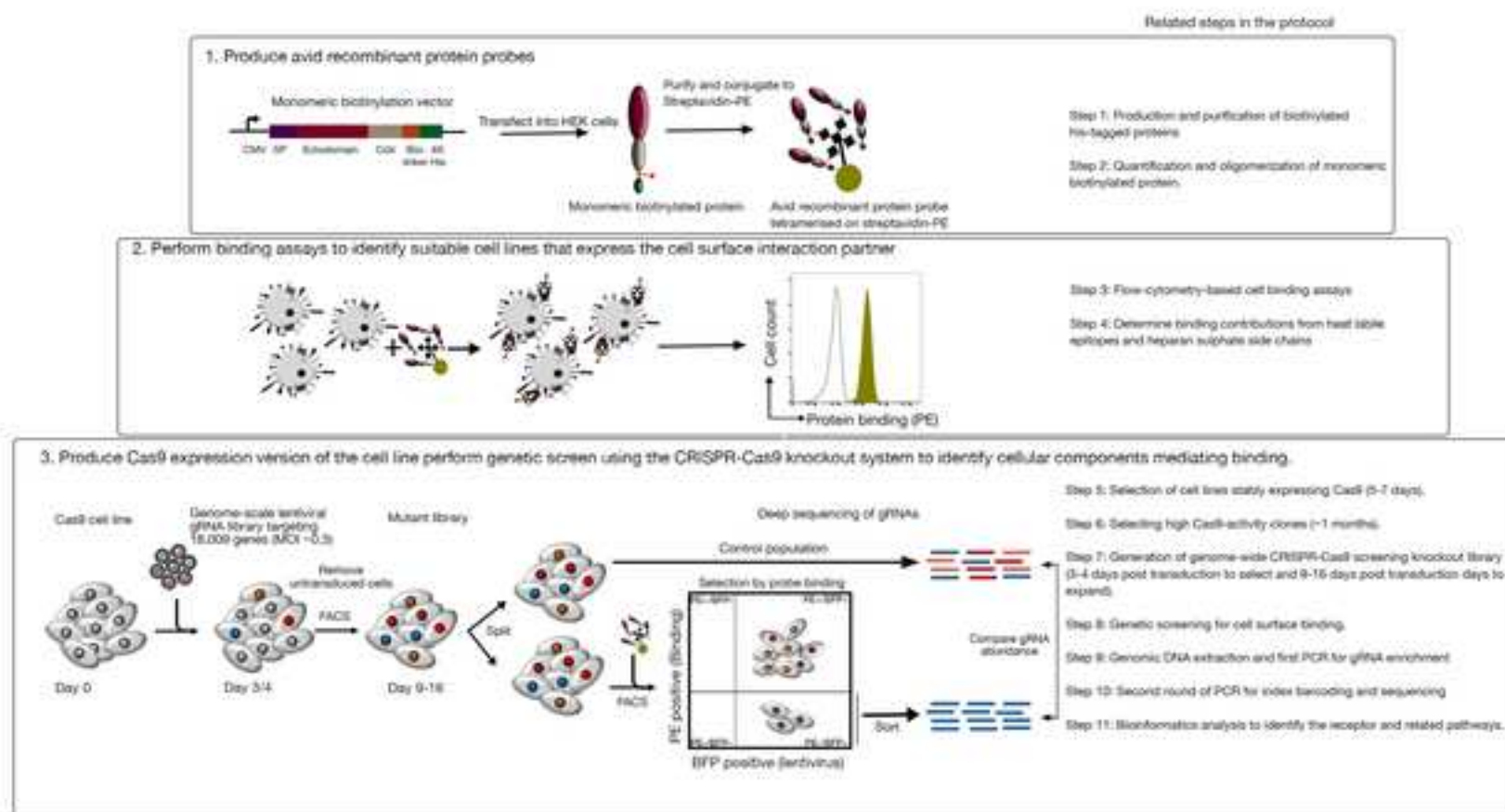
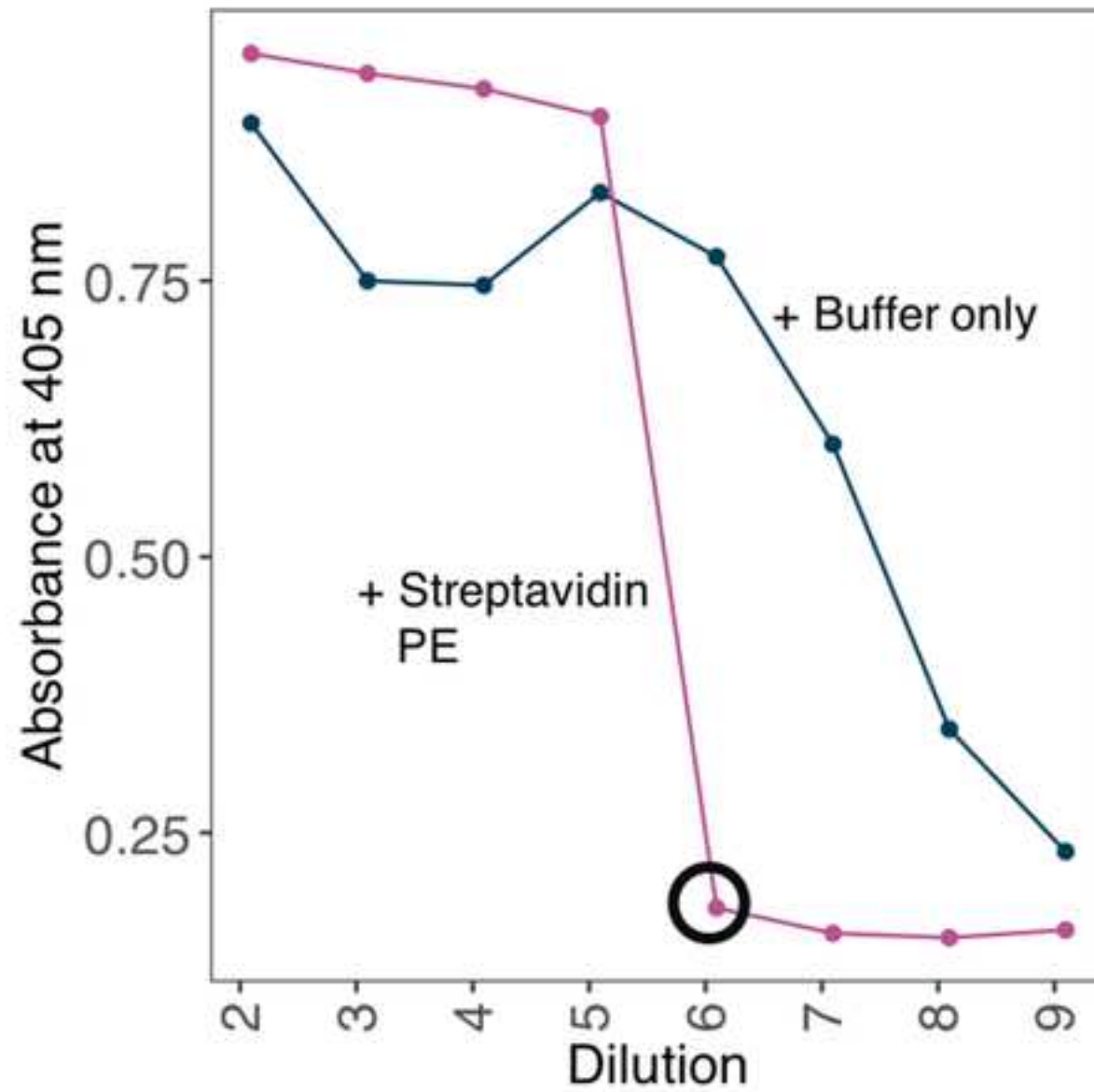
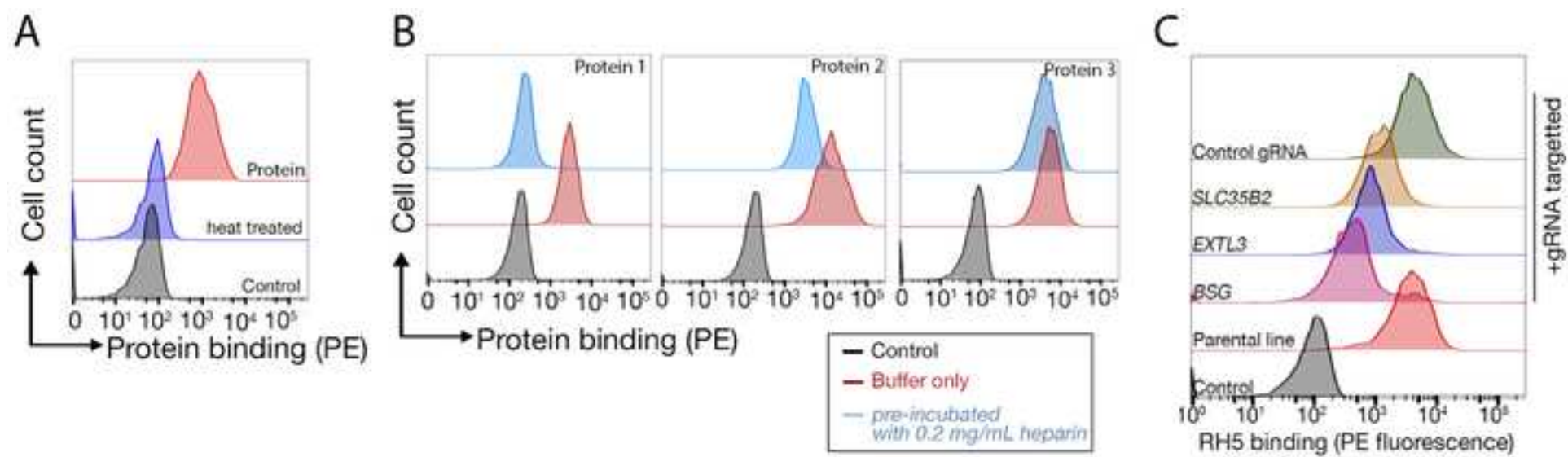


Figure 2





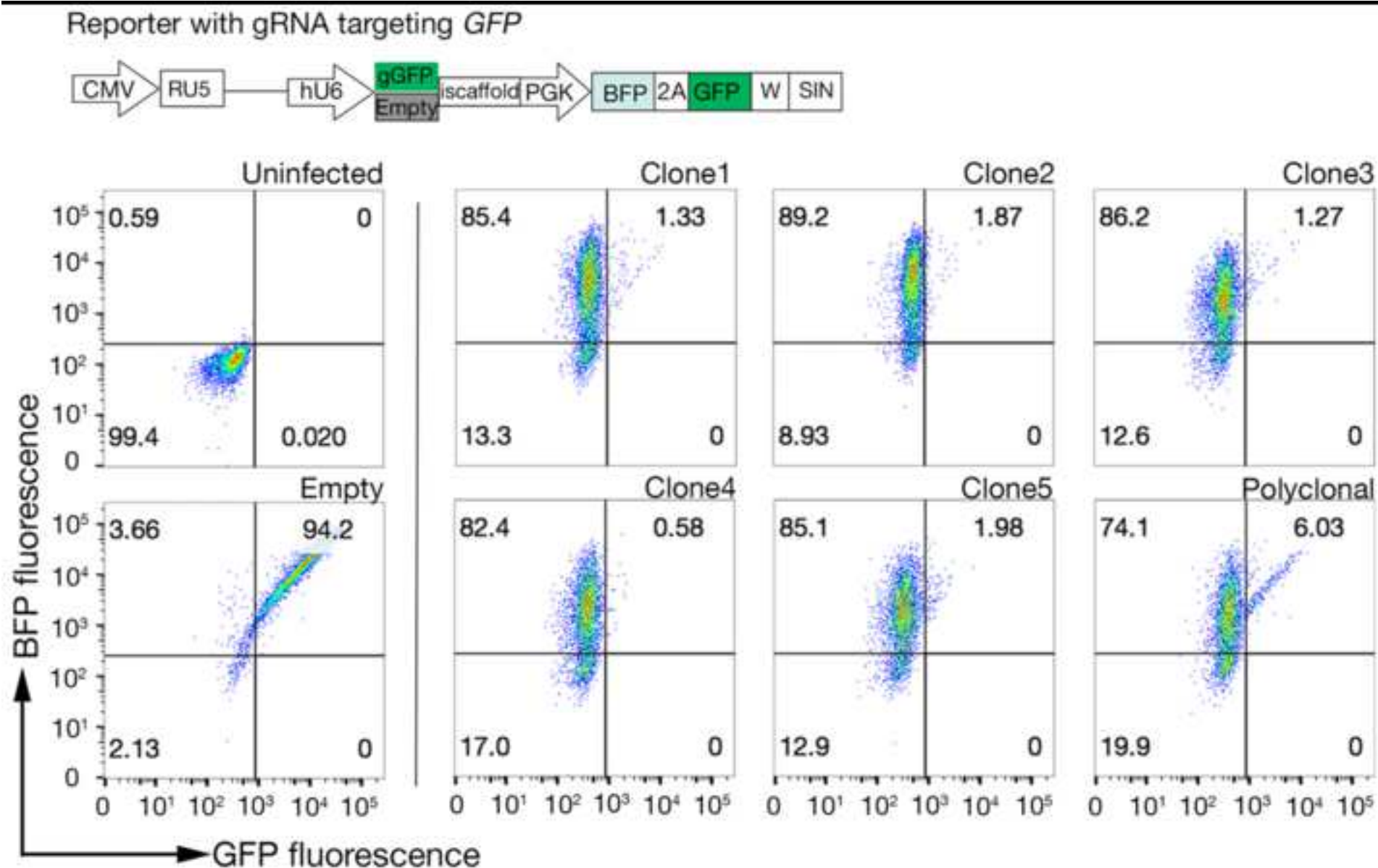
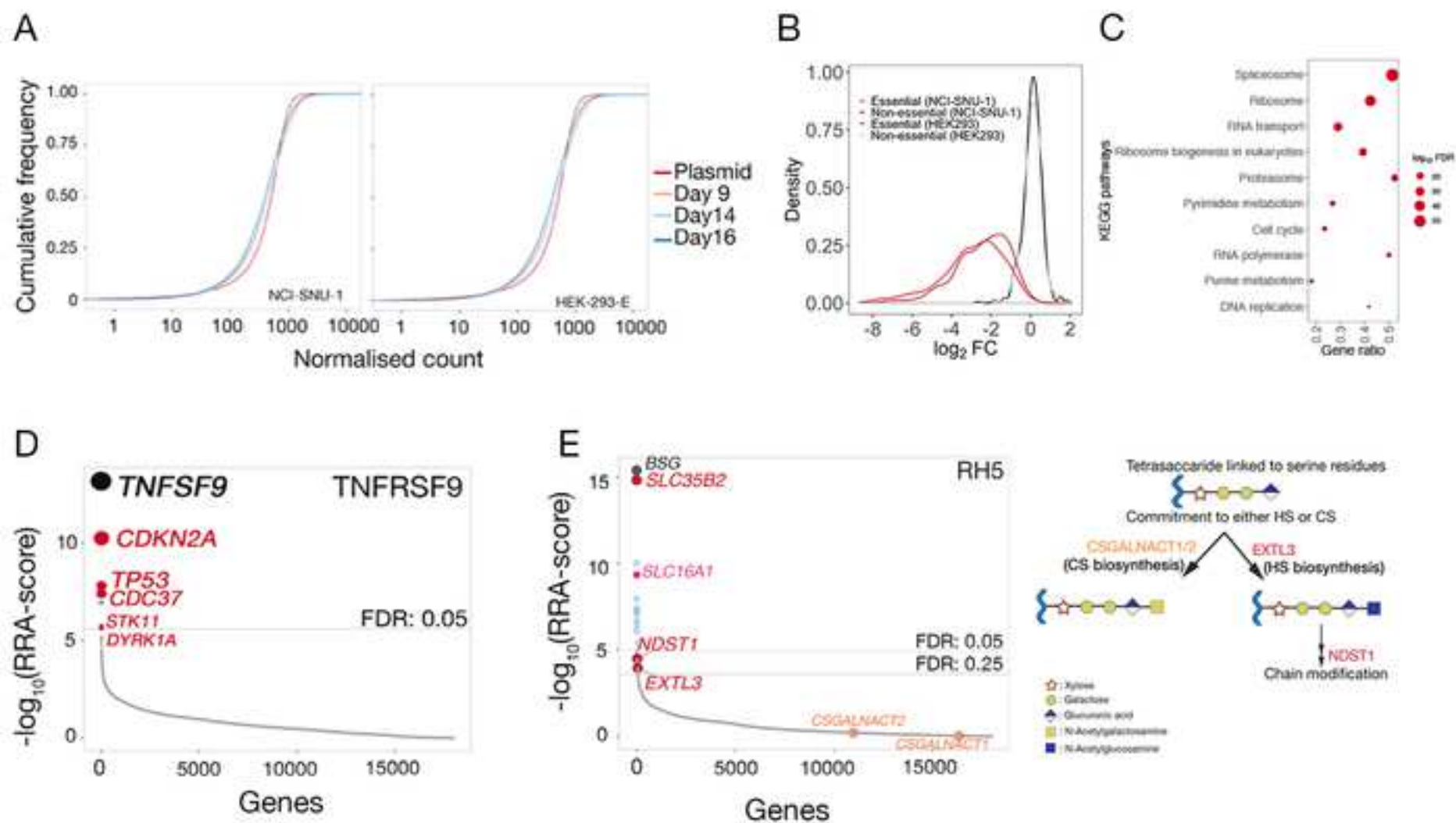


Figure 5

[Click here to access/download;Figure;Figure 5.psd](#)

**Table 1.** Plasmids helpful for this approach

Plasmid name	Plasmid #
Protein expression construct: CD200RCD4d3+4-bio-linker-his	Addgene: 36153
pMD2.G	Addgene: 12259
psPAX2	Addgene: 12260
Cas9-construct: pKLV2-EF1a-Cas9Bsd-W	Addgene: 68343
gRNA expression construct: pKLV2-U6gRNA5(BbsI)-PGKpuro2ABFP-W	Addgene: 67974
Human Improved Genome-wide Knockout CRISPR Library	Addgene: 67989
GFP-BFP construct: pKLV2-U6gRNA5(gGFP)-PGKBFP2AGFP-W	Addgene: 67980
Empty construct: pKLV2-U6gRNA5(empty)-PGKBFP2AGFP-W	Addgene: 67979

## Use

Production of recombinant Protein with CD4d3+4, biotin and 6-his tags.

VSV-G envelope expressing plasmid; production of lentivirus

Lentiviral packaging plasmid, production of lentivirus

Production of constitutively expressing Cas9 line

CRISPR gRNA expression vector with an improved scaffold and puro/BFP markers

A gRNA library against 18,010 human genes, designed for use in lentivirus.

Cas9 activity reporter with BFP and GFP.

Cas9 activity reporter (control) with BFP and GFP.

**Table 2.** Buffer's used

Buffer name
HBS (10X)
PBS (10X)
Sodium Phosphate Buffer (80mM stock)
His-purification binding buffer
His-purification elution buffer
Diethanolamine buffer
D10

## Components

1.5 M NaCl and 200 mM HEPES in MiliQ water, adjust to pH 7.4

80 g NaCl, 2 g KCl, 14.4 g  $\text{Na}_2\text{HPO}_4$  and 2.4 g  $\text{KH}_2\text{PO}_4$  in MiliQ water, adjust to pH 7

7.1 g  $\text{Na}_2\text{HPO}_4 \cdot 2\text{H}_2\text{O}$ , 5.55 g  $\text{NaH}_2\text{PO}_4$ , adjust to pH 7.4

20 mM Sodium Phosphate Buffer, 0.5 M NaCl and 20 mM Imidazole, adjust to pH

20 mM Sodium Phosphate Buffer, 0.5M NaCl and 400 mM Imidazole, adjust to pH

10 % diethanolamine and 0.5 mM  $\text{MgCl}_2$  in MiliQ water, adjust to pH 9.2:

DMEM, 1% penicillin-streptomycin (100 units/mL) and 10 % heat inactivated FBS

.4

7.4

| 7.4

<b>Table 3.</b> Amounts and volumes of reagents with the Addgene’s lentivirus packaging mix		
Components	10-cm dish	6-well plate
293FT cells	70 –80% confluent	70 –80% confluent
Transfection compatible media (Opti-MEM) (Step 5.1.	3 mL	500 µL
Transfection compatible media (Opti-MEM) (Step 5.1.	5 mL	2 mL
Lentiviral transfer vector	3 µg	0.5 µg
psPax2 (see table 1)	7.4 µg	1.2 µg
pMD2.G (see table 1)	1.6 µg	0.25 µg
PLUS reagent	12 µL	2 µL
Lipofectamine LTX	36 µL	6 µL
D10 (Step 7.1.7)	5 mL	1.5 mL
D10 (Step 7.1.8 and 7.1.10)	8 mL	2 mL

Table 4: Primer sequences for amplifying gRNA and N	
Name	
U1	
L1	
PE1.0	
iPCRTag1	
iPCRTag2	
iPCRTag3	
iPCRTag4	
iPCRTag5	
Custom sequencing primer	

GS 35. Order PAGE-purified oligos.

Sequence 5'-3' (restriction sites are indicated in bold, \* represents phosphorothioate)

ACACTCTTTCCCTACACGACGCTCTTCCGATCTCTTGTGGAAAGGACGAAACA

TCGGCATTCTGCTGAACCGCTCTTCCGATCTCTAAAGCGCATGCTCCAGAC

AATGATACGGCGACCACCGAGATCTACACTCTTTCCCTACACGACGCTCTTCCGATC\*T

CAAGCAGAAGACGGCATACGAGATAACGTGATGAGATCGGTCTCGGCATTCCTGCTGAACCGCTCTTCCGA

CAAGCAGAAGACGGCATACGAGATAAACATCGGAGATCGGTCTCGGCATTCCTGCTGAACCGCTCTTCCGA

CAAGCAGAAGACGGCATACGAGATATGCCTAAGAGATCGGTCTCGGCATTCCTGCTGAACCGCTCTTCCGA

CAAGCAGAAGACGGCATACGAGATAGTGGTCAGAGATCGGTCTCGGCATTCCTGCTGAACCGCTCTTCCGA

CAAGCAGAAGACGGCATACGAGATACCACTGTGAGATCGGTCTCGGCATTCCTGCTGAACCGCTCTTCCGA

TCTTCCGATCTCTTGTGGAAAGGACGAAACACCG

## Notes

Sense primer to amplify gRNAs from genomic DNA

Antisense primer to amplify gRNAs from genomic DNA

Common sense primer to amplify first PCR product

Next generation sequencing primers

Next generation sequencing primers with barcodes

Next generation sequencing primers with barcodes

Next generation sequencing primers with barcodes

Next generation sequencing primers with barcodes

Custom primer that only sequences the variable 19bps gRNA sequence

**Table 5:** PCR for the amplification of gRNAs from high complexity samples

Reagent	Volume per reaction	Master mix (x38)
Q5 Hot Start High-Fidelity 2x	25 µL	950 µL
Primer (L1/U1) mix (10 µM each)	1µL	38 µL
Genomic DNA (1 mg/mL)	2 µL	72 µL
H <sub>2</sub> O	22 µL	1100 µL
Total	50 µL	1900 µL

**Table 6:** PCR conditions for the first PCR

Cycle number	Denature	Annealing	Extension
1	98 °C, 30s		
2-24	98 °C, 10s	61 °C, 15s	72 °C, 20 s
25			72 °C, 2 min

**Table 7:** PCR for the index tagging of sgRNAs from genetic screens

Reagent	Volume per reaction
KAPA HiFi HotStart ReadyMix	25 µL
Primer (PE1.0/index primer) mix (5 µM each)	2µL
First PCR product (40 pg/µL)	5 µL
H <sub>2</sub> O	18 µL
Total	50 µL

**Table 8:** PCR conditions for second PCR

Cycle number	Denature	Annealing	Extension
1	98 °C, 30s		
2-15	98 °C, 10s	66 °C, 15s	72 °C, 20 s
16			72 °C, 5 min

Name of Material/Equipment	Company	Catalog Number	Comments/Description
Anti-mouse alkaline phosphatase	Sigma	A4656	
Blasticidin	Chem-Cruz	SC-204655	
Blood & Cell Culture DNA Maxi Kit	Qiagen	13362	
BSA	Sigma	A9647-100G	
Diethanolamine	Sigma	398179	
DMEM	Gibco	31966-021	
Dneasy Blood and Tissue kit	Qiagen	69504	
DynaMag-96 Side Magnet	Invitrogen	12331D	
HEK293T packaging cells	ATCC	CRL-3216	
Heparin	Sigma	H4784-1G	
KAPA HiFi HotStart ReadyMix	Kapa	KK2602	
Lipofectamine LTX with PLUS reagent	Invitrogen	15338100	
MoFlo XDP cell sorter	BD		
Ni <sup>2+</sup> -NTA agarose beads	Jena Bioscience	AC-501-25	
OPTI-MEM	Life Technologies	31985-070	
OX-68 antibody	AbD Serotec	MCA1022R	
p-nitrophenyl phosphate	Sigma	1040-506	
PD-10 desalting columns	GE healthcare	17085101	
Polybrene	Millipore	TR-1003-G	
Polypropylene tubes with 5 mL bed volume	Qiagen	34964	
Proteinase K, recombinant, PCR Grade	Roche	3115879001	
Puromycin	Gibco	A11138-03	
Q5 Hot Start High-Fidelity 2× Master Mix	NEB	M0494L	
QIAquick PCR purification kit	Qiagen	28104	
SCFA filter	Nalgene	190-2545	
Sony Cell sorter	Sony		
SPRI beads (Agencourt AMPure XP beads)	Beckman	A63881	
Streptavidin-coated microtitre plates	Nalgene	734-1284	
Streptavidin-PE	Biolegend	405204	

## Point by point response to reviewers and editorial comments

### Reviewer #1:

#### Manuscript Summary:

The manuscript by Sharma and Wright describes the deployment of genome-scale CRISPR/Cas9 screens in the identification of cell surface receptors using an oligomerized streptavidin-PE recombinant ectodomain probe; hence permitting the identification of receptor-ligand interactions. While comprehensive and clearly within the scope of the journal, certain sections require further clarification to maximize adoption of this approach in a more general setting:

#### Major Concerns:

Line 63: Here the authors state that overexpression of cDNAs to screen for a phenotype of interest can be used to identify new protein-protein interactions. Within this section (or later on in the discussion if more appropriate), authors should also discuss the pitfalls of relying on protein overexpression and how this can lead to an increased risk of false positives, particularly if the native binding avidity of the receptor is also altered. In particular, how a CRISPR/Cas9 based approach overcomes this problem should also be discussed.

We thank the reviewer for this comment. The CRISPR/Cas9 based loss-of-function based screening approach described here identifies the interactions mediated by cell surface molecules natively displayed on cell surfaces and does not, by contrast to cDNA expression cloning methods, depend on the over-expression of receptors. This also makes this genetic screening approach suitable to study interactions mediated by all types of cell surface molecules regardless of their biological properties.

To highlight this, we have now changed the discussion section:

**‘However, one major advantage of this approach is that it is applicable to interactions mediated by surface molecules natively displayed on the cell and does not depend on the over-expression of receptors, which can influence the binding avidity of the receptor. Unlike other methods, therefore, this technique makes no assumptions regarding the biochemical nature or cell biology of the receptors and provides an opportunity to study interactions mediated by proteins that are normally difficult to study using biochemical approaches; for example, very large proteins, or those that traverse the membrane multiple times or form complexes with other proteins, and molecules than proteins such as glycans, glycolipids and phospholipids.’**

Line 269: If an EDTA-based cell detachment method is favored, why is FBS-containing media required for neutralization and washing of the cells? Can the cells not directly be resuspended in wash buffer (PBS/1%BSA)?

We thank the reviewer for pointing this out. We have found that if EDTA-based cell detachment buffer is used, the cells can be directly resuspended in wash buffer. We have now added this point to the protocol.

Line 289: What does "...step 5.12..." refer to?

Thank you for pointing out the typo. This has now been corrected.

Line 300: "...ideally over 80 % of the cells..." How did the authors determine that this value (and not lower) is appropriate for identification of a positive binding cell line? Why would a cell line that demonstrates <80% binding not be considered a positive line? Based on the previous gating steps if less than 1% is set at the positive gate then anything higher should be considered a positive result?

The 80% cut-off is based on our experience of performing genetic screens, but we agree that the way it is written here could be construed as an absolute cut off, which is not the case. Because we are using a single stringent sorting approach for performing genetic screens (1-5% of non-binding population), if the binding population is very low, the signal to noise ratio will be very low and the screen might not be successful.

We have reworded this and added the following sentence as a note:

**"[Note: Cell lines that display a higher binding population are desired for genetic screens, as they have a higher signal to noise ratio. Ideally over 80 % of the cells should fall within this gate.]"**

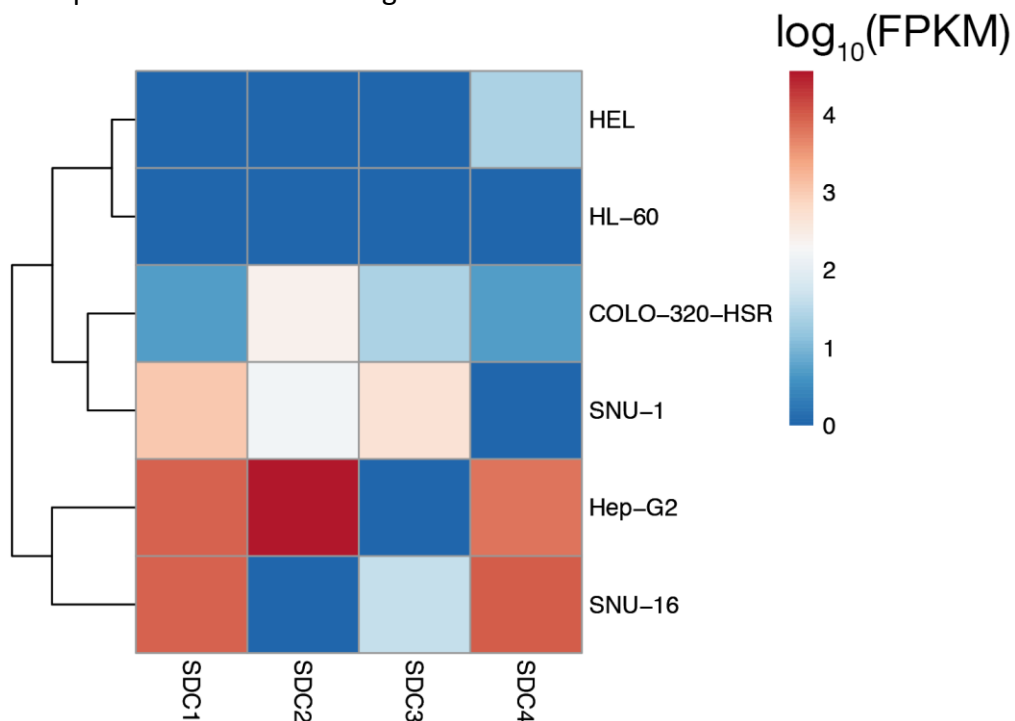
Line 316: It is unclear to me what is exactly meant by "Binding to HS on the cell surface is not necessarily 'non-specific' but rather a property of the protein." Given the authors' experiences with this phenomenon during their screening assays, can they give further comment as to how this HS binding property varies between cell lines that indicate positive binding to the probe? Does the additive binding interfere with the sensitivity of the assay?

Following this point, a large section appears to be devoted in discussing practical ways around the HS binding issues, which suggests that it has a significant impact on the deployment of the assay. It would therefore be helpful if the authors detail all the cell lines they have tested that have this property and provide some statistical analysis demonstrating the magnitude of the additive binding and how the proposed genetic targeting approaches targeting the HS/GAG biosynthesis pathways improves the assay sensitivity. Additionally, the impact of interference of HS/GAG biosynthesis on the biological properties of the cells (and indeed any impact on receptor-ligand interactions or screening) need to be discussed further.

In the cell binding assay designed here, we have generally found that the contribution of HS towards binding is additive rather than co-dependent on other receptors ([Sharma et al. 2018](#)). Also, if a protein has HS-binding property, it should, in principle, bind to *all* cell lines that display HS molecules. Pragmatically, it also means that if the protein has HS-binding properties, it will

likely bind to any HS-expressing cell surface, even in the absence of another receptor. The initial HS-binding test is to aid the understanding of the biochemical properties of the recombinant protein. If a complete loss of binding is observed in this heparin blocking experiment, it is likely that the genes identified in the full genetic screen would be dominated by those involved in HS biosynthesis. We have shown this to be a case in both *P. falciparum* and mammalian proteins (in at least 6 cases) (Sharma 2018).

We have tested a number of cell lines originating from different tissue types and have noticed that some cell lines tend to bind to proteins in a HS-dependent manner more than others, presumably due to higher cell surface levels of HS. We have observed a greater number of proteins binding to cell lines such as HEK293, NCI-SNU1, NCI-SNU-16, COLO-320-HSR, and HepG2. Cell lines originating from the haematopoietic lineage such as HEL, HL-60 generally bind to fewer proteins which are not dependent on HS. One of our observations from CRISPR activation based experiments has been that upregulations of syndecans (proteoglycans that display HS at the cell surface) in cell lines lead to increased binding of proteins that have HS-binding properties (Chong et al. 2018). We have also found that expression of syndecans tend to be higher in the cell lines in which we tend to see more HS-dependent binding (figure 1). We have, however, not performed a comprehensive analysis to study the binding behavior of different cell lines so we would be cautious in confidently providing the identity of the cell lines that express differential binding behavior.



**Figure 1:** Expression levels of syndecan family member proteins in different cell lines. Normalized FPKM values for each cell line was obtained from RNA-seq expression analysis from cell model passport.

However, we have found that the additive binding of HS does not interfere with the binding to a specific receptor: if a protein has a HS-binding property but also binds to a second receptor, the protein will continue to bind to cells in the absence of HS on the surface of the cells. Heparin blocking experiments in such cases will reveal residual binding. We have illustrated this in figure 3 of the manuscript. In genetic screens, we have been able to identify both the receptor and the HS biosynthesis pathway (also see figure 3). We have dedicated a substantial section in our original paper to describe this phenomenon using the example of interaction mediated by the *P. falciparum* invasion ligand RH5. We therefore believe that it would be more suitable to point the readers to that section of the original paper rather than rewrite the whole section again.

To address these points, we have included this in the discussion section :

**One of the major limitations of this method when using it to identify the receptor of an orphan protein is the initial requirement to first identify a cell line that binds to the protein. This is not always easy, and identifying a cell line that displays a binding phenotype which is also permissive to genetic screens, can be the time-limiting step for deploying this assay. We have also noticed that some cell lines tend to bind to more proteins than others. This is especially relevant for proteins that have a HS-binding property, because these proteins tend to bind to *any* cell line that displays HS side chains, irrespective of the native binding context. Additionally, we have observed that upregulations of syndecans (proteoglycans that contain HS) in cell lines leads to increased binding of HS-binding proteins (Chong *et al*, 2018). This could be a factor to take into consideration when selecting the cell line for screening. However, also important to note is that the additive binding of HS does not interfere with the binding to a specific receptor. Because a number of proteins are known to bind to HS and the binding mediated by HS in this assay is additive rather than co-dependent, in a practical sense, it means that if a binding is observed, it is possible that it is mediated solely by HS. In such a scenario, the heparin blocking approach that we have described can identify such behaviors without having the need to perform a full genetic screen.**

Line 498: The basis of 30% transduction efficiency is based on the Poisson probability distribution that ensures each cell receives a single gRNA, ensuring that the resultant phenotype is linked to a specific knockout. This point should be included in the section as to make it clear to a novice reader.

We thank the reviewer for this comment. We have now added this point in the figure legend of figure 1.

Given that the control population has not undergone the same processing through FACS the sorted BFP+ population, would the assay not benefit from additionally analysing the BFP+/PE+ population that currently does not get sorted? This may give some insight as to any FACS-related or other technical related loss of gRNA when compared to the control population as  $gRNA(\text{control}) = gRNA(\text{BFP+}/\text{PE+}) + gRNA(\text{BFP+}/\text{PE-})$ . This may additionally inform on any biological response elicited by probe binding that could alter gRNA distributions. Furthermore, some exemplary FACS plots showing a real distribution of the fluorescent populations would be

informative, are the populations clearly defined so the gating is straightforward or is the distribution broader?

During the development of this method, we tried a number of different approaches including the one where we compared gRNA(BFP+/PE+) with gRNA(BFP+/PE-) as suggested by the reviewer. We found it technically difficult to perform a two-way sort because 94-99% of the cells were being sorted into the (BFP+/PE+) tube and the smaller number of (BFP+/PE-) cells in the other. Technically, this resulted in cell “spraying” contaminating the smaller fraction. We observed that the “single sort” approach could overcome this technical limitation. For one-way sorting to work efficiently, we tend to keep the binding population be as high as 80% such that the signal to noise ratio is high. We have now included an example of what we consider to be an ideal binding profile. Gating for these samples is very straightforward as seen from the figure 2. We have added this figure in supplementary figure 1 and referred in text in section 8.10 of the revised manuscript:

**“An example of a sorting gate is provided in supplementary figure 1.”**

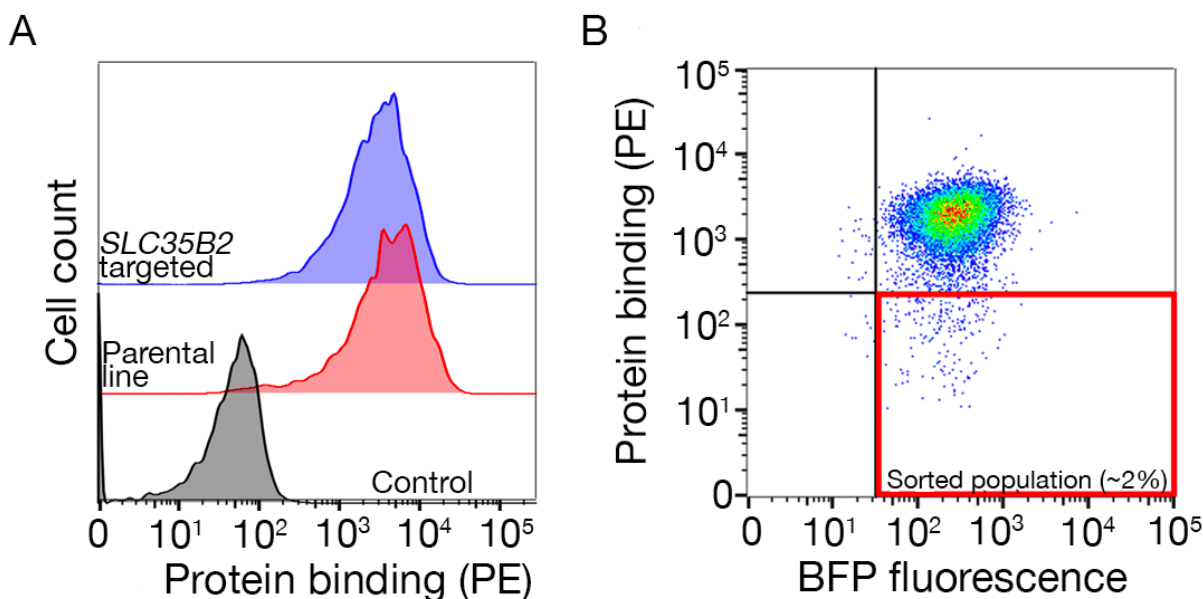


Figure 2: A guide to drawing gates for sorting the non-binding population. **A.** An ideal protein candidate for screening should have a clear shift of binding population compared to the control population and the binding should be retained on cells lacking machinery for HS biosynthesis. A heparin blocking experiment can be used in place of testing on *SLC35B2* targeted cell lines. **B.** Cells lacking the surface staining from the protein ectodomain but expressing BFP fluorescence (from lentiviral transduction) were collected. The cells displayed are from a screen for the identification of receptor for GABBR2 (figure 2A modified from Sharma et al. 2018).

The readout of the assay depends on the loss of the streptavidin-PE probe from the cell, however loss of binding may also be induced gRNA targeting general cellular processes such as

receptor trafficking to the plasma membrane or receptor protein folding within the ER, which may not be necessarily lethal and thus not classed as an essential gene. Does the scoring and ranking take this into consideration?

The reviewer is correct in recognition that the loss of binding can be due to indirect non-lethal processes and this is something we initially considered. However, as we started performing these types of genetic screens, it became less of a concern for us because it was evident that the gRNAs targeting the receptor usually dominate the screen. For the ~10 antibody screens and 9 recombinant protein screens that we have reported to date, we have always identified the receptor within FDR of 0.05 and usually ranked 1. Therefore, we think that it is not crucial to have an additional scoring system for the identification of the receptor itself. We have now added this sentence in the discussion section to make this clear.

**“When using this approach to identify the interaction partner, one should therefore expect the receptor and the factors contributing to its presentation on the surface to be identified with a high statistical confidence.”**

Minor Concerns:

Although obvious, "D10 media" needs to be formally defined.

This is defined in table 2.

**Reviewer #2:**

Manuscript Summary:

Manuscript describes application of GeCKO library screening using fluorescent avid protein probes and negative FACS-sorting procedure. The manuscript is well written and protocol is clear and can be recommended for acceptance without a major revision.

We thank the reviewer for the positive feedback.

Minor Concerns:

1. Protocol #1. In my opinion, the brief description of principles how proteins expressed from vectors are biotinylated, and major advantages or disadvantages, using different multimerization motifs (leucin zippers, COMP and so on) would be very informative. As an option, schematic illustration of this can increase the perception of multimer formation and process understanding.

We thank the reviewer and we agree that that would be very useful for the readers. We have previously reviewed the concepts of multimerization and different multimerization techniques so we believe it would be more useful to point the readers to the review for a more comprehensive understanding (Wright 2009). We have now done this in the introduction section.

2. Previous works on affinity reagents (mAbs and others) screenings using GeCKO library should be cited (J Immunol Methods. 2016 Dec;439:8-14, Cell Rep. 2018 Apr 10;23(2):596-607, others)

Citations for studies that use high-affinity ligands and reviews on pooled genetic screens for the purposes of studying a range of biological processes such as cell signaling and virus-host interactions have now been added in the introduction section.

**Reviewer #3:**

Manuscript Summary:

In the present manuscript, Sharma et al. describe a genome-scale CRISPR/Cas9 knockout genetic screening approach to identify receptors and cellular pathways required for specific cell surface recognition events. Overall, the manuscript is well written and provides detailed information about the methodology and associated protocols.

We thank the reviewer for the positive feedback.

Major Concerns:

Thorough discussion of the methodology's limitations is missing. The authors may consider highlighting more the biological and technical drawbacks in order to provide the reader with realistic expectations on the success of the screening.

We thank the reviewer for this comment. A similar point was raised by reviewer #1 so we will reiterate the point we made before:

The major limitation of this approach is the initial screening required to identify a screen-permissive cell line that interacts with the protein. This is not always straightforward. This is additionally complicated for proteins that have HS-binding property. This is because, such a protein, in principle, binds to *all* cell lines that display HS molecules. Pragmatically, it also means that if the protein has HS-binding properties, it will likely bind to any HS-expressing cell surface, even in the absence of another receptor. We have included an initial HS-binding test to aid the understanding of the biochemical properties of the recombinant protein. If a complete loss of binding is observed in this heparin blocking experiment, it is likely that the genes identified in the full genetic screen would be dominated by those involved in HS biosynthesis. As per the technical aspect of the screens, we have found the CRISPR/Cas9 screening system itself to be very robust. We have always identified the screen among the highest ranking genes and usually on position 1.

Addressing both reviewers, we have now added the following paragraph in the discussion section to address this comment.

**One of the major limitations of this method when using it to identify the receptor of an orphan protein is the initial requirement to first identify a cell line that binds to the protein. This is not always easy, and identifying a cell line that displays a binding phenotype which is also permissive to genetic screens, can be the time-limiting step for deploying this assay. We have also noticed that some cell lines tend to bind to more proteins than others. This is especially relevant for proteins that have a HS-binding property, because these proteins tend**

to bind to *any* cell line that displays HS side chains, irrespective of the native binding context. Additionally, we have observed that upregulations of Syndecans (proteoglycans that contain HS) in cell lines leads to increased binding of HS-binding proteins (Chong *et al*, 2018). This could be a factor to take into consideration when selecting the cell line for screening. However, also important to note is that the additive binding of HS does not interfere with the binding to a specific receptor. Because a number of proteins are known to bind to HS and the binding mediated by HS in this assay is additive rather than co-dependent, in a practical sense, it means that if a binding is observed, it is possible that it is mediated solely by HS. In such a scenario, the heparin blocking approach that we have described can identify such behaviors without having the need to perform a full genetic screen.

Minor Concerns:

Some minor questions/comments regarding the experimental procedure are listed below:  
Overall: the authors could consider to include timelines in each section to facilitate planning and provide a rough idea regarding the duration of the experiment.

We have now added a timeline in the schematic in figure 1. We have highlights the sections in the protocol that require considerable time (more than 1 week) and planning.

Line223: transfer the total volume of the sample?

This refers to transferring the total of 200  $\mu$ L from the dilution plate to the streptavidin plate. We have changed this to the following:

**“Transfer total volume (200  $\mu$ L) from dilution plates to individual wells of the streptavidin-coated plates and incubate for 1 hour at room temperature (RT).**

Line226: what is the volume of the washing buffer?

It should be 200  $\mu$ L, we have now added this.

Line289. step 5.12 does not exist; probable typo

Yes, we have corrected it.

Line295:could be useful to elaborate more on what a control probe is.

Control probe is protein that contains the same tags as a recombinant protein of interest. We include a control protein to ensure that the observed binding is not mediated by the tags. We have defined this in section 3.2. In the edited manuscript this is in section 2.2.

Section 5: could be useful to provide a control graph.

The control graphs are demonstrated in figure 3. We have now also added a representative graph that showcases an ideal sorting gate.

**“An example of a sorting gate is provided in supplementary figure 1.”**

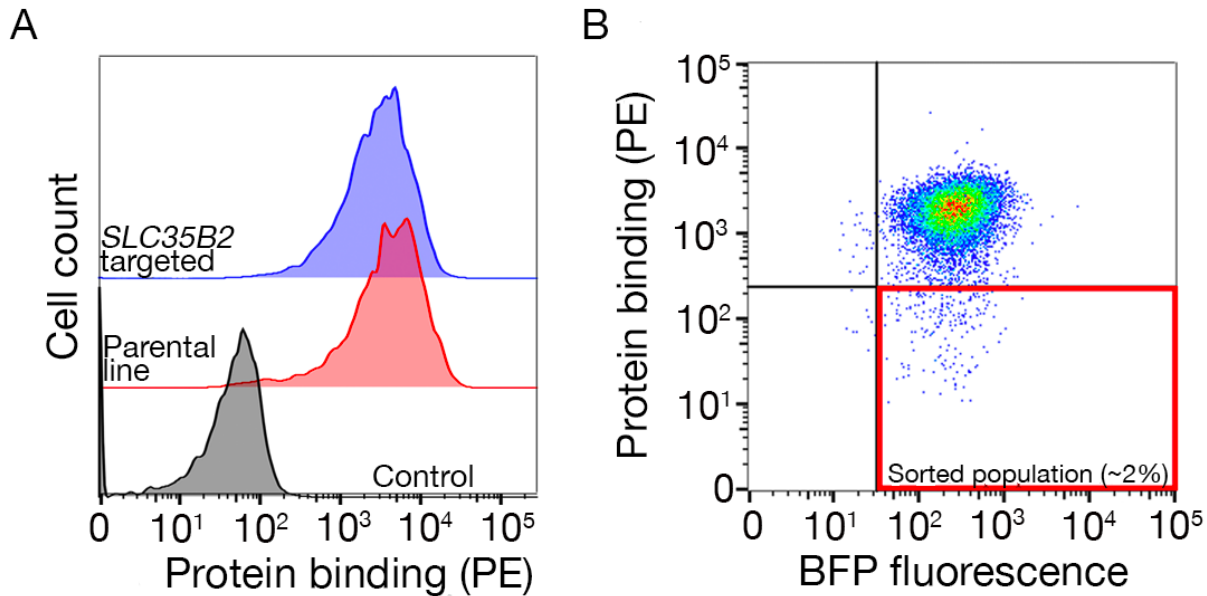


Figure 2: A guide to drawing gates for sorting the non-binding population. **A.** An ideal protein candidate for screening should have a clear shift of binding population compared to the control population and the binding should be retained on cells lacking machinery for HS biosynthesis. A heparin blocking experiment can be used in place of testing on *SLC35B2* targeted cell lines. **B.** Cells lacking the surface staining from the protein ectodomain but expressing BFP fluorescence (from lentiviral transduction) were collected. The cells displayed are from a screen for the identification of receptor for GABBR2 (figure 2A modified from Sharma et al. 2018).

Line527: please indicate the unit (9-16 days?)

We have changed it to 9-16 days.

#### Reviewer #4:

In this protocol, authors have described the steps to employ the CRISPR screen to identify genes that directly and indirectly interfere with the binding of the proteins (of choice) on the cell membranes' array of receptors. The protein of choice is oligomerized on a streptavidin-PE. The oligomerized protein is then subjected to the cell culture. The oligomerized protein molecules will bind the receptors on the cell membrane. The cell culture (expressing Cas9) is pre-subjected to a library of gRNAs, in amounts such that each cell will undergo genetic disruption in roughly one gene only. The cells of the culture, are then FACS sorted according to the fluorescence of the PE (which would be bound to the cell membrane via the protein molecules oligomerized over streptavidin-EP). Further sequencing of such cells and their comparisons with a control population will reveal the genes that play important roles in mediating the binding of the protein of choice onto the cell membrane's receptors. The CRISPR screens are now being widely used to uncover the genetic networks of many biological interactions. This protocol is a timely, elegant and well-detailed adaptation of the CRISPR screen to uncover the genetic networks of protein-membrane receptor interactions.

The protocol is extremely long and well-detailed. I applaud the authors for including these many details. The manuscript should be published. Although I think the current version protocol is not the easiest of the read and it may be taxing for new readers. I would suggest that the protocol be made a little tighter. The protocol must not be stripped of details, as I strongly feel that they are keys to a good protocol. But following (or related) changes can be made to make the protocol a little tighter, without compromising on necessary details:

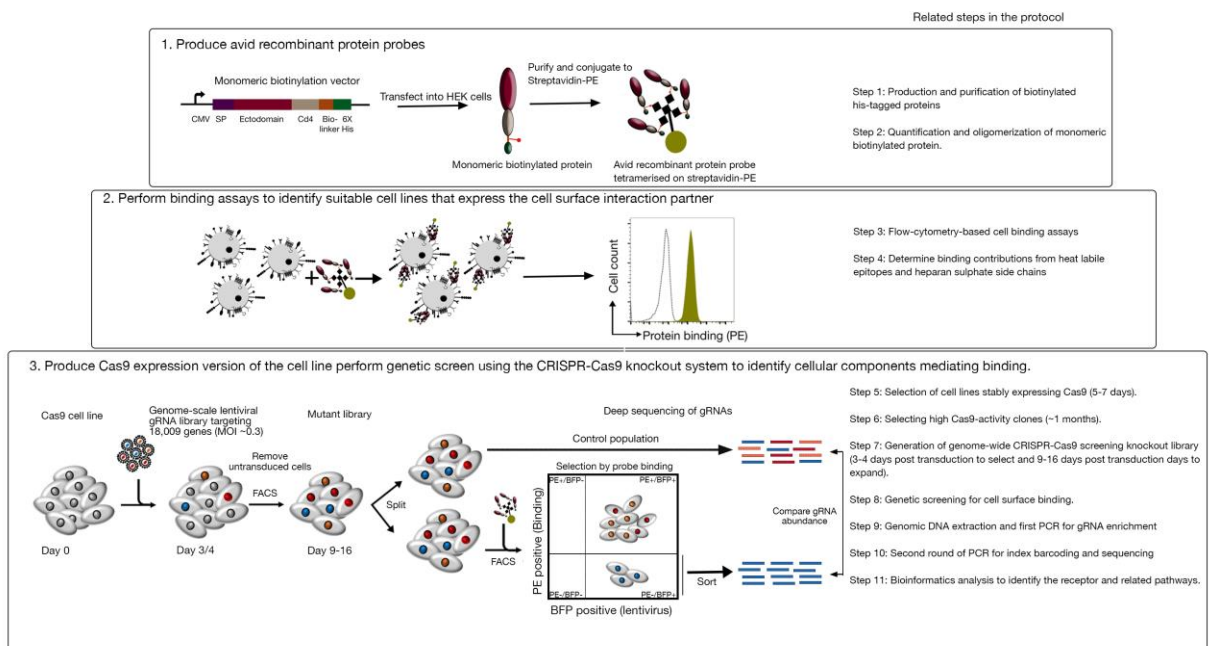
1. The protocol has too many abbreviations. I would suggest removing abbreviations for terms that are not even germane to the manuscript. A simple example: Mass spectrometry (MS). Also remove abbreviations for some other terms and feel free to use the full term instead of abbreviations in the rest of the manuscript.

We thank the reviewer for this comment. We have removed a number of abbreviations as suggested by the reviewer. Examples include: Mass spectrometry (MS), monoclonal antibodies (mAbs), over-night (O/N).

2. The protocol needs a figure flowchart in figure 1, describing all the keys steps.

Thank you for this comment. We have taken this into consideration and changed the figure to make it clear how the individual steps in the protocol relate to the three key parts of the protocol.

Modified figure 1:



3. The abbreviation PE has never been defined in the manuscript. Please define it as Phycoerythrin. Authors also need to specify that the PE is fluorescent and its emission wavelength etc.

Thank you - we have now added this in the introduction section as follows.

**The protocol here will describe the steps for generating fluorescent avid probes from monomeric biotinylated proteins by conjugating them to streptavidin conjugated to a fluorochrome (e.g Phycoerythrin (PE)), which can be used directly in cell-based binding assays and has the advantage of not requiring a secondary antibody for detection.**

The emission wavelength and the detection lasers are mentioned in the protocol section 3.8.

4. All the sections are not necessarily in order of the workflow. E.g., Ways of determining specificity is not necessarily a direct part of the workflow. Is there a way to present these sections separately? I understand that JOVE has a particular format. But if possible, authors may consider highlighting the necessary sections of the workflow, a little separately than the other indirect sections.

We thank the reviewer for this comment. It is true that specificity is not directly related but we have found it to be a very important check that must be performed before proceeding with the screen. That said, we note that we have spent a very long section on this point, which can be distracting. We have considerably trimmed this section to make clearer. Section 4 of the protocol is the new revised section.

5. CRISPR screens are becoming increasingly common, but the protocol assumes that the reader would know about the idea and concept of the CRISPR screen already. I would add a few more details about the idea behind the CRISPR screen.

We agree that it would be useful to have background on CRISPR screens but because of space limitations we are unable to do this in the main text. We have added more details on the concept of CRISPR screen in figure 1.

6. In the introduction, the authors mention that the assay consists of three major steps. But then those three major steps are broken into many sections. A better synergy is needed between the introduction and the sections of the protocol. Perhaps the author can mention that the first step includes section 1 to m or something similar.

We have edited the schematic figure to indicate which section refer to which three steps of the protocol. Hopefully this will make the protocol easier to follow.

7. In data analysis section, It would be better if terms like RRA score & Cumulative distribution function plots are explained in one or two sentences. I understand that readers must look to other manuscripts for details about the analysis and interpretation of the sequencing data. But at the same time, it may be better to ensure that any specialized/uncommon technical terms like RRA, etc are given some context/explanation in the protocol.

We have added more details on specialized terms in the results section. Due to space limitations, we were not able to provide details in the protocol section, but we have included the concepts in the discussion and figure legends. Here are some examples:

**The RRA-score in MAGeCK provides a measure of which gRNAs are ranked consistently higher than expected.**—discussion section.

**For any given number, cumulative density function reports percent of datapoints that are below that threshold.**— Figure 5A legend.

**Editorial comments:**

Changes to be made by the Author(s):

1. Please take this opportunity to thoroughly proofread the manuscript to ensure that there are no spelling or grammar issues. The JoVE editor will not copy-edit your manuscript and any errors in the submitted revision may be present in the published version. Please use American English throughout.

Thanks you, this has been done.

2. Please provide at least 6 keywords or phrases.

Keywords are :CRISPR/Cas9, recombinant proteins, cell adhesion, cell surface receptor, genome-scale screening, extracellular protein interactions.

3. Please ensure that the long Abstract is within 150-300-word limit and clearly states the goal of the protocol.

The current abstract length is 150 words.

4. Please remove all commercial language from your manuscript and use generic terms instead. All commercial products should be sufficiently referenced in the Table of Materials and Reagents.

For example: Addgene, Opti-MEM, Lipofectamine, polybrene, Falcon tube, AMPure XP paramagnetic beads, Illumina NGS, iPCR-11mer tags, etc.

This has now been done.

5. Please ensure that all text in the protocol section is written in the imperative tense as if telling someone how to do the technique (e.g., “Do this,” “Ensure that,” etc.). The actions should be described in the imperative tense in complete sentences wherever possible. Avoid usage of phrases such as “could be,” “should be,” and “would be” throughout the Protocol. Any text that cannot be written in the imperative tense may be added as a “Note.”

The manuscript has been edited to ensure this.

6. The Protocol should contain only action items that direct the reader to do something. Please move the discussion about the protocol to the Discussion.

This has now been done.

7. The Protocol should be made up almost entirely of discrete steps without large paragraphs of text between sections.

Given the complexity of the protocol, we have included a section head for some of the steps so that anyone reading the protocol is able to understand why a certain step is being performed. Large text from protocol itself has now been moved as a NOTE.

8. Please ensure that individual steps of the protocol should only contain 2-3 actions per step. This has now been done.

9. Please revise the protocol text to avoid the use of any personal pronouns in the protocol (e.g., "we", "you", "our" etc.).

This has now been except in [Tip] sections.

10. Please add more details to your protocol steps. Please ensure you answer the “how” question, i.e., how is the step performed?

We have tried to edit the manuscript to include as much information as possible for how the step should be performed.

11. 1.1 and 1.2: Please explain how this is done, else please consider moving this section to the introduction as well.

We have moved this to the introduction section.

12. 4: This section doesn't describe the action in the protocol, this can be moved to the discussion instead.

We have moved this section to the discussion section.

13. 5.1: Please include the cell line used in your example experiment.

HEK293 and NCI-SNU-1 cell was used in the example experiments. However, it is not necessary to use these cell lines. The identity of these cell lines are provided in the representative result section. We would prefer not to provide the identity in the protocol section as it might imply that these cell lines *must* be used.

14. Lines 197-204, 683-685: Only numbered steps describing the action can be filmed. If this need filming, please convert to numbered action step in imperative tense. Else please move some of the details to discussion. Sentences which directly are relevant with the protocol can stay as a NOTE. In the JoVE Protocol format, “Notes” should be concise and used sparingly. They should only be used to provide extraneous details, optional steps, or recommendations that are not critical to a step. Any text that provides details about how to perform a particular step should either be included in the step itself or added as a sub-step.

Lines 197-204 have been moved. Lines 683-683 includes steps from another published protocol. This has now been referred and the action steps have been changed to imperative tense.

15. 6, 7.1, 9, 10, 12.1: Please consider making action steps describing the actual action being performed in your experiment and move the other details to the discussion section.

These have now been addressed as suggested.

16. Please add more specific details for the software steps (e.g. button clicks for software actions, numerical values for settings, etc.).

For MAGECK, a separate protocol for installation has already been published. We have referred to that publication as it is a method of its own. For other analysis steps, we have now included the column header from the output of MAGECK software that are relevant for this study.

17. There is a 10-page limit for the Protocol (including headings and spacings), but there is a 2.75-page limit for filmable content (including headings and spacings). Please highlight 2.75 pages or less of the Protocol (including headings and spacing) that identifies the essential steps of the protocol for the video, i.e., the steps that should be visualized to tell the most cohesive story of the Protocol. Presently the protocol length is more than 10 pages.

The protocol is within 10 pages now.

18. Please obtain explicit copyright permission to reuse any figures from a previous publication. Explicit permission can be expressed in the form of a letter from the editor or a link to the editorial policy that allows re-prints. Please upload this information as a .doc or .docx file to your Editorial Manager account. The Figure must be cited appropriately in the Figure Legend, i.e. "This figure has been modified from [citation]."

All the modified figures have been published under creative common's license. We have included the original paper where the license is printed.

19. As we are a methods journal, please ensure that the Discussion explicitly covers the following in detail in 3-6 paragraphs with citations:

- a) Critical steps within the protocol
- b) Any modifications and troubleshooting of the technique
- c) Any limitations of the technique
- d) The significance with respect to existing methods
- e) Any future applications of the technique

The discussion section has been edited to address these points.

20. JoVE manuscripts do not have a separate conclusion section. Please combine the conclusions with the discussion.

This has now been done.

21. Please do not abbreviate the journal titles in the references section.

This has now been done.

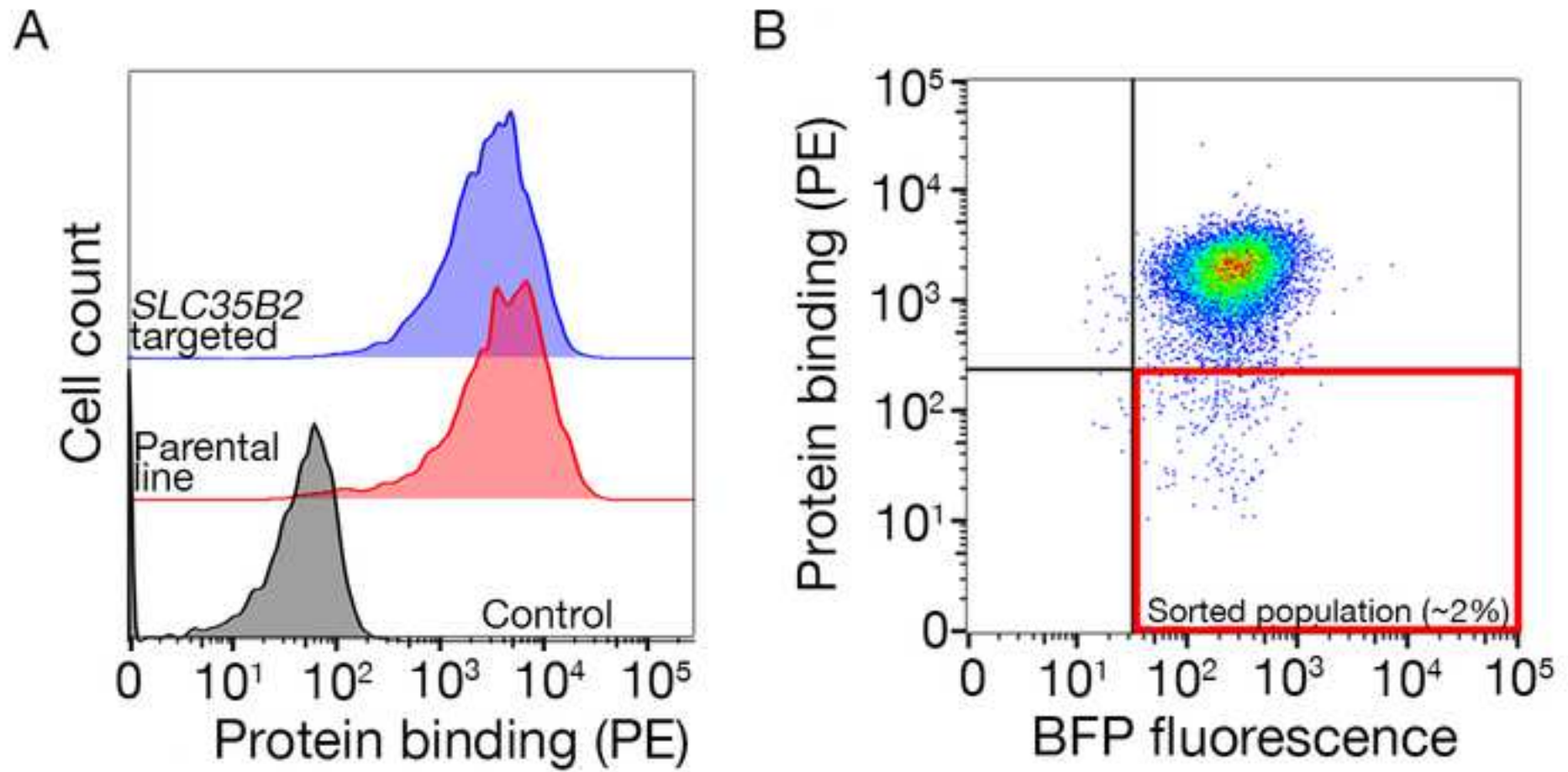
22. Please remove all the commercial terms referred to in the tables/figures. Please use the generic term instead. All commercial terms can be referred to in the table of materials instead.

This has now been done.

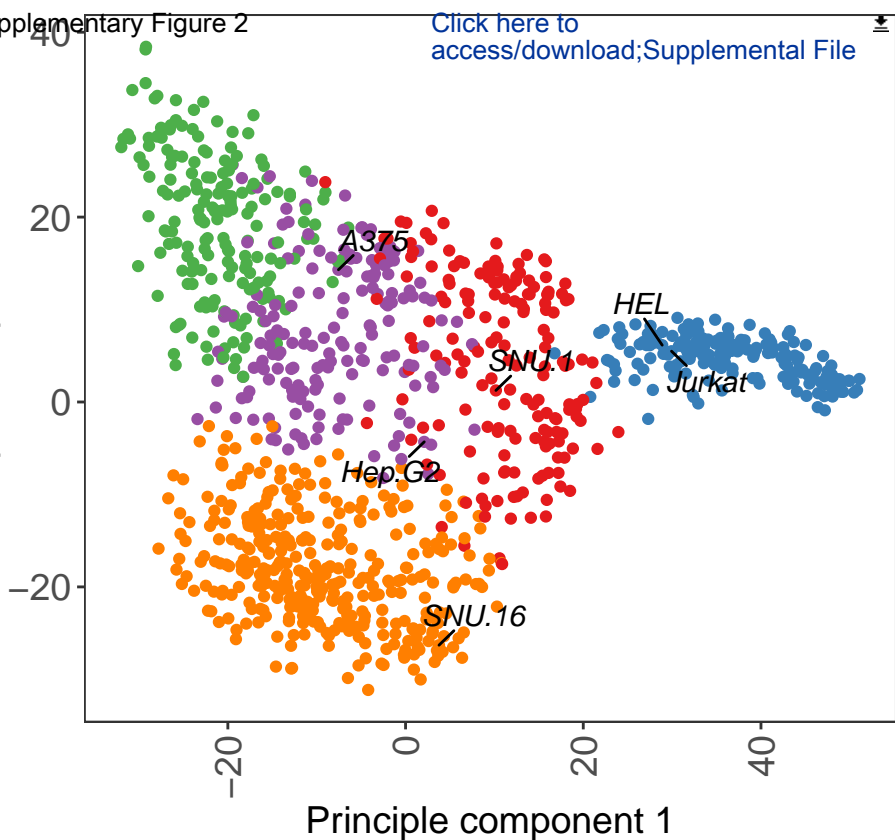
23. Please sort the materials table in alphabetical order.

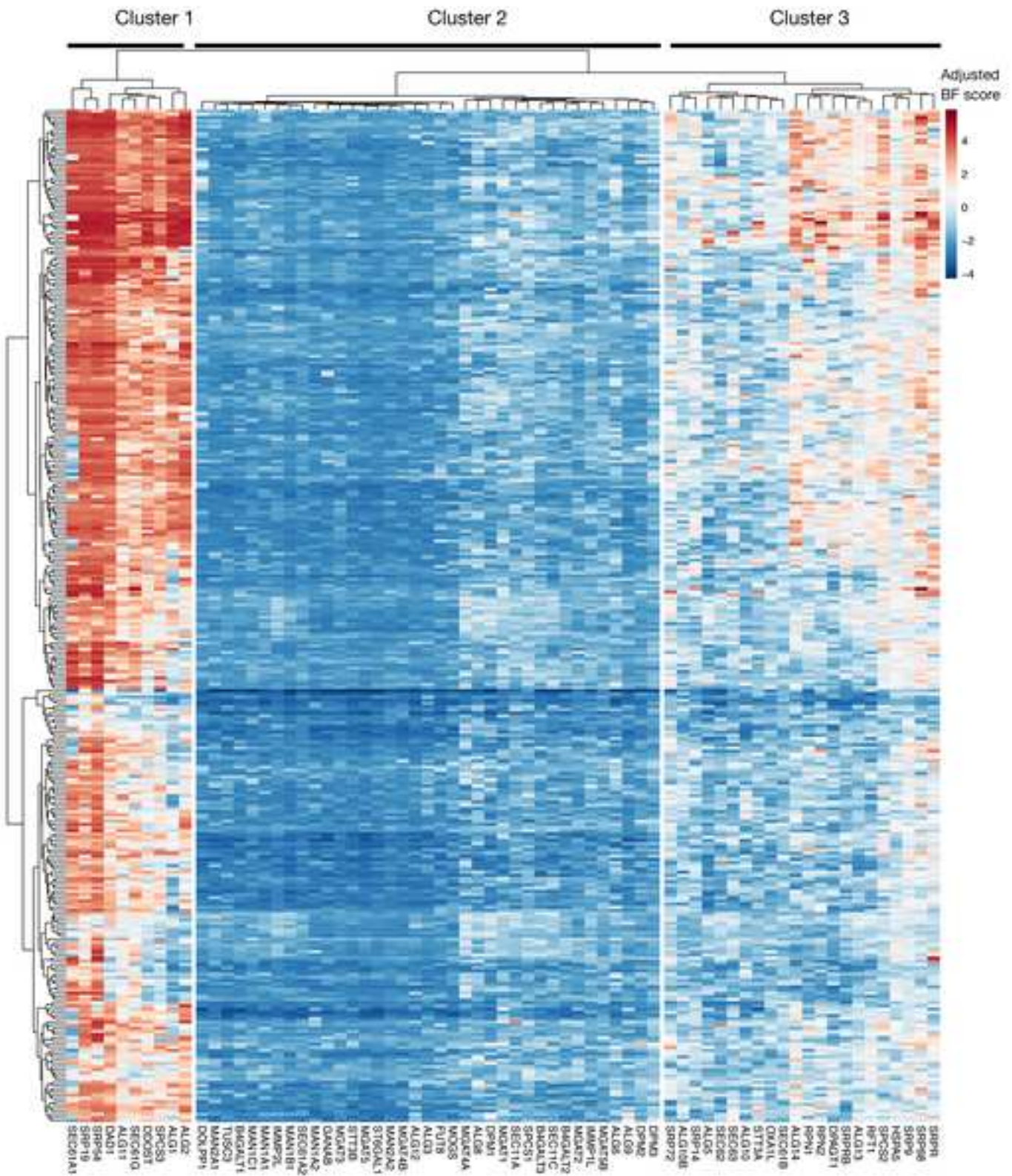
This has now been done.

- Chong, Zheng-Shan, Shuhei Ohnishi, Kosuke Yusa, and Gavin J. Wright. 2018. "Pooled Extracellular Receptor-Ligand Interaction Screening Using CRISPR Activation." *Genome Biology*. <https://doi.org/10.1186/s13059-018-1581-3>.
- Sharma, Sumana. 2018. "Genome-Scale Identification of Cellular Pathways Required for Cell Surface Recognition." University of Cambridge. <https://doi.org/10.17863/CAM.18823>.
- Sharma, Sumana, S. Josefin Bartholdson, Amalie C. M. Couch, Kosuke Yusa, and Gavin J. Wright. 2018. "Genome-Scale Identification of Cellular Pathways Required for Cell Surface Recognition." *Genome Research* 28 (9): 1372–82.
- Wright, Gavin J. 2009. "Signal Initiation in Biological Systems: The Properties and Detection of Transient Extracellular Protein Interactions." *Molecular bioSystems* 5 (12): 1405–12.



Principle component 2





Your MS Excel document "Supplementary table 1.xlsx" cannot be opened and processed. Please see the common list of problems, and suggested resolutions below.

#### External Data

-----

If you are submitting an Excel file, please make sure that your document does not have links to external data. If it does, break the links, save the document and resend. To break the links please do the following.

-On the Edit menu, click Links.

-In the Source list, click the link you want to break.

-To select multiple linked objects, hold down CTRL and click each linked object. To select all links, press CTRL+A.

-Click Break Link.

#### Embedded Macros

-----

Your submission should not contain macros. If they do, an alert box may appear when you open your document (this alert box prevents EM from automatically converting your Excel document into the PDF that Editors and Reviewers will use). You must adjust your Excel document to remove these macros.

#### Excel 2002/Excel XP files

-----

At the present time, EM supports Excel files in Excel 2000 and earlier formats. If you are using a more recent version of MS Excel, try saving your Excel document in a format compatible with Excel 2000, and resubmit to EM.

#### Other Problems

-----

If you are able to get your Excel document to open with no alert boxes appearing, and you have submitted it in Excel 2000 (or earlier) format, and you still see an error indication in your PDF file (where your Excel document should be appearing). please contact the journal via the 'Contact Us' button on the Navigation Bar.'

You will need to reformat your Excel document, and then re-submit it.

Cell_line	sanger_ID	tissue	cluster_info	Growth_condition
201T	SIDM00055	Lung	4	Unknown
22RV1	SIDM00499	Prostate	1	Adherent
23132.87	SIDM00980	Stomach	5	Adherent
42.MG.BA	SIDM00982	Central Nervc	3	Adherent
451Lu	SIDM01240	Skin	4	Adherent
5637	SIDM00807	Bladder	5	Adherent
639.V	SIDM00999	Bladder	4	Adherent
647.V	SIDM00983	Bladder	5	Adherent
697	SIDM01076	Haematopoie	2	Suspension
769.P	SIDM00803	Kidney	4	Adherent
8.MG.BA	SIDM00998	Central Nervc	3	Adherent
8305C	SIDM00997	Thyroid	3	Adherent
8505C	SIDM00996	Thyroid	3	Adherent
A101D	SIDM00801	Skin	4	Adherent
A172	SIDM00799	Central Nervc	3	Adherent
A204	SIDM00798	Soft Tissue	4	Adherent
A2058	SIDM00797	Skin	4	Adherent
A253	SIDM00796	Head and Nei	5	Adherent
A2780	SIDM00210	Ovary	1	Adherent
A3.KAW	SIDM00495	Haematopoie	2	Suspension

A375	SIDM00795	Skin	4 Adherent
A388	SIDM00794	Skin	5 Adherent
A4.Fuk	SIDM00502	Haematopoie	2 Suspension
A427	SIDM00810	Lung	1 Adherent
A431	SIDM00995	Vulva	5 Adherent
A498	SIDM00124	Kidney	4 Adherent
A549	SIDM00903	Lung	4 Adherent
A673	SIDM00848	Bone	1 Adherent
A704	SIDM00849	Kidney	4 Unknown
ABC.1	SIDM00494	Lung	5 Unknown
ACHN	SIDM00123	Kidney	4 Adherent
ACN	SIDM00057	Peripheral Ne	4 Unknown
AGS	SIDM00850	Stomach	5 Adherent
ALL.PO	SIDM00211	Haematopoie	2 Suspension
ALL.SIL	SIDM00994	Haematopoie	2 Suspension
AM.38	SIDM00492	Central Nervc	3 Unknown
AMO.1	SIDM00993	Haematopoie	2 Suspension
AN3.CA	SIDM00901	Endometriurr	1 Adherent
ARH.77	SIDM00900	Haematopoie	2 Suspension
ASH.3	SIDM00491	Thyroid	3 Adherent

AsPC.1	SIDM00899	Pancreas	5 Adherent
ATN.1	SIDM00233	Haematopoie	2 Suspension
AU565	SIDM00898	Breast	5 Adherent
B.CPAP	SIDM00992	Thyroid	3 Adherent
BALL.1	SIDM00287	Haematopoie	2 Suspension
BB30.HNC	SIDM00172	Head and Nei	5 Unknown
BB49.HNC	SIDM00174	Head and Nei	3 Unknown
BB65.RCC	SIDM00191	Kidney	4 Unknown
BC.1	SIDM00896	Haematopoie	2 Suspension
BC.3	SIDM00895	Haematopoie	2 Suspension
BE.13	SIDM01245	Haematopoie	2 Suspension
BE2.M17	SIDM01229	Peripheral Ne	1 Adherent
Becker	SIDM00490	Central Nervc	3 Adherent
BEN	SIDM00990	Lung	1 Adherent
BFTC.905	SIDM00989	Bladder	5 Adherent
BFTC.909	SIDM00988	Kidney	3 Adherent
BHT.101	SIDM00987	Thyroid	4 Adherent
BHY	SIDM00986	Head and Nei	5 Adherent
BICR10	SIDM00077	Head and Nei	3 Adherent
BICR22	SIDM00485	Head and Nei	5 Adherent
BICR31	SIDM00078	Head and Nei	5 Adherent

BICR78	SIDM00501	Head and Nei	5 Adherent
BL.41	SIDM00985	Haematopoie	2 Suspension
BL.70	SIDM00074	Haematopoie	2 Unknown
BONNA.12	SIDM00984	Haematopoie	2 Adherent
BPH.1	SIDM00964	Prostate	5 Adherent
BT.20	SIDM00893	Breast	5 Adherent
BT.474	SIDM00963	Breast	5 Adherent
BT.483	SIDM00892	Breast	5 Adherent
BT.549	SIDM00122	Breast	3 Adherent
BV.173	SIDM00962	Haematopoie	2 Suspension
BxPC.3	SIDM00132	Pancreas	5 Adherent
C.33.A	SIDM00889	Cervix	1 Adherent
C.4.I	SIDM00905	Cervix	5 Adherent
C2BBe1	SIDM01233	Large Intestin	5 Adherent
C32	SIDM00890	Skin	4 Adherent
C3A	SIDM01237	Liver	4 Adherent
Ca.Ski	SIDM00923	Cervix	5 Adherent
CA46	SIDM00907	Haematopoie	2 Suspension
Ca9.22	SIDM00489	Head and Nei	5 Adherent
CADO.ES1	SIDM00961	Bone	1 Adherent

CAKI.1	SIDM00941	Kidney	4 Adherent
CAL.120	SIDM00940	Breast	4 Adherent
CAL.12T	SIDM00939	Lung	5 Adherent
CAL.148	SIDM00938	Breast	1 Adherent
CAL.27	SIDM00937	Head and Neck	5 Adherent
CAL.29	SIDM00936	Bladder	5 Adherent
CAL.33	SIDM00942	Head and Neck	5 Adherent
CAL.39	SIDM00935	Vulva	5 Adherent
CAL.51	SIDM00933	Breast	4 Adherent
CAL.54	SIDM00932	Kidney	4 Adherent
CAL.62	SIDM00931	Thyroid	4 Adherent
CAL.72	SIDM00930	Bone	3 Adherent
CAL.78	SIDM00929	Bone	3 Adherent
CAL.85.1	SIDM00928	Breast	5 Adherent
Calu.3	SIDM00922	Lung	5 Adherent
Calu.6	SIDM00921	Lung	4 Adherent
CAMA.1	SIDM00920	Breast	5 Adherent
Caov.3	SIDM00919	Ovary	5 Adherent
Caov.4	SIDM00918	Ovary	5 Adherent
CAPAN.1	SIDM00934	Pancreas	5 Adherent
CAPAN.2	SIDM00943	Pancreas	5 Adherent

CaR.1	SIDM00488	Large Intestin	5 Adherent
CAS.1	SIDM00212	Central Nervc	3 Adherent
CCF.STTG1	SIDM00131	Central Nervc	3 Adherent
CCK.81	SIDM00487	Large Intestin	5 Adherent
CCRF.CEM	SIDM00121	Haematopoie	2 Suspension
CESS	SIDM00917	Haematopoie	2 Suspension
CFPAC.1	SIDM00130	Pancreas	5 Adherent
CGTH.W.1	SIDM01248	Thyroid	3 Unknown
ChaGo.K.1	SIDM00924	Lung	5 Adherent
CHL.1	SIDM01228	Skin	4 Adherent
CHP.126	SIDM00286	Peripheral Ne	1 Adherent
CHP.134	SIDM00285	Peripheral Ne	1 Adherent
CHP.212	SIDM00910	Peripheral Ne	3 Adherent
CHSA0011	SIDM00051	Bone	3 Unknown
CHSA0108	SIDM00050	Bone	3 Unknown
CHSA8926	SIDM00176	Bone	3 Unknown
CL.11	SIDM00944	Large Intestin	5 Adherent
CL.34	SIDM00945	Large Intestin	5 Adherent
CL.40	SIDM00960	Large Intestin	5 Adherent
CMK	SIDM00959	Haematopoie	2 Suspension

CML.T1	SIDM00958	Haematopoie	2 Suspension
COLO.201	SIDM00823	Large Intestin	5 Semi-Adherent
COLO.205	SIDM00826	Large Intestin	5 Semi-Adherent
COLO.320	SIDM00071	Large Intestin	1 Unknown
COLO.320.HS	SIDM00842	Large Intestin	1 Semi-Adherent
COLO.668	SIDM00143	Lung	1 Semi-Adherent
COLO.678	SIDM00957	Large Intestin	5 Adherent
COLO.679	SIDM00284	Skin	4 Adherent
COLO.680N	SIDM00956	Esophagus	5 Adherent
COLO.684	SIDM00142	Endometriurr	1 Suspension
COLO.783	SIDM00955	Skin	4 Adherent
COLO.792	SIDM00140	Skin	4 Adherent
COLO.800	SIDM00515	Skin	4 Adherent
COLO.824	SIDM00954	Breast	1 Adherent
COLO.829	SIDM00909	Skin	4 Adherent
COR.L105	SIDM00513	Lung	5 Semi-Adherent
COR.L23	SIDM00512	Lung	5 Adherent
COR.L279	SIDM00511	Lung	1 Suspension
COR.L303	SIDM00510	Lung	1 Suspension
COR.L311	SIDM00509	Lung	1 Suspension
COR.L32	SIDM00508	Lung	1 Suspension

COR.L321	SIDM00507	Lung	1 Suspension
COR.L88	SIDM00506	Lung	1 Semi-Adherent
COR.L95	SIDM00521	Lung	1 Suspension
CP50.MEL.B	SIDM00175	Skin	4 Unknown
CP66.MEL	SIDM00190	Skin	4 Unknown
CP67.MEL	SIDM00189	Skin	4 Unknown
CPC.N	SIDM00953	Lung	1 Suspension
CRO.AP2	SIDM00952	Haematopoie	2 Suspension
CRO.AP3	SIDM00070	Haematopoie	2 Suspension
CS1	SIDM00199	Bone	3 Adherent
CTB.1	SIDM00283	Haematopoie	2 Suspension
CTV.1	SIDM00951	Haematopoie	2 Suspension
CW.2	SIDM00282	Large Intestin	5 Adherent
D.247MG	SIDM00622	Central Nervc	3 Unknown
D.263MG	SIDM00732	Central Nervc	3 Unknown
D.283MED	SIDM00888	Central Nervc	1 Suspension
D.336MG	SIDM00620	Central Nervc	3 Unknown
D.392MG	SIDM00542	Central Nervc	3 Unknown
D.423MG	SIDM00209	Central Nervc	3 Unknown
D.502MG	SIDM00208	Central Nervc	3 Unknown

D.542MG	SIDM00207	Central Nervc	3 Unknown
D.566MG	SIDM00206	Central Nervc	3 Unknown
DAN.G	SIDM00950	Pancreas	5 Adherent
Daoy	SIDM00887	Central Nervc	3 Adherent
Daudi	SIDM00845	Haematopoie	2 Suspension
DB	SIDM00886	Haematopoie	2 Suspension
DBTRG.05MG	SIDM00867	Central Nervc	3 Adherent
DEL	SIDM00949	Haematopoie	2 Suspension
Detroit562	SIDM00831	Head and Nei	5 Adherent
DG.75	SIDM00948	Haematopoie	2 Suspension
DiFi	SIDM00049	Large Intestin	5 Unknown
DJM.1	SIDM00281	Skin	5 Adherent
DK.MG	SIDM00947	Central Nervc	3 Adherent
DMS.114	SIDM00865	Lung	1 Adherent
DMS.273	SIDM00522	Lung	1 Adherent
DMS.53	SIDM00523	Lung	1 Adherent
DMS.79	SIDM00524	Lung	1 Suspension
DND.41	SIDM00946	Haematopoie	2 Suspension
DOHH.2	SIDM00981	Haematopoie	2 Suspension
DOK	SIDM00540	Head and Nei	5 Adherent
DoTc2.4510	SIDM00864	Cervix	5 Adherent

DOV13	SIDM00969	Ovary	3 Adherent
DSH1	SIDM00970	Bladder	5 Adherent
DU.145	SIDM00120	Prostate	5 Adherent
DU.4475	SIDM01001	Breast	1 Suspension
EB.1	SIDM00862	Haematopoie	2 Suspension
EB.3	SIDM00793	Haematopoie	2 Suspension
EB2	SIDM00863	Haematopoie	2 Suspension
EBC.1	SIDM00486	Lung	5 Unknown
EC.GI.10	SIDM00278	Esophagus	5 Adherent
ECC10	SIDM00280	Stomach	1 Suspension
ECC12	SIDM00279	Stomach	1 Adherent
EFM.19	SIDM01056	Breast	5 Adherent
EFM.192A	SIDM01002	Breast	5 Adherent
EFO.21	SIDM01052	Ovary	4 Adherent
EFO.27	SIDM01051	Ovary	5 Adherent
EGI.1	SIDM01050	Biliary Tract	5 Adherent
EHEB	SIDM01049	Haematopoie	2 Suspension
EJM	SIDM01048	Haematopoie	2 Unknown
EKVX	SIDM00119	Lung	5 Adherent
EM.2	SIDM01047	Haematopoie	2 Suspension

EMC.BAC.1	SIDM00048	Lung	5 Unknown
EMC.BAC.2	SIDM00047	Lung	5 Unknown
EN	SIDM01045	Endometriurr	1 Adherent
EoL.1.cell	SIDM00277	Haematopoie	2 Suspension
EPLC.272H	SIDM01044	Lung	5 Adherent
ES.2	SIDM00861	Ovary	3 Adherent
ES1	SIDM00266	Bone	1 Unknown
ES3	SIDM00265	Bone	1 Unknown
ES4	SIDM00264	Bone	1 Unknown
ES5	SIDM00263	Bone	1 Unknown
ES6	SIDM00262	Bone	1 Unknown
ES7	SIDM00269	Bone	1 Unknown
ES8	SIDM00261	Bone	1 Unknown
ESO26	SIDM00539	Esophagus	5 Suspension
ESO51	SIDM00538	Esophagus	5 Suspension
ESS.1	SIDM01043	Endometriurr	3 Adherent
ETK.1	SIDM00276	Biliary Tract	5 Unknown
EVSA.T	SIDM01042	Breast	5 Adherent
EW.11	SIDM00203	Bone	1 Unknown
EW.12	SIDM00200	Bone	1 Unknown
EW.13	SIDM00621	Bone	1 Unknown

EW.16	SIDM00198	Bone	1 Unknown
EW.18	SIDM00197	Bone	1 Unknown
EW.22	SIDM00196	Bone	1 Unknown
EW.24	SIDM00195	Bone	1 Unknown
EW.3	SIDM00202	Bone	1 Unknown
EW.7	SIDM00201	Bone	1 Unknown
FADU	SIDM00860	Head and Neck	5 Adherent
Farage	SIDM00859	Haematopoietic	2 Suspension
FLO.1	SIDM01041	Esophagus	4 Adherent
FTC.133	SIDM00213	Thyroid	3 Adherent
FTC.238	SIDM00228	Thyroid	3 Adherent
FU.OV.1	SIDM01040	Ovary	4 Adherent
FU97	SIDM00503	Stomach	4 Adherent
G.292.Clone.1	SIDM00858	Bone	3 Adherent
G.361	SIDM00857	Skin	4 Adherent
G.401	SIDM00856	Kidney	1 Adherent
G.402	SIDM00855	Soft Tissue	4 Adherent
G.MEL	SIDM00129	Skin	4 Unknown
GA.10	SIDM00854	Haematopoietic	2 Suspension
GAK	SIDM00543	Skin	4 Adherent

GAMG	SIDM01055	Central Nervc	3 Adherent
GB.1	SIDM00581	Central Nervc	3 Unknown
GCIY	SIDM00274	Stomach	4 Adherent
GCT	SIDM00853	Soft Tissue	3 Adherent
GDM.1	SIDM01057	Haematopoie	2 Suspension
Geo	SIDM00068	Large Intestin	5 Unknown
GI.1	SIDM00273	Central Nervc	3 Adherent
GI.ME.N	SIDM00227	Peripheral Ne	3 Adherent
GMS.10	SIDM01074	Central Nervc	3 Adherent
GOTO	SIDM00544	Peripheral Ne	1 Adherent
GP2d	SIDM00536	Large Intestin	5 Adherent
GP5d	SIDM00537	Large Intestin	5 Adherent
GR.ST	SIDM01259	Haematopoie	2 Suspension
GRANTA.519	SIDM01058	Haematopoie	2 Suspension
GT3TKB	SIDM01231	Lung	4 Adherent
H.EMC.SS	SIDM00535	Bone	1 Adherent
H4	SIDM00852	Central Nervc	3 Adherent
H9	SIDM01251	Haematopoie	2 Suspension
HA7.RCC	SIDM00188	Kidney	4 Unknown
HAL.01	SIDM00289	Haematopoie	2 Suspension
HARA	SIDM00598	Lung	5 Unknown

HC.1	SIDM01073	Haematopoie	2 Suspension
HCC.15	SIDM01072	Lung	5 Adherent
HCC.33	SIDM01071	Lung	1 Adherent
HCC.366	SIDM01070	Lung	5 Adherent
HCC.44	SIDM01069	Lung	4 Adherent
HCC.78	SIDM01068	Lung	5 Adherent
HCC.827	SIDM01067	Lung	5 Adherent
HCC1143	SIDM00866	Breast	5 Adherent
HCC1187	SIDM00885	Breast	5 Semi-Adherent
HCC1395	SIDM00884	Breast	3 Adherent
HCC1419	SIDM00882	Breast	5 Adherent
HCC1428	SIDM00881	Breast	5 Adherent
HCC1500	SIDM00879	Breast	5 Adherent
HCC1569	SIDM00878	Breast	4 Semi-Adherent
HCC1599	SIDM00877	Breast	5 Suspension
HCC1806	SIDM00875	Breast	5 Adherent
HCC1937	SIDM00874	Breast	5 Semi-Adherent
HCC1954	SIDM00872	Breast	5 Adherent
HCC202	SIDM00870	Breast	5 Semi-Adherent
HCC2157	SIDM00774	Breast	5 Suspension

HCC2218	SIDM00772	Breast	5 Suspension
HCC2998	SIDM00118	Large Intestin	5 Unknown
HCC38	SIDM00675	Breast	5 Semi-Adherent
HCC70	SIDM00673	Breast	5 Adherent
HCE.4	SIDM00052	Esophagus	3 Unknown
HCT.116	SIDM00783	Large Intestin	5 Adherent
HCT.15	SIDM00789	Large Intestin	5 Adherent
HD.MY.Z	SIDM01063	Haematopoie	3 Unknown
HDLM.2	SIDM01064	Haematopoie	2 Suspension
HDQ.P1	SIDM01062	Breast	5 Adherent
HEC.1	SIDM00596	Endometriurr	4 Adherent
HEC.1.A	SIDM00595	Endometriurr	5 Adherent
HEC.1.B	SIDM01241	Endometriurr	4 Adherent
HEL	SIDM00594	Haematopoie	2 Suspension
HEL.92.1.7	SIDM00593	Haematopoie	2 Unknown
HeLa	SIDM00846	Cervix	4 Adherent
Hep.G2	SIDM00904	Liver	4 Adherent
Hep3B2.1.7	SIDM00672	Liver	4 Adherent
Hey	SIDM00968	Ovary	3 Adherent
Hey.A8	SIDM01215	Ovary	4 Adherent
HGC.27	SIDM00290	Stomach	1 Adherent

HH	SIDM00671	Haematopoie	2 Suspension
HL.60	SIDM00829	Haematopoie	2 Suspension
HLE	SIDM00064	Liver	4 Unknown
HLF	SIDM00065	Liver	4 Unknown
HMCB	SIDM00914	Skin	4 Adherent
HMV.II	SIDM00291	Skin	4 Adherent
HN	SIDM01061	Head and Nei	5 Adherent
HO.1.N.1	SIDM00592	Head and Nei	5 Adherent
HO.1.u.1	SIDM00591	Head and Nei	5 Unknown
HOP.62	SIDM00133	Lung	3 Unknown
HOP.92	SIDM00134	Lung	3 Unknown
HOS	SIDM00806	Bone	3 Adherent
HPAC	SIDM00670	Pancreas	5 Adherent
HPAF.II	SIDM00669	Pancreas	5 Adherent
Hs.294.T	SIDM00800	Skin	3 Adherent
Hs.445	SIDM00668	Haematopoie	2 Suspension
Hs.578.T	SIDM00135	Breast	3 Adherent
Hs.633T	SIDM00667	Soft Tissue	3 Adherent
Hs.683	SIDM00666	Central Nervc	3 Adherent
Hs.746T	SIDM00665	Stomach	3 Adherent

Hs.766T	SIDM00664	Pancreas	5 Adherent
Hs.939.T	SIDM00663	Skin	4 Adherent
Hs.940.T	SIDM00662	Skin	3 Adherent
Hs.944.T	SIDM00063	Skin	4 Unknown
HSC.2	SIDM00590	Head and Nei	5 Adherent
HSC.3	SIDM00589	Head and Nei	5 Unknown
HSC.39	SIDM00045	Stomach	5 Unknown
HSC.4	SIDM00588	Head and Nei	5 Adherent
HT	SIDM00661	Haematopoie	2 Suspension
HT.1080	SIDM00828	Soft Tissue	3 Adherent
HT.115	SIDM00534	Large Intestin	5 Adherent
HT.1197	SIDM00676	Bladder	5 Adherent
HT.1376	SIDM00678	Bladder	5 Adherent
HT.144	SIDM00695	Skin	4 Adherent
HT.29	SIDM00136	Large Intestin	5 Adherent
HT.3	SIDM00679	Cervix	5 Adherent
HT55	SIDM00541	Large Intestin	5 Adherent
HTC.C3	SIDM00496	Thyroid	5 Adherent
HuCCT1	SIDM00587	Biliary Tract	5 Unknown
huH.1	SIDM00586	Liver	4 Unknown
HuH.6	SIDM00062	Liver	4 Unknown

HUH.6.clone5	SIDM01257	Liver	1 Unknown
HuH.7	SIDM00585	Liver	4 Unknown
HuO.3N1	SIDM00584	Bone	3 Unknown
HuO9	SIDM00599	Bone	3 Adherent
HuP.T3	SIDM00533	Pancreas	5 Adherent
HuP.T4	SIDM00531	Pancreas	5 Adherent
HuT.78	SIDM00851	Haematopoie	2 Suspension
HuTu.80	SIDM00694	Small Intestin	1 Adherent
IA.LM	SIDM00306	Lung	4 Adherent
IGR.1	SIDM01060	Skin	4 Adherent
IGR.37	SIDM01066	Skin	4 Adherent
IGROV.1	SIDM00151	Ovary	4 Unknown
IHH.4	SIDM00600	Thyroid	4 Unknown
IM.9	SIDM00632	Haematopoie	2 Suspension
IM.95	SIDM00602	Stomach	5 Unknown
IMR.32	SIDM00226	Peripheral Ne	1 Adherent
IMR.5	SIDM01206	Peripheral Ne	1 Adherent
IPC.298	SIDM01059	Skin	4 Adherent
Ishikawa(Her	SIDM00037	Endometriurr	5 Unknown
IST.MEL1	SIDM00225	Skin	3 Adherent

IST.MES1	SIDM00224	Lung	3 Adherent
IST.SL1	SIDM00223	Lung	1 Semi-Adherent
IST.SL2	SIDM00230	Lung	1 Suspension
J82	SIDM00693	Bladder	3 Adherent
JAR	SIDM01039	Placenta	5 Adherent
JEG.3	SIDM01218	Placenta	5 Adherent
JEKO.1	SIDM01038	Haematopoie	2 Suspension
JHH.1	SIDM00618	Liver	5 Adherent
JHH.2	SIDM00617	Liver	3 Unknown
JHH.4	SIDM00616	Liver	4 Adherent
JHH.6	SIDM00615	Liver	3 Adherent
JHH.7	SIDM00614	Liver	4 Adherent
JHOS.2	SIDM00305	Ovary	5 Adherent
JHOS.3	SIDM00304	Ovary	5 Adherent
JHOS.4	SIDM00303	Ovary	5 Adherent
JHU.011	SIDM00128	Head and Nei	5 Unknown
JHU.022	SIDM00127	Head and Nei	5 Unknown
JHU.029	SIDM00690	Head and Nei	5 Adherent
JIMT.1	SIDM01037	Breast	5 Adherent
JiyoyeP.2003	SIDM00808	Haematopoie	2 Suspension
JJN.3	SIDM01036	Haematopoie	2 Suspension

JM1	SIDM00689	Haematopoie	2 Suspension
JSC.1	SIDM00688	Haematopoie	2 Suspension
Jurkat	SIDM01016	Haematopoie	2 Suspension
JURL.MK1	SIDM01015	Haematopoie	2 Suspension
JVM.2	SIDM01013	Haematopoie	2 Suspension
JVM.3	SIDM01012	Haematopoie	2 Suspension
K.562	SIDM00791	Haematopoie	2 Suspension
K052	SIDM00018	Haematopoie	2 Suspension
K2	SIDM00126	Skin	4 Unknown
K5	SIDM00056	Thyroid	3 Unknown
KALS.1	SIDM00613	Central Nervc	3 Unknown
KARPAS.1106	SIDM01011	Haematopoie	2 Suspension
KARPAS.231	SIDM01017	Haematopoie	2 Suspension
KARPAS.299	SIDM01010	Haematopoie	2 Suspension
KARPAS.422	SIDM01008	Haematopoie	2 Suspension
KARPAS.45	SIDM01007	Haematopoie	2 Suspension
KARPAS.620	SIDM01006	Haematopoie	2 Suspension
KASUMI.1	SIDM01005	Haematopoie	2 Suspension
KATOIII	SIDM00687	Stomach	5 Semi-Adherent
KCL.22	SIDM01004	Haematopoie	2 Suspension

KE.37	SIDM01003	Haematopoie	2 Suspension
KELLY	SIDM01009	Peripheral Ne	1 Adherent
KG.1	SIDM00785	Haematopoie	2 Suspension
KG.1.C	SIDM00784	Central Nervc	3 Unknown
KGN	SIDM00302	Ovary	3 Adherent
KINGS.1	SIDM00612	Central Nervc	3 Unknown
KLE	SIDM00686	Endometrium	4 Adherent
KM.H2	SIDM01018	Haematopoie	2 Suspension
KM12	SIDM00150	Large Intestin	5 Unknown
KMH.2	SIDM00619	Thyroid	3 Unknown
KMOE.2	SIDM01019	Haematopoie	2 Suspension
KMRC.1	SIDM00611	Kidney	4 Adherent
KMRC.20	SIDM00609	Kidney	4 Adherent
KMS.11	SIDM00608	Haematopoie	2 Unknown
KMS.12.BM	SIDM01020	Haematopoie	2 Suspension
KNS.42	SIDM00607	Central Nervc	4 Unknown
KNS.62	SIDM00606	Lung	5 Adherent
KNS.81	SIDM00605	Central Nervc	3 Unknown
KNS.81.FD	SIDM01210	Central Nervc	3 Unknown
KON	SIDM00604	Head and Nei	5 Adherent
KOPN.8	SIDM01034	Haematopoie	2 Suspension

KOSC.2	SIDM00603	Head and Nei	5 Unknown
KP.1N	SIDM00583	Pancreas	4 Adherent
KP.2	SIDM00582	Pancreas	5 Adherent
KP.3	SIDM00571	Pancreas	5 Adherent
KP.4	SIDM00301	Pancreas	4 Adherent
KP.N.RT.BM.1	SIDM00557	Peripheral Ne	1 Unknown
KP.N.YN	SIDM00556	Peripheral Ne	1 Unknown
KP.N.YS	SIDM00157	Peripheral Ne	1 Unknown
KPL.1	SIDM00147	Breast	5 Unknown
KS.1	SIDM00555	Central Nervc	3 Adherent
KU.19.19	SIDM01033	Bladder	5 Adherent
KU812	SIDM00308	Haematopoie	2 Suspension
KURAMOCHI	SIDM00554	Ovary	4 Unknown
KY821	SIDM00553	Haematopoie	2 Unknown
KYAE.1	SIDM00530	Esophagus	5 Adherent
KYM.1	SIDM00552	Soft Tissue	1 Unknown
KYSE.140	SIDM01032	Esophagus	5 Adherent
KYSE.150	SIDM01031	Esophagus	5 Adherent
KYSE.180	SIDM01030	Esophagus	5 Adherent
KYSE.220	SIDM00551	Esophagus	5 Adherent

KYSE.270	SIDM01029	Esophagus	5 Adherent
KYSE.30	SIDM00015	Esophagus	5 Unknown
KYSE.410	SIDM01028	Esophagus	5 Adherent
KYSE.450	SIDM01027	Esophagus	5 Adherent
KYSE.50	SIDM00550	Esophagus	5 Unknown
KYSE.510	SIDM01026	Esophagus	5 Adherent
KYSE.520	SIDM01025	Esophagus	5 Adherent
KYSE.70	SIDM01024	Esophagus	5 Adherent
L.1236	SIDM00313	Haematopoie	2 Unknown
L.363	SIDM00312	Haematopoie	2 Suspension
L.428	SIDM00327	Haematopoie	2 Suspension
L.540	SIDM00329	Haematopoie	2 Suspension
LAMA.84	SIDM00346	Haematopoie	2 Suspension
LAN.6	SIDM00330	Peripheral Ne	1 Semi-Adherent
LB1047.RCC	SIDM00187	Kidney	3 Unknown
LB2241.RCC	SIDM00186	Kidney	4 Unknown
LB2518.MEL	SIDM00185	Skin	4 Unknown
LB373.MEL.D	SIDM00184	Skin	4 Unknown
LB647.SCLC	SIDM00183	Lung	1 Unknown
LB771.HNC	SIDM00182	Head and Nei	4 Unknown
LB831.BLC	SIDM00181	Bladder	3 Unknown

LB996.RCC	SIDM00180	Kidney	4 Unknown
LC.1.sq	SIDM00300	Lung	5 Adherent
LC.1F	SIDM00298	Lung	5 Suspension
LC.2.ad	SIDM00297	Lung	5 Adherent
LC4.1	SIDM00549	Haematopoie	2 Suspension
LCLC.103H	SIDM00345	Lung	3 Adherent
LCLC.97TM1	SIDM00344	Lung	5 Adherent
LIM1215	SIDM00014	Large Intestin	5 Unknown
LK.2	SIDM00548	Lung	1 Adherent
LN.18	SIDM00685	Central Nervc	3 Adherent
LN.229	SIDM00684	Central Nervc	3 Adherent
LN.405	SIDM00343	Central Nervc	3 Adherent
LNCaP.Clone.	SIDM00683	Prostate	1 Semi-Adherent
LNZTA3WT4	SIDM01220	Central Nervc	3 Adherent
LOU.NH91	SIDM00341	Lung	3 Adherent
LOUCY	SIDM00342	Haematopoie	2 Suspension
LoVo	SIDM00839	Large Intestin	5 Adherent
LOXIMVI	SIDM00149	Skin	4 Unknown
LP.1	SIDM00340	Haematopoie	2 Suspension
LS.1034	SIDM00681	Large Intestin	5 Adherent

LS.123	SIDM00776	Large Intestin	5 Adherent
LS.180	SIDM00680	Large Intestin	5 Adherent
LS.411N	SIDM00660	Large Intestin	5 Adherent
LS.513	SIDM00677	Large Intestin	5 Adherent
LU.134.A	SIDM00296	Lung	1 Suspension
LU.135	SIDM00294	Lung	1 Semi-Adherent
LU.139	SIDM00293	Lung	1 Suspension
LU.165	SIDM00292	Lung	1 Suspension
LU.65	SIDM00547	Lung	5 Unknown
LU.99	SIDM00546	Lung	4 Unknown
LXF.289	SIDM00339	Lung	3 Adherent
M059J	SIDM00659	Central Nervc	3 Adherent
M059K	SIDM00639	Central Nervc	3 Adherent
M14	SIDM00003	Skin	4 Unknown
Malme.3M	SIDM00011	Skin	4 Unknown
MC.1010	SIDM00009	Haematopoie	2 Unknown
MC.CAR	SIDM00636	Haematopoie	2 Suspension
MC.IXC	SIDM01246	Bone	1 Adherent
MC116	SIDM00637	Haematopoie	2 Suspension
MCAS	SIDM00008	Ovary	5 Unknown
MCC13	SIDM00007	Skin	3 Unknown

MCC26	SIDM00006	Skin	3 Unknown
MCF7	SIDM00148	Breast	5 Adherent
MDA.MB.134	SIDM00005	Breast	1 Adherent
MDA.MB.157	SIDM00529	Breast	3 Adherent
MDA.MB.175	SIDM00633	Breast	5 Semi-Adherent
MDA.MB.231	SIDM00146	Breast	4 Adherent
MDA.MB.330	SIDM00631	Breast	5 Adherent
MDA.MB.361	SIDM00528	Breast	5 Adherent
MDA.MB.415	SIDM00630	Breast	5 Adherent
MDA.MB.435	SIDM01222	Skin	4 Unknown
MDA.MB.436	SIDM00629	Breast	4 Adherent
MDA.MB.453	SIDM00272	Breast	5 Adherent
MDA.MB.468	SIDM00628	Breast	5 Adherent
MDST8	SIDM00527	Large Intestin	3 Adherent
ME.1	SIDM00338	Haematopoie	2 Suspension
ME.180	SIDM00627	Cervix	5 Adherent
MEC.1	SIDM00001	Haematopoie	2 Unknown
MEG.01	SIDM00526	Haematopoie	2 Semi-Adherent
MEL.HO	SIDM00337	Skin	4 Adherent
MEL.JUSO	SIDM00336	Skin	4 Adherent

MES.SA	SIDM00625	Uterus	4 Adherent
Mewo	SIDM00545	Skin	4 Adherent
MFE.280	SIDM00335	Endometriurr	1 Adherent
MFE.296	SIDM00334	Endometriurr	4 Adherent
MFE.319	SIDM00333	Endometriurr	5 Adherent
MFH.ino	SIDM00299	Soft Tissue	3 Adherent
MFM.223	SIDM00332	Breast	5 Adherent
MG.63	SIDM00525	Bone	3 Adherent
MHH.CALL.2	SIDM00331	Haematopoie	2 Suspension
MHH.CALL.4	SIDM00376	Haematopoie	2 Suspension
MHH.ES.1	SIDM00386	Bone	1 Adherent
MHH.NB.11	SIDM00387	Peripheral Ne	1 Adherent
MHH.PREB.1	SIDM00388	Haematopoie	2 Suspension
MIA.PaCa.2	SIDM00505	Pancreas	4 Semi-Adherent
MKN1	SIDM00271	Stomach	3 Adherent
MKN28	SIDM00260	Stomach	5 Adherent
MKN45	SIDM00247	Stomach	5 Adherent
MKN7	SIDM00560	Stomach	5 Unknown
MKN74	SIDM00248	Stomach	5 Adherent
ML.1	SIDM00440	Haematopoie	3 Suspension
ML.2	SIDM00441	Haematopoie	2 Suspension

MLMA	SIDM00561	Haematopoie	2 Unknown
MM1S	SIDM01265	Haematopoie	2 Semi-Adherent
MMAc.SF	SIDM01242	Skin	4 Adherent
MN.60	SIDM00438	Haematopoie	2 Suspension
Mo.T	SIDM00623	Haematopoie	2 Suspension
MOG.G.CCM	SIDM00532	Central Nervc	3 Adherent
MOG.G.UVW	SIDM00504	Central Nervc	3 Adherent
MOLM.13	SIDM00437	Haematopoie	2 Suspension
MOLM.16	SIDM00435	Haematopoie	2 Suspension
MOLP.8	SIDM00434	Haematopoie	2 Suspension
MOLT.13	SIDM00433	Haematopoie	2 Suspension
MOLT.16	SIDM00431	Haematopoie	2 Suspension
MOLT.3	SIDM00153	Haematopoie	2 Unknown
MOLT.4	SIDM00145	Haematopoie	2 Suspension
MONO.MAC.:	SIDM01022	Haematopoie	2 Suspension
MONO.MAC.:	SIDM01023	Haematopoie	2 Suspension
MPP.89	SIDM00222	Lung	3 Adherent
MRK.nu.1	SIDM00562	Breast	5 Unknown
MS.1	SIDM00408	Skin	1 Suspension
MS751	SIDM00638	Cervix	5 Adherent

MSTO.211H	SIDM00640	Lung	4 Adherent
MV.4.11	SIDM00657	Haematopoie	2 Suspension
MY.M12	SIDM01217	Haematopoie	2 Suspension
MZ1.PC	SIDM00053	Pancreas	5 Unknown
MZ2.MEL	SIDM00179	Skin	4 Unknown
MZ7.mel	SIDM00516	Skin	4 Unknown
NALM.6	SIDM00429	Haematopoie	2 Suspension
NAMALWA	SIDM00641	Haematopoie	2 Suspension
NB(TU)1.10	SIDM00579	Peripheral Ne	1 Unknown
NB1	SIDM00578	Peripheral Ne	1 Unknown
NB10	SIDM00259	Peripheral Ne	1 Unknown
NB12	SIDM00258	Peripheral Ne	1 Unknown
NB13	SIDM00257	Peripheral Ne	1 Unknown
NB14	SIDM00256	Peripheral Ne	1 Unknown
NB17	SIDM00255	Peripheral Ne	1 Unknown
NB4	SIDM00428	Haematopoie	2 Suspension
NB5	SIDM00254	Peripheral Ne	1 Unknown
NB6	SIDM00253	Peripheral Ne	1 Unknown
NB69	SIDM00244	Peripheral Ne	1 Adherent
NB7	SIDM00156	Peripheral Ne	1 Unknown
NBsusSR	SIDM00002	Peripheral Ne	1 Unknown

NCC010	SIDM00231	Kidney	4 Unknown
NCC021	SIDM00232	Kidney	4 Unknown
NCCIT	SIDM00655	Testis	1 Adherent
NCI.H1048	SIDM00654	Lung	1 Adherent
NCI.H1092	SIDM00653	Lung	1 Suspension
NCI.H1105	SIDM00652	Lung	1 Suspension
NCI.H1155	SIDM00651	Lung	1 Suspension
NCI.H128	SIDM00650	Lung	1 Suspension
NCI.H1299	SIDM00649	Lung	4 Adherent
NCI.H1304	SIDM00648	Lung	1 Suspension
NCI.H1341	SIDM00646	Lung	1 Suspension
NCI.H1355	SIDM00645	Lung	4 Suspension
NCI.H1395	SIDM00644	Lung	5 Adherent
NCI.H1417	SIDM00642	Lung	1 Suspension
NCI.H1435	SIDM00658	Lung	5 Semi-Adherent
NCI.H1436	SIDM00697	Lung	1 Suspension
NCI.H1437	SIDM00734	Lung	5 Adherent
NCI.H146	SIDM00698	Lung	1 Suspension
NCI.H1563	SIDM00751	Lung	3 Adherent
NCI.H1568	SIDM00750	Lung	5 Adherent

NCI.H1573	SIDM00749	Lung	5 Adherent
NCI.H1581	SIDM00748	Lung	1 Semi-Adherent
NCI.H1623	SIDM00747	Lung	5 Semi-Adherent
NCI.H1648	SIDM00746	Lung	5 Adherent
NCI.H1650	SIDM00745	Lung	5 Adherent
NCI.H1651	SIDM00744	Lung	4 Adherent
NCI.H1666	SIDM00743	Lung	5 Semi-Adherent
NCI.H1688	SIDM00792	Lung	1 Adherent
NCI.H1693	SIDM00742	Lung	5 Adherent
NCI.H1694	SIDM00741	Lung	1 Suspension
NCI.H1703	SIDM00740	Lung	4 Adherent
NCI.H1734	SIDM00739	Lung	5 Adherent
NCI.H1755	SIDM00738	Lung	3 Semi-Adherent
NCI.H1770	SIDM00737	Lung	1 Suspension
NCI.H1781	SIDM00754	Lung	5 Adherent
NCI.H1792	SIDM00771	Lung	3 Adherent
NCI.H1793	SIDM00755	Lung	4 Adherent
NCI.H1836	SIDM00770	Lung	1 Suspension
NCI.H1838	SIDM00769	Lung	5 Adherent
NCI.H1869	SIDM00768	Lung	5 Adherent
NCI.H187	SIDM00767	Lung	1 Suspension

NCI.H1876	SIDM00766	Lung	1 Adherent
NCI.H1915	SIDM00763	Lung	4 Adherent
NCI.H1944	SIDM00762	Lung	5 Adherent
NCI.H196	SIDM00761	Lung	3 Adherent
NCI.H1963	SIDM00760	Lung	1 Suspension
NCI.H1975	SIDM00759	Lung	4 Adherent
NCI.H1993	SIDM00758	Lung	5 Adherent
NCI.H2009	SIDM00756	Lung	5 Adherent
NCI.H2023	SIDM00753	Lung	4 Adherent
NCI.H2029	SIDM00735	Lung	1 Adherent
NCI.H2030	SIDM00715	Lung	4 Adherent
NCI.H2052	SIDM00713	Lung	3 Adherent
NCI.H2066	SIDM00711	Lung	1 Adherent
NCI.H2073	SIDM00757	Lung	5 Adherent
NCI.H2081	SIDM00710	Lung	1 Suspension
NCI.H2085	SIDM00709	Lung	5 Adherent
NCI.H2087	SIDM00708	Lung	5 Adherent
NCI.H209	SIDM00706	Lung	1 Suspension
NCI.H211	SIDM00704	Lung	1 Suspension
NCI.H2110	SIDM00703	Lung	5 Adherent

NCI.H2122	SIDM00702	Lung	5 Adherent
NCI.H2126	SIDM00775	Lung	5 Adherent
NCI.H2135	SIDM00700	Lung	4 Adherent
NCI.H2141	SIDM00699	Lung	1 Suspension
NCI.H2170	SIDM00716	Lung	5 Adherent
NCI.H2171	SIDM00733	Lung	1 Suspension
NCI.H2172	SIDM00926	Lung	4 Adherent
NCI.H2196	SIDM00731	Lung	1 Adherent
NCI.H2227	SIDM00730	Lung	1 Semi-Adherent
NCI.H2228	SIDM00729	Lung	4 Adherent
NCI.H226	SIDM00139	Lung	3 Adherent
NCI.H2286	SIDM00017	Lung	4 Unknown
NCI.H2291	SIDM00728	Lung	5 Adherent
NCI.H23	SIDM00138	Lung	4 Adherent
NCI.H2342	SIDM00727	Lung	5 Adherent
NCI.H2347	SIDM00726	Lung	5 Adherent
NCI.H2369	SIDM00104	Lung	3 Unknown
NCI.H2373	SIDM00103	Lung	3 Unknown
NCI.H2405	SIDM00724	Lung	4 Semi-Adherent
NCI.H2444	SIDM00723	Lung	5 Adherent
NCI.H2452	SIDM00722	Lung	3 Adherent

NCI.H2461	SIDM00102	Lung	3 Unknown
NCI.H250	SIDM00721	Lung	1 Suspension
NCI.H2591	SIDM00101	Lung	3 Unknown
NCI.H2595	SIDM00100	Lung	3 Unknown
NCI.H2722	SIDM00099	Lung	3 Unknown
NCI.H2731	SIDM00098	Lung	3 Unknown
NCI.H2795	SIDM00154	Lung	5 Unknown
NCI.H28	SIDM00720	Lung	4 Adherent
NCI.H2803	SIDM00309	Lung	3 Unknown
NCI.H2804	SIDM00310	Lung	3 Unknown
NCI.H2810	SIDM00311	Lung	3 Unknown
NCI.H2818	SIDM00520	Lung	3 Unknown
NCI.H2869	SIDM00519	Lung	3 Unknown
NCI.H290	SIDM00518	Lung	3 Unknown
NCI.H292	SIDM00493	Lung	5 Adherent
NCI.H3118	SIDM00517	Head and Neck	5 Unknown
NCI.H3122	SIDM00137	Lung	5 Unknown
NCI.H322M	SIDM00117	Lung	5 Unknown
NCI.H345	SIDM00719	Lung	1 Semi-Adherent
NCI.H358	SIDM00718	Lung	5 Adherent

NCI.H378	SIDM00915	Lung	1 Suspension
NCI.H441	SIDM00925	Lung	5 Adherent
NCI.H446	SIDM00965	Lung	1 Semi-Adherent
NCI.H460	SIDM00144	Lung	4 Adherent
NCI.H508	SIDM00777	Large Intestin	5 Suspension
NCI.H510A	SIDM00927	Lung	1 Semi-Adherent
NCI.H513	SIDM00114	Lung	5 Unknown
NCI.H520	SIDM01130	Lung	1 Adherent
NCI.H522	SIDM00116	Lung	1 Adherent
NCI.H524	SIDM01129	Lung	1 Suspension
NCI.H526	SIDM01128	Lung	1 Suspension
NCI.H596	SIDM01127	Lung	5 Adherent
NCI.H64	SIDM01126	Lung	1 Suspension
NCI.H647	SIDM01125	Lung	5 Adherent
NCI.H650	SIDM01124	Lung	4 Suspension
NCI.H660	SIDM01123	Prostate	1 Suspension
NCI.H661	SIDM01122	Lung	4 Adherent
NCI.H69	SIDM01121	Lung	1 Suspension
NCI.H716	SIDM00779	Large Intestin	1 Suspension
NCI.H720	SIDM01120	Lung	1 Suspension
NCI.H727	SIDM01119	Lung	5 Adherent

NCI.H740	SIDM01118	Lung	1 Suspension
NCI.H747	SIDM00778	Large Intestin	5 Semi-Adherent
NCI.H748	SIDM01117	Lung	1 Suspension
NCI.H810	SIDM01116	Lung	1 Adherent
NCI.H82	SIDM01131	Lung	1 Suspension
NCI.H820	SIDM00019	Lung	5 Unknown
NCI.H835	SIDM01133	Lung	1 Suspension
NCI.H838	SIDM01150	Lung	4 Adherent
NCI.H841	SIDM01134	Lung	1 Semi-Adherent
NCI.H847	SIDM01149	Lung	1 Suspension
NCI.H929	SIDM01148	Haematopoie	2 Suspension
NCI.N87	SIDM01147	Stomach	5 Adherent
NEC8	SIDM00243	Testis	1 Unknown
NH.12	SIDM00577	Peripheral Ne	1 Unknown
NKM.1	SIDM00576	Haematopoie	2 Unknown
NMC.G1	SIDM00575	Central Nervc	3 Unknown
no.10	SIDM00574	Central Nervc	3 Adherent
no.11	SIDM00573	Central Nervc	3 Unknown
NOMO.1	SIDM00580	Haematopoie	2 Unknown
NOS.1	SIDM00242	Bone	3 Adherent

NTERA.2.cl.D	SIDM01203	Testis	1 Adherent
NU.DUL.1	SIDM00443	Haematopoie	2 Suspension
NUGC.3	SIDM00572	Stomach	5 Unknown
NUGC.4	SIDM00570	Stomach	5 Unknown
NY	SIDM00569	Bone	3 Unknown
OACM5.1	SIDM00444	Esophagus	4 Semi-Adherent
OACp4C	SIDM00445	Esophagus	5 Adherent
OAW.28	SIDM00481	Ovary	5 Adherent
OAW.42	SIDM00480	Ovary	4 Adherent
OC.314	SIDM00220	Ovary	4 Adherent
OCI.AML2	SIDM00446	Haematopoie	2 Suspension
OCI.AML3	SIDM00462	Haematopoie	2 Suspension
OCI.AML5	SIDM00461	Haematopoie	2 Suspension
OCI.LY.19	SIDM00460	Haematopoie	2 Suspension
OCI.LY7	SIDM00459	Haematopoie	2 Suspension
OCI.M1	SIDM00458	Haematopoie	2 Suspension
OCUB.M	SIDM00241	Breast	5 Adherent
OCUM.1	SIDM00568	Stomach	5 Unknown
OE19	SIDM00479	Esophagus	5 Adherent
OE21	SIDM00478	Esophagus	5 Adherent
OE33	SIDM00477	Esophagus	5 Adherent

OMC.1	SIDM00240	Cervix	5 Adherent
ONCO.DG.1	SIDM00093	Ovary	5 Unknown
ONS.76	SIDM00567	Central Nervc	3 Adherent
OPM.2	SIDM00457	Haematopoie	2 Suspension
OS.RC.2	SIDM00239	Kidney	3 Adherent
OSC.19	SIDM00566	Head and Nei	5 Unknown
OSC.20	SIDM00565	Head and Nei	5 Unknown
OUMS.23	SIDM00036	Large Intestin	5 Unknown
OV.17R	SIDM00476	Ovary	4 Adherent
OV.56	SIDM00475	Ovary	4 Adherent
OV.7	SIDM00474	Ovary	3 Adherent
OV.90	SIDM01141	Ovary	5 Adherent
OVCA420	SIDM00967	Ovary	5 Adherent
OVCA433	SIDM00966	Ovary	3 Adherent
OVCAR.3	SIDM00105	Ovary	5 Adherent
OVCAR.4	SIDM00092	Ovary	5 Unknown
OVCAR.5	SIDM00091	Ovary	5 Unknown
OVCAR.8	SIDM00090	Ovary	4 Unknown
OVISE	SIDM00564	Ovary	5 Adherent
OVK.18	SIDM00238	Ovary	1 Adherent

OVKATE	SIDM00466	Ovary	5 Unknown
OVMIU	SIDM00465	Ovary	5 Unknown
OVTOKO	SIDM00425	Ovary	4 Unknown
P12.ICHIKAW	SIDM00463	Haematopoie	2 Suspension
P30.OHK	SIDM00364	Haematopoie	2 Unknown
P31.FUJ	SIDM00363	Haematopoie	2 Unknown
P32.ISH	SIDM00362	Haematopoie	2 Unknown
P3HR.1	SIDM01254	Haematopoie	2 Suspension
PA.1	SIDM01140	Ovary	4 Adherent
PA.TU.8902	SIDM00455	Pancreas	5 Adherent
PA.TU.8988S	SIDM00452	Pancreas	5 Adherent
PA.TU.8988T	SIDM00453	Pancreas	4 Adherent
PANC.02.03	SIDM01139	Pancreas	5 Adherent
PANC.03.27	SIDM01138	Pancreas	5 Adherent
PANC.04.03	SIDM01137	Pancreas	5 Adherent
PANC.08.13	SIDM01136	Pancreas	5 Adherent
PANC.1	SIDM00610	Pancreas	4 Adherent
PANC.10.05	SIDM01135	Pancreas	5 Adherent
PC.14	SIDM00237	Lung	5 Adherent
PC.3	SIDM00088	Prostate	5 Adherent
PCI.15A	SIDM00270	Head and Nei	5 Unknown

PCI.30	SIDM00044	Head and Neck	3 Unknown
PCI.38	SIDM00155	Head and Neck	5 Unknown
PCI.4B	SIDM00043	Head and Neck	5 Unknown
PCI.6A	SIDM00115	Head and Neck	5 Unknown
PE.CA.PJ15	SIDM00473	Head and Neck	5 Adherent
PEER	SIDM00991	Haematopoietic	2 Suspension
PEO1	SIDM00472	Ovary	5 Adherent
PF.382	SIDM00451	Haematopoietic	2 Suspension
PFSK.1	SIDM01132	Central Nervous	4 Adherent
PL.21	SIDM00450	Haematopoietic	2 Suspension
PL18	SIDM00054	Pancreas	5 Unknown
PL4	SIDM00042	Pancreas	5 Unknown
PSN1	SIDM00469	Pancreas	4 Adherent
PWR.1E	SIDM01114	Prostate	3 Adherent
QGP.1	SIDM00360	Pancreas	1 Unknown
QIMR.WIL	SIDM00467	Haematopoietic	2 Suspension
Raji	SIDM00844	Haematopoietic	2 Suspension
Ramos.2G6.4	SIDM01247	Haematopoietic	2 Suspension
RC.K8	SIDM00448	Haematopoietic	2 Suspension
RCC.AB	SIDM00821	Kidney	4 Adherent

RCC.ER	SIDM00820	Kidney	4 Adherent
RCC.FG2	SIDM00819	Kidney	4 Adherent
RCC.JF	SIDM00818	Kidney	4 Adherent
RCC.JW	SIDM00817	Kidney	4 Adherent
RCC.MF	SIDM00816	Kidney	4 Adherent
RCC10RGB	SIDM00235	Kidney	4 Adherent
RCH.AC.V	SIDM00449	Haematopoie	2 Suspension
RCM.1	SIDM00359	Large Intestin	5 Unknown
RD	SIDM00847	Soft Tissue	3 Adherent
REH	SIDM00447	Haematopoie	2 Suspension
RERF.GC.1B	SIDM00358	Stomach	5 Unknown
RERF.LC.AI	SIDM00307	Lung	4 Adherent
RERF.LC.KJ	SIDM00356	Lung	5 Unknown
RERF.LC.MS	SIDM00355	Lung	3 Unknown
RERF.LC.Sq1	SIDM00354	Lung	3 Unknown
RF.48	SIDM01092	Haematopoie	2 Unknown
RH.1	SIDM00427	Bone	1 Adherent
RH.18	SIDM00454	Soft Tissue	3 Adherent
RH.41	SIDM00426	Soft Tissue	1 Adherent
RKN	SIDM00353	Soft Tissue	3 Unknown
RKO	SIDM01090	Large Intestin	4 Adherent

RL	SIDM01089	Haematopoie	2 Suspension
RL95.2	SIDM01088	Endometriurr	5 Adherent
RMG.I	SIDM00352	Ovary	5 Unknown
RO82.W.1	SIDM00218	Thyroid	3 Adherent
ROS.50	SIDM00415	Haematopoie	2 Suspension
RPMI.2650	SIDM00809	Head and Nei	1 Adherent
RPMI.6666	SIDM00813	Haematopoie	2 Suspension
RPMI.7951	SIDM01087	Skin	3 Adherent
RPMI.8226	SIDM00087	Haematopoie	2 Suspension
RPMI.8402	SIDM00403	Haematopoie	2 Suspension
RPMI.8866	SIDM00482	Haematopoie	2 Suspension
RS4.11	SIDM01086	Haematopoie	2 Suspension
RT.112	SIDM00402	Bladder	5 Unknown
RT4	SIDM01085	Bladder	5 Adherent
RVH.421	SIDM00401	Skin	4 Adherent
RXF393	SIDM00086	Kidney	3 Unknown
Saos.2	SIDM01084	Bone	3 Adherent
Sarc9371	SIDM00033	Bone	3 Unknown
SAS	SIDM00351	Head and Nei	5 Unknown
SAT	SIDM00350	Head and Nei	5 Unknown

SBC.1	SIDM00365	Lung	1 Unknown
SBC.3	SIDM00367	Lung	1 Unknown
SBC.5	SIDM00368	Lung	1 Unknown
SC.1	SIDM00400	Haematopoie	2 Suspension
SCaBER	SIDM00032	Bladder	5 Adherent
SCC.15	SIDM01083	Head and Nei	5 Adherent
SCC.25	SIDM01082	Head and Nei	5 Adherent
SCC.3	SIDM00384	Haematopoie	4 Unknown
SCC.4	SIDM01081	Head and Nei	5 Adherent
SCC.9	SIDM01080	Head and Nei	5 Adherent
SCC90	SIDM00399	Head and Nei	5 Adherent
SCH	SIDM00383	Stomach	5 Unknown
SCLC.21H	SIDM00029	Lung	1 Unknown
Set2	SIDM00397	Haematopoie	2 Suspension
SF126	SIDM00382	Central Nervc	3 Unknown
SF268	SIDM00085	Central Nervc	3 Unknown
SF295	SIDM00084	Central Nervc	3 Unknown
SF539	SIDM00083	Central Nervc	3 Unknown
SH.4	SIDM01079	Skin	4 Adherent
SHP.77	SIDM01078	Lung	1 Semi-Adherent
SIG.M5	SIDM00396	Haematopoie	2 Suspension

SiHa	SIDM01093	Cervix	4 Adherent
SIMA	SIDM00395	Peripheral Ne	1 Adherent
SISO	SIDM00394	Cervix	5 Adherent
SJRH30	SIDM01095	Soft Tissue	1 Adherent
SJSA.1	SIDM01112	Bone	3 Adherent
SK.BR.3	SIDM00897	Breast	5 Adherent
SK.CO.1	SIDM01096	Large Intestin	5 Adherent
SK.ES.1	SIDM01111	Bone	1 Unknown
SK.GT.2	SIDM00393	Stomach	5 Adherent
SK.GT.4	SIDM00483	Esophagus	5 Adherent
SK.HEP.1	SIDM01110	Liver	4 Adherent
SK.LMS.1	SIDM01109	Soft Tissue	3 Adherent
SK.LU.1	SIDM01108	Lung	4 Adherent
SK.MEL.1	SIDM01107	Skin	4 Suspension
SK.MEL.2	SIDM00082	Skin	4 Adherent
SK.MEL.24	SIDM01106	Skin	4 Adherent
SK.MEL.28	SIDM00081	Skin	4 Adherent
SK.MEL.3	SIDM01105	Skin	4 Adherent
SK.MEL.30	SIDM00392	Skin	4 Adherent
SK.MEL.31	SIDM01104	Skin	3 Adherent

SK.MEL.5	SIDM00080	Skin	4 Adherent
SK.MES.1	SIDM01103	Lung	3 Adherent
SK.MG.1	SIDM00379	Central Nervc	3 Unknown
SK.MM.2	SIDM00391	Haematopoie	2 Suspension
SK.N.AS	SIDM01101	Peripheral Ne	3 Adherent
SK.N.BE.2	SIDM00894	Peripheral Ne	1 Adherent
SK.N.DZ	SIDM01100	Peripheral Ne	1 Adherent
SK.N.FI	SIDM01099	Peripheral Ne	1 Adherent
SK.N.MC	SIDM00634	Bone	1 Adherent
SK.N.SH	SIDM01098	Peripheral Ne	1 Adherent
SK.NEP.1	SIDM00027	Kidney	1 Unknown
SK.OV.3	SIDM00079	Ovary	4 Adherent
SK.PN.DW	SIDM01097	Central Nervc	1 Adherent
SK.UT.1	SIDM01113	Uterus	4 Adherent
SKG.IIIa	SIDM00381	Cervix	5 Unknown
SKM.1	SIDM00380	Haematopoie	2 Suspension
SKN	SIDM00377	Uterus	3 Unknown
SKN.3	SIDM00375	Head and Nei	5 Unknown
SLVL	SIDM00374	Haematopoie	2 Unknown
SN12C	SIDM00094	Kidney	3 Adherent
SNB75	SIDM00095	Central Nervc	3 Unknown

SNG.M	SIDM00373	Endometrium	5 Unknown
SNU.1	SIDM01146	Stomach	1 Suspension
SNU.1033	SIDM00192	Large Intestine	5 Adherent
SNU.1040	SIDM00217	Large Intestine	5 Adherent
SNU.1041	SIDM00173	Head and Neck	5 Adherent
SNU.1066	SIDM00170	Head and Neck	5 Adherent
SNU.1076	SIDM00168	Head and Neck	5 Adherent
SNU.1077	SIDM00166	Uterus	4 Adherent
SNU.1079	SIDM00164	Biliary Tract	4 Adherent
SNU.1196	SIDM00163	Biliary Tract	5 Adherent
SNU.1197	SIDM00171	Large Intestine	5 Adherent
SNU.1214	SIDM00167	Head and Neck	5 Adherent
SNU.16	SIDM01145	Stomach	5 Suspension
SNU.175	SIDM00216	Large Intestine	5 Suspension
SNU.182	SIDM01151	Liver	3 Adherent
SNU.201	SIDM00158	Central Nervous System	3 Unknown
SNU.245	SIDM00162	Biliary Tract	5 Adherent
SNU.283	SIDM00215	Large Intestine	5 Adherent
SNU.308	SIDM00161	Biliary Tract	5 Adherent
SNU.387	SIDM01180	Liver	3 Adherent

SNU.398	SIDM01181	Liver	1 Semi-Adherent
SNU.407	SIDM00214	Large Intestin	5 Adherent
SNU.423	SIDM01199	Liver	3 Adherent
SNU.449	SIDM01194	Liver	4 Adherent
SNU.475	SIDM01196	Liver	3 Adherent
SNU.478	SIDM00160	Biliary Tract	5 Adherent
SNU.5	SIDM01144	Stomach	5 Suspension
SNU.503	SIDM00497	Large Intestin	5 Adherent
SNU.61	SIDM00194	Large Intestin	5 Adherent
SNU.685	SIDM00165	Uterus	3 Adherent
SNU.81	SIDM00193	Large Intestin	5 Adherent
SNU.869	SIDM00159	Biliary Tract	5 Adherent
SNU.899	SIDM00169	Head and Nei	5 Adherent
SNU.C1	SIDM01197	Large Intestin	5 Suspension
SNU.C2A	SIDM00780	Large Intestin	5 Semi-Adherent
SNU.C2B	SIDM00781	Large Intestin	5 Semi-Adherent
SNU.C4	SIDM00500	Large Intestin	5 Adherent
SNU.C5	SIDM00498	Large Intestin	5 Adherent
SR	SIDM00096	Haematopoie	2 Suspension
ST486	SIDM01198	Haematopoie	2 Suspension
STS.0421	SIDM00041	Stomach	3 Unknown

SU.DHL.1	SIDM00390	Haematopoie	2 Suspension
SU.DHL.10	SIDM00389	Haematopoie	2 Suspension
SU.DHL.16	SIDM00404	Haematopoie	2 Suspension
SU.DHL.4	SIDM00405	Haematopoie	2 Suspension
SU.DHL.5	SIDM00406	Haematopoie	2 Suspension
SU.DHL.6	SIDM00407	Haematopoie	2 Suspension
SU.DHL.8	SIDM00423	Haematopoie	2 Suspension
SU8686	SIDM01188	Pancreas	5 Adherent
SUIT.2	SIDM00371	Pancreas	5 Unknown
SUP.B15	SIDM01176	Haematopoie	2 Suspension
SUP.B8	SIDM01177	Haematopoie	2 Unknown
SUP.HD1	SIDM00422	Haematopoie	2 Suspension
SUP.M2	SIDM00421	Haematopoie	2 Suspension
SUP.T1	SIDM01154	Haematopoie	2 Suspension
SW1088	SIDM01155	Central Nervc	3 Adherent
SW1116	SIDM00835	Large Intestin	5 Adherent
SW1271	SIDM01156	Lung	4 Adherent
SW13	SIDM00814	Adrenal Glan	1 Adherent
SW1417	SIDM00812	Large Intestin	5 Adherent
SW1463	SIDM00834	Large Intestin	5 Adherent

SW156	SIDM01162	Kidney	4 Adherent
SW1573	SIDM01163	Lung	4 Adherent
SW1710	SIDM00420	Bladder	3 Adherent
SW1783	SIDM01164	Central Nervc	3 Adherent
SW1990	SIDM01165	Pancreas	5 Adherent
SW403	SIDM00836	Large Intestin	5 Adherent
SW48	SIDM00837	Large Intestin	5 Adherent
SW480	SIDM00840	Large Intestin	5 Adherent
SW579	SIDM01167	Thyroid	3 Adherent
SW620	SIDM00841	Large Intestin	5 Adherent
SW626	SIDM01168	Large Intestin	5 Adherent
SW684	SIDM01175	Soft Tissue	3 Adherent
SW756	SIDM01174	Cervix	4 Adherent
SW780	SIDM01160	Bladder	5 Adherent
SW837	SIDM00833	Large Intestin	5 Adherent
SW872	SIDM01159	Soft Tissue	3 Adherent
SW900	SIDM01158	Lung	5 Adherent
SW948	SIDM00832	Large Intestin	5 Adherent
SW954	SIDM01157	Vulva	5 Adherent
SW962	SIDM01178	Vulva	3 Adherent
SW982	SIDM01161	Soft Tissue	3 Adherent

T.T	SIDM00322	Esophagus	5 Unknown
T24	SIDM01184	Bladder	4 Adherent
T47D	SIDM00097	Breast	5 Adherent
T84	SIDM00782	Large Intestin	5 Adherent
T98G	SIDM01171	Central Nervc	3 Adherent
TALL.1	SIDM00370	Haematopoie	2 Suspension
TASK1	SIDM00026	Central Nervc	1 Unknown
TC.71	SIDM00419	Bone	1 Semi-Adherent
TC.YIK	SIDM00234	Cervix	1 Suspension
TCCSUP	SIDM01190	Bladder	3 Adherent
TE.1	SIDM00369	Esophagus	5 Unknown
TE.10	SIDM00349	Esophagus	5 Unknown
TE.11	SIDM00348	Esophagus	5 Unknown
TE.12	SIDM00023	Esophagus	5 Unknown
TE.15	SIDM00249	Esophagus	5 Adherent
TE.4	SIDM00250	Esophagus	5 Adherent
TE.441.T	SIDM01173	Soft Tissue	1 Adherent
TE.5	SIDM00347	Esophagus	5 Unknown
TE.6	SIDM00328	Esophagus	5 Unknown
TE.8	SIDM00326	Esophagus	3 Unknown

TE.9	SIDM00325	Esophagus	5 Unknown
TGBC11TKB	SIDM00251	Stomach	5 Adherent
TGBC1TKB	SIDM00252	Biliary Tract	5 Adherent
TGBC24TKB	SIDM00267	Biliary Tract	5 Adherent
TGW	SIDM00324	Peripheral Ne	1 Unknown
THP.1	SIDM01172	Haematopoie	2 Suspension
TK	SIDM00323	Haematopoie	2 Unknown
TK10	SIDM00113	Kidney	4 Unknown
TMK.1	SIDM00040	Stomach	5 Unknown
TOV.112D	SIDM01170	Ovary	1 Adherent
TOV.21G	SIDM01169	Ovary	4 Adherent
TT	SIDM00484	Thyroid	1 Adherent
TT2609.C02	SIDM00418	Thyroid	3 Adherent
TUR	SIDM01205	Haematopoie	2 Suspension
TYK.nu	SIDM00321	Ovary	4 Unknown
U.118.MG	SIDM01193	Central Nervc	3 Adherent
U.2.OS	SIDM01191	Bone	3 Adherent
U.266	SIDM00417	Haematopoie	2 Suspension
U.698.M	SIDM00424	Haematopoie	2 Suspension
U.87.MG	SIDM01189	Central Nervc	3 Adherent
U.937	SIDM01195	Haematopoie	2 Suspension

U.CH2	SIDM01185	Bone	3 Adherent
U031	SIDM00112	Kidney	3 Unknown
U251	SIDM00111	Central Nervc	3 Adherent
UACC.257	SIDM00108	Skin	4 Unknown
UACC.62	SIDM00107	Skin	4 Unknown
UACC.812	SIDM01187	Breast	5 Adherent
UACC.893	SIDM01186	Breast	5 Adherent
UDSCC2	SIDM00025	Head and Nei	5 Unknown
UM.UC.3	SIDM01182	Bladder	4 Adherent
UMC.11	SIDM01183	Lung	1 Adherent
UWB1.289	SIDM00815	Ovary	5 Unknown
VA.ES.BJ	SIDM01179	Soft Tissue	4 Adherent
VAL	SIDM00416	Haematopoie	2 Suspension
VCaP	SIDM01077	Prostate	1 Adherent
VM.CUB.1	SIDM00414	Bladder	5 Adherent
VMRC.LCD	SIDM00320	Lung	1 Unknown
VMRC.MELG	SIDM00319	Skin	3 Adherent
VMRC.RCW	SIDM00318	Kidney	4 Unknown
VMRC.RCZ	SIDM00317	Kidney	4 Unknown
WIL2.NS	SIDM01102	Haematopoie	2 Suspension

WM.115	SIDM01000	Skin	3 Adherent
WM.266.4	SIDM00979	Skin	4 Adherent
WM1552C	SIDM00976	Skin	3 Adherent
WM278	SIDM00975	Skin	4 Adherent
WM35	SIDM00974	Skin	4 Adherent
WM793B	SIDM00973	Skin	3 Adherent
WSU.DLCL2	SIDM00413	Haematopoie	2 Suspension
WSU.NHL	SIDM00412	Haematopoie	2 Suspension
YAPC	SIDM00411	Pancreas	5 Adherent
YH.13	SIDM00316	Central Nervc	3 Unknown
YKG.1	SIDM00315	Central Nervc	3 Adherent
YT	SIDM00410	Haematopoie	2 Suspension
ZR.75.1	SIDM00314	Breast	5 Unknown
ZR.75.30	SIDM00971	Breast	5 Adherent

## Method

# Genome-scale identification of cellular pathways required for cell surface recognition

Sumana Sharma,<sup>1</sup> S. Josefin Bartholdson,<sup>1,3</sup> Amalie C.M. Couch,<sup>1</sup> Kosuke Yusa,<sup>2</sup> and Gavin J. Wright<sup>1</sup>

<sup>1</sup>Cell Surface Signalling Laboratory, Wellcome Trust Sanger Institute, Cambridge CB10 1SA, United Kingdom; <sup>2</sup>Stem Cell Genetics Laboratory, Wellcome Trust Sanger Institute, Cambridge CB10 1SA, United Kingdom

Interactions mediated by cell surface receptors initiate important instructive signaling cues but can be difficult to detect in biochemical assays because they are often highly transient and membrane-embedded receptors are difficult to solubilize in their native conformation. Here, we address these biochemical challenges by using a genome-scale, cell-based genetic screening approach using CRISPR gene knockout technology to identify cellular pathways required for specific cell surface recognition events. By using high-affinity monoclonal antibodies and low-affinity ligands, we determined the necessary screening parameters, including the importance of establishing binding contributions from the glycocalyx, that permitted the unequivocal identification of genes encoding directly interacting membrane-embedded receptors with high statistical confidence. Importantly, we show that this genome-wide screening approach additionally identified receptor-specific pathways that are required for functional display of receptors on the cell surface that included chaperones, enzymes that add post-translational modifications, trafficking proteins, and transcription factors. Finally, we demonstrate the utility of the approach by identifying IGF2R (insulin like growth factor 2 receptor) as a binding partner for the R2 subunit of GABA<sub>B</sub> receptors. We show that this interaction is direct and is critically dependent on mannose-6-phosphate, providing a mechanism for the internalization and regulation of GABA<sub>B</sub> receptor signaling. We conclude that this single approach can reveal both the molecular nature and the genetic pathways required for functional cell surface display of receptors recognized by antibodies, secreted proteins, and membrane-embedded ligands without the need to make any prior assumptions regarding their biochemical properties.

[Supplemental material is available for this article.]

Membrane-compartmentalized cells receive instructional information from their surroundings by extracellular signaling cues that are often initiated by specific binding events made by plasma membrane-embedded receptors. These extracellular interactions are crucial for the normal development and function of multicellular organisms and can be exploited therapeutically because they are directly accessible to soluble drugs such as monoclonal antibodies (mAbs) (Weiner 2015). Investigating extracellular cell signaling interactions mediated by membrane receptor proteins can be challenging because the proteins are amphipathic, making them difficult to solubilize in their native conformation and because the interactions are typified by weak interaction affinities; consequently, most commonly used methods are generally unsuitable to detect this class of protein interactions (Wright 2009). The biochemical features of low-affinity membrane receptor interactions have necessitated the development of bespoke techniques to detect them, and one approach involves expressing the entire ectodomain of a receptor as a soluble recombinant protein. The ectodomains are usually purposefully oligomerized so that they can be used as highly avid probes to identify binding partners by expression cloning or biochemical purifications (Wright et al. 2010). More recently, we and others have developed large-scale systematic methods to identify novel receptor–ligand interactions

by screening for direct interactions within large protein libraries containing hundreds of receptor ectodomains using ELISA (enzyme-linked immunosorbent assay)-style approaches (Bushell et al. 2008; Ozkan et al. 2013; Visser et al. 2015). While successful, this general approach has drawbacks that prevent its wider use by most laboratories because compiling protein libraries containing hundreds of proteins is resource intensive, and most researchers' interests are usually focused on a single or small number of proteins rather than the networks of interactions within receptor protein families. Importantly, this technique requires that the receptor binding function is retained when expressed by heterologous cells out of the context of the plasma membrane as a soluble recombinant protein. While this is generally the case for proteins that span the membrane once, this is more difficult for receptor complexes and membrane proteins that span the membrane multiple times, presenting additional challenges to characterize their interactions. Moreover, methods detecting binding events between recombinant proteins do not account for the complex environment in which receptor interactions would normally occur at the cell surface, which includes contributions from a charged glycocalyx of carbohydrates and lipids displayed on a dynamic membrane.

The recent development of cell-based genetic screening approaches using highly efficient CRISPR methods now presents the possibility to interrogate the genetic basis of cellular phenotypes on a genome-wide scale (Koike-Yusa et al. 2014; Shalem

<sup>3</sup>Present address: Department of Medicine, University of Cambridge, Addenbrooke's Hospital, Cambridge, CB2 0QQ, UK  
Corresponding author: gw2@sanger.ac.uk

Article published online before print. Article, supplemental material, and publication date are at <http://www.genome.org/cgi/doi/10.1101/gr.231183.117>. Freely available online through the *Genome Research* Open Access option.

© 2018 Sharma et al. This article, published in *Genome Research*, is available under a Creative Commons License (Attribution 4.0 International), as described at <http://creativecommons.org/licenses/by/4.0/>.

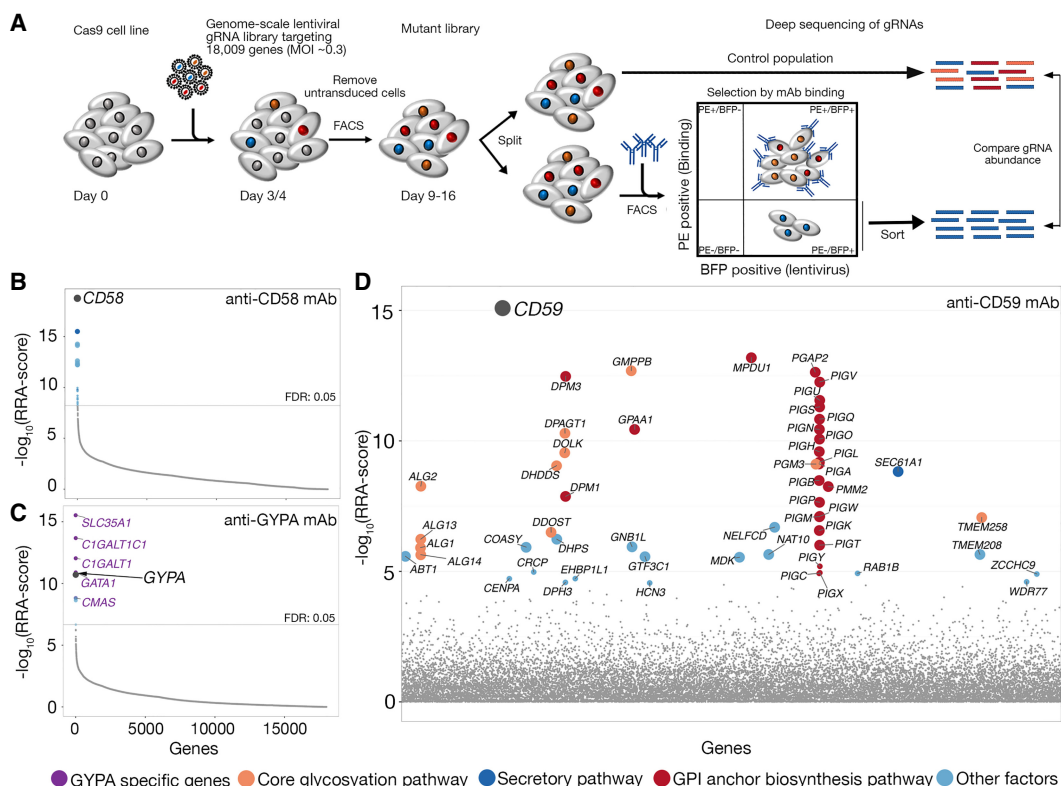
et al. 2014, 2015; Wang et al. 2014). Libraries of cells that contain biallelic targeted loss-of-function alleles can be created, and by selecting those cells with a phenotype of interest, the gene products involved can be identified (Ma et al. 2015; Parnas et al. 2015; Zhang et al. 2016). Here, we use genome-scale, cell-based CRISPR knockout (KO) screens to determine the molecular basis of cell surface recognition events made by mAbs, secreted proteins, and receptors. We show that this technique is able to not only identify genes encoding cell surface proteins that directly interact with these binding probes but also reveal receptor-specific pathways required for receptor display at the cell surface in a functional form, including enzymes required for essential post-translational modifications, chaperones, and trafficking proteins.

## Results

### Genetic determinants of mAb cell surface epitope display by genome-scale CRISPR screens

To determine if a genome-scale, cell-based CRISPR-KO approach could identify genes that are required for specific cell surface recognition events within the context of a plasma membrane, we first selected a panel of six mAbs that brightly stained five different cell surface receptors (Supplemental Table S1). To create a population of mutant cells, the Cas9-expressing cells were trans-

duced at a low multiplicity of infection (~0.3) with a library of lentiviruses, each encoding a single gRNA from a pool of 90,709 individual gRNAs targeting 18,009 human genes (Tzelepis et al. 2016). Transduced cells that had lost the antibody epitope at the cell surface were isolated by FACS, and the genes responsible for this loss of binding were identified by comparing the relative abundance of the different gene-specific gRNAs present in the sorted cells compared with the total unsorted population using deep sequencing of gRNA PCR products and enrichment analysis (Fig. 1A; Li et al. 2014). Initial experiments established that the day of selection, the number of sorted cells, and sorting thresholds influenced the success of the approach (see Methods), as well as selecting high-activity clonal Cas9-expressing versions for each cell line to homogenize and maximize the efficiency of genome editing (Supplemental Fig. S1). By using this optimized approach, gRNAs targeting the gene encoding the antibody receptor epitope were specifically enriched in the sorted cells for each of the six antibodies (false-discovery rate [FDR] <0.05) (see Supplemental Table S1; Supplemental Data S1). In the case of selections using anti-CD58, gRNAs targeting *CD58* were the most highly enriched (Fig. 1B). In cells selected with the anti-glycophorin A mAb, the most enriched genes targeted *GYP A* encoding the receptor, as well as genes required for sialylated O-linked glycan biosynthesis (*C1GALT1*, *C1GALT1C1*, *SLC35A1*, *CMAS*)—which are presumably critical for creating the extracellular antibody epitope—and

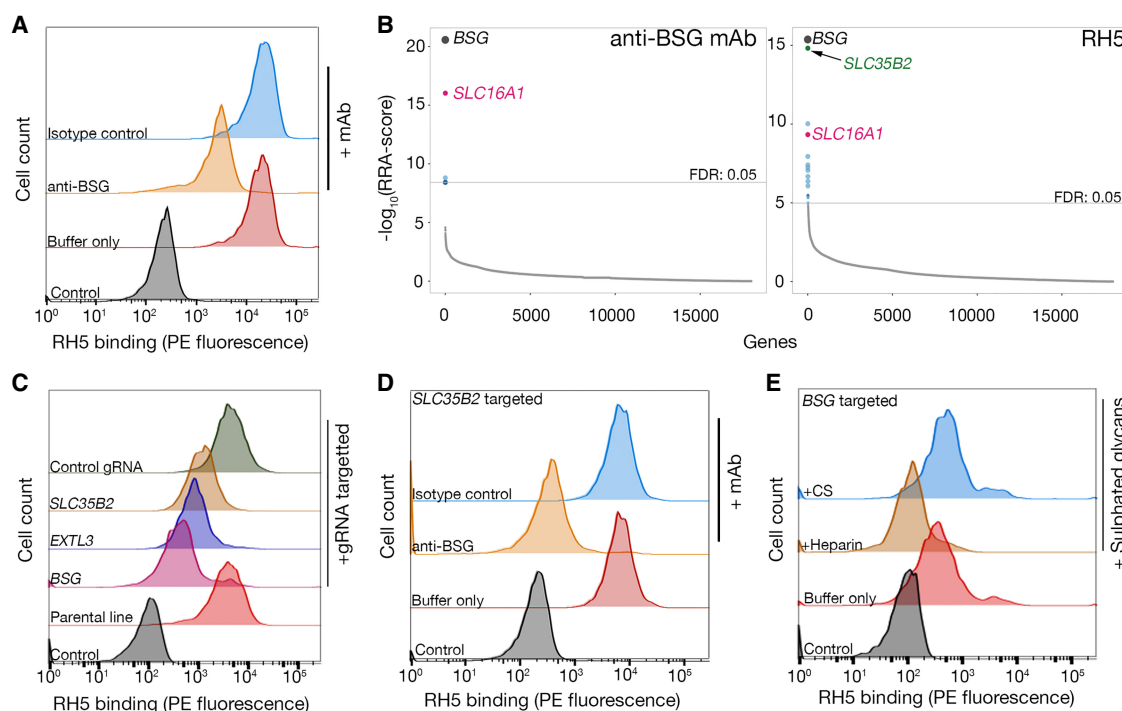


the erythroid-specific transcription factor GATA1, which is likely to be necessary for *GYP*A transcription in these cells (Fig. 1C). Similarly, a mAb recognizing the glycosylphosphatidylinositol (GPI)-anchored CD59 receptor identified not only *CD59* as the gene with the most enriched gRNAs but also 21 out of 27 enzymes known to be required for GPI-anchor biosynthesis (Fig. 1D). In addition, a mAb recognizing the  $\alpha 2\beta 1$  integrin identified *ITGB1* and components of the Arp2/3 complex, demonstrating the epitope of this mAb is located within the  $\beta 1$  and not  $\alpha 2$  chain, and the critical requirement of actin regulation, consistent with the known cell biology of integrin function (Supplemental Table S1; Brakebusch and Fassler 2003). A pathway analysis of all enriched genes that were shared between antibody selection encoded proteins required for protein secretion and glycosylation, as expected, but also identified housekeeping pathways such as ribosome biosynthesis and RNA metabolism (Supplemental Fig. S2). We observed that the representation of these pathways was often reduced when selections were performed several days later, suggesting these genes are required for long-term cell viability in culture and that antibody staining was reduced on moribund cells. An independent repeat of the selections using the anti-CD59 antibody again identified *CD59* as one of the most enriched genes together with genes involved in the GPI-anchor biosynthesis pathway, showing that the experimental parameters were tolerant of biological variation (Supplemental Fig. S3; Supplemental Data S1). Together, these data demonstrated that the genome-scale CRISPR-KO screening

approach could not only robustly identify the gene encoding the antibody epitope for all of the six antibodies tested but also reveal genes and pathways that are important in the cell biology for specific receptors.

### CRISPR-KO screening identifies receptors and a role for heparan sulfate in low-affinity interactions

While high-affinity mAbs are useful research tools and a few studies have previously shown the utility of CRISPR screens to identify receptor-related cellular pathways (Parnas et al. 2015; Zotova et al. 2016; Burr et al. 2017), we next sought to determine if this approach could be used to identify low-affinity receptors for cell signaling ligands. As a model system, we selected the interaction between *Plasmodium falciparum* RH5 and its host receptor basigin (BSG) because it is a low-affinity interaction ( $K_D \sim 1 \mu\text{M}$ ), is biochemically and structurally well characterized and because BSG was highly expressed on our Cas9-expressing HEK293 cell line (Crosnier et al. 2011; Wright et al. 2014). To detect low-affinity interactions, we increased binding avidity by clustering biotinylated RH5 around a fluorescent streptavidin conjugate, and this reagent bound to the surface of HEK293 cells as expected; however, precoating the cells with a blocking anti-BSG mAb did not prevent all RH5 binding, suggesting there was an additional receptor(s) for RH5 on HEK293 cells (Fig. 2A). This additional binding was not due to a subfraction of inactive protein in the RH5



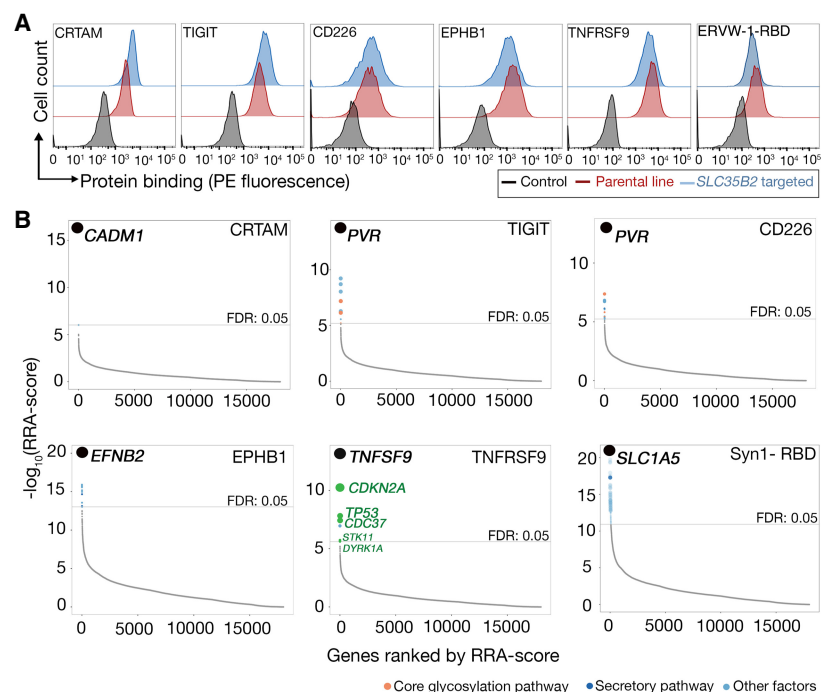
**Figure 2.** Identification of BSG and heparan sulfate as independent receptors for *P. falciparum* RH5 on HEK293 cells. (A) Biotinylated RH5 was clustered around a streptavidin–PE conjugate and binding to HEK293 cells was analyzed by flow cytometry. RH5 binding is only partially reduced by a blocking anti-BSG mAb relative to controls. (B) Rank-ordered genes identified from gRNA enrichment analysis required for cell surface display of an anti-BSG mAb (left) and RH5 binding (right). Significantly enriched genes with a FDR < 0.05 are colored (full screening results available in Supplemental Data S2); genes encoding the receptor (BSG) and chaperone (SLC16A1) were common to both screens, and a gene involved in GAG-biosynthesis (SLC35B2) was additionally required for RH5 binding. (C) Binding of RH5 to cells is reduced when transduced with lentiviruses encoding gRNAs targeting either the receptor (BSG) or enzymes required for HS synthesis (SLC35B2, EXTL3) relative to controls. Transduced polyclonal lines were used for this experiment. (D) RH5 binding to SLC35B2-targeted HEK293 cells could be completely prevented if preincubated with a blocking anti-BSG mAb but not an isotype-matched control. (E) RH5 binding to BSG-targeted HEK293 cells could be completely blocked if preincubated with 200  $\mu\text{g}/\text{mL}$  heparin but not 200  $\mu\text{g}/\text{mL}$  CS. A representative of three independent (A and C) or technical (D and E) replicate experiments is shown.

preparation since all binding could be prevented by heat treatment (Supplemental Fig. S4A). To identify the receptor(s) for RH5 other than BSG in this cellular context, we compared the genes required for RH5 binding versus those necessary for surface expression of BSG by using an anti-BSG mAb. The enriched gRNAs common to both selections beyond those targeting general secretory pathway genes corresponded to *BSG*, as expected, but also to a gene encoding a monocarboxylate transporter, *SLC16A1*, which is a known chaperone required for surface expression of BSG (Fig. 2B; Kirk et al. 2000). The most highly enriched gene in the cells sorted using RH5 compared with anti-BSG was *SLC35B2* (solute carrier family 35 member B2), which encodes a protein that transports 3'-phosphoadenosine-5'-phosphosulfate from the cytosol into the lumen of the Golgi apparatus, where sulfotransferases use it as a universal donor for the sulfation of major constituents of the cellular glycocalyx, including glycoproteins, glycolipids, and glycosaminoglycans (GAGs) (Fig. 2B; Kamiyama et al. 2003). An analysis of the enriched genes required for RH5 binding using KEGG (Kanehisa et al. 2017) identified the heparan sulfate (HS) biosynthesis pathway, and consistent with this, RH5 binding could be inhibited to a threshold value by heparin, but not the related GAG, chondroitin sulfate (CS) (Supplemental Fig. S4B,C). This is in agreement with the reported presence of heparin binding motifs in RH5 and its ability to bind heparin-coated agarose (Baum et al. 2009). HS is a component of the cellular glycocalyx surrounding most cells and is known to adsorb a wide range of extracellular proteins both as a coreceptor for signaling proteins and to interact with the extracellular matrix (Esko et al. 2009). To further investigate the role of *SLC35B2* and HS in RH5 binding, we first demonstrated that surface expression of BSG was not affected in cells where genes required for GAG biosynthesis were targeted (Supplemental Fig. S4D). We then showed that cells targeted for genes required for HS biosynthesis (*SLC35B2* and *EXTL3*) showed a partial reduction in RH5 binding (Fig. 2C) and that the residual binding was specifically due to BSG because it could be completely abrogated by preincubating the cells with the blocking anti-BSG antibody (Fig. 2D). This suggested that the contributions of BSG and HS to RH5 binding were independent, and this was further confirmed by showing that soluble heparin, but not CS, could block all RH5 binding to BSG-deficient cells (Fig. 2E). These experiments revealed a role for HS within the glycocalyx for interactions at the cell surface and demonstrated that this technique was able to identify the direct receptor of low-affinity ligands, including any required receptor cofactors such as chaperones.

### Genome-scale, cell-based CRISPR-KO screens identify directly interacting receptors

Many extracellular proteins are known to bind HS, and the finding that HS binding

in our assay could be additive rather than codependent on other receptors suggested that HS may represent a factor responsible for cell surface binding for a range of ligands even in the absence of another receptor. To examine and address this, we took advantage of our *SLC35B2*-targeted cell line to rapidly determine the contribution of cell staining due to extracellular sulfate adsorption by comparing ligand binding events between the parental and *SLC35B2*-targeted cells. By using this approach, we identified six ligands with known protein receptors, which, when presented as avid binding reagents, bound HEK293 cells and exhibited no loss of staining when *SLC35B2* had been targeted. These ligands were as follows: CRTAM, TIGIT, CD226, EPHB1, TNFRSF9, and N-terminal receptor binding domain (RBD) of ERVW-1 (also known as Syncytin 1) (Fig. 3A). When these protein probes were used in our CRISPR screening approach, the gene with the most enriched gRNAs corresponded to a known receptor in each and every case: *CADM1* was the top-ranked gene when selected with CRTAM, *PVR* for both TIGIT and CD226 ligands, *EFNB2* for EPHB1, *TNFSF9* for TNFRSF9, and *SLC1A5* for the RBD of ERVW-1 (ERVW-1-RBD) (Fig. 3B). For TNFRSF9, several genes involved in the p53 pathway (*CDKN2A*, *CDC37*, *STK11*, and *DYRK1A*) and *TP53* itself were also enriched in the nonbinding population, suggesting a role for the p53 pathway in presenting TNFSF9 in a functional ligand-binding form on the cell surface. We validated this by independently targeting *TP53*, which resulted in a decrease in the binding of the TNFRSF9 ligand (Supplemental Fig. S5A). An independent repeat of the selections using the TIGIT protein again



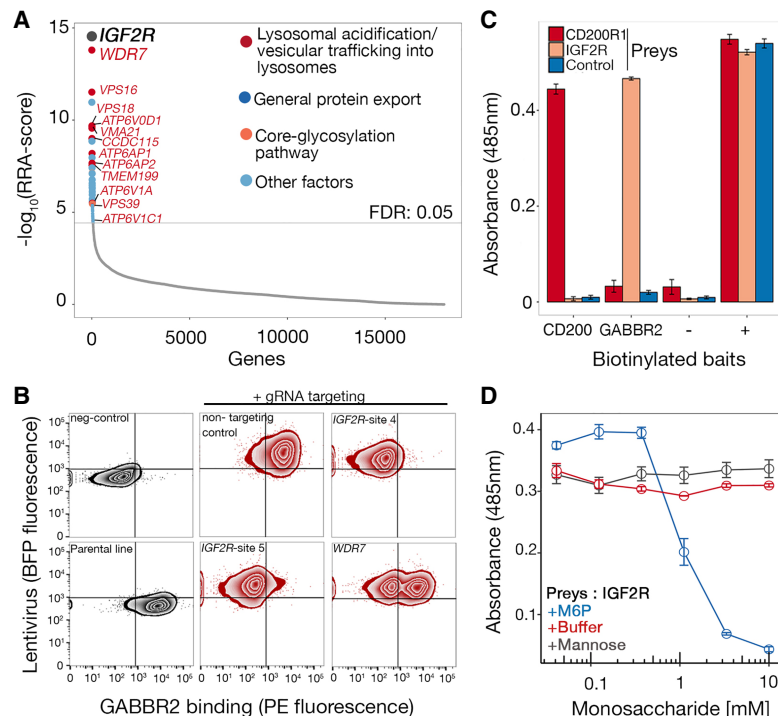
**Figure 3.** Identification of cell adhesion receptors and pathways using cell-based genetic screens. (A) The indicated oligomeric probes were tested for binding to the unmodified parental cell line (red histograms) or polyclonal *SLC35B2*-targeted cells (blue); representative experiments of three technical replicates are shown. The HEK293 cell line was used for all proteins except TNFRSF9, which is the NCI-SNU-1 line. (B) RRA-score rank-ordered genes identified from gRNA enrichment analysis of sorted cells that had lost binding to CRTAM, TIGIT, CD226, EPHB1, TNFRSF9, and ERVW-1 (ERVW-1-RBD); in all six cases, the gene encoding the known receptor was identified as the most significantly enriched gene. In B, genes with FDR < 5% are labeled with colors corresponding to related functions (full screening results available in Supplemental Data S3).

identified the gene encoding its receptor, PVR, as the most highly enriched gene, demonstrating that the experimental parameters were tolerant of biological variation (Supplemental Fig. S5B). These experiments demonstrate that the directly interacting receptor could be unambiguously identified with high statistical confidence in every case, and for some proteins, the cellular pathways responsible for the cell biology of the receptor can be determined.

### IGF2R interacts with GABBR2 in a M6P-dependent manner

The results of these experiments suggested that the cell-based, genome-scale CRISPR-KO approach could be used to identify not only directly interacting cell surface receptors but also the cellular pathways required for functional receptor presentation at the cell surface. To demonstrate this, we identified the gamma-aminobutyric acid (GABA) type B receptor subunit 2 (GABBR2) as a receptor whose ectodomain bound HEK293 cells in an *SLC35B2*-independent manner. GABBR2 has an important but mechanistically poorly understood role in the internalization of inhibitory GABA<sub>B</sub> receptor complexes to regulate neurotransmission (Hannan et al. 2011; Benke 2013). Sorted mutant cells that had lost the ability to bind the GABBR2 ectodomain were enriched in gRNAs targeting the gene *IGF2R* (insulin like growth factor 2 receptor, also known as the cation-independent mannose-6-phosphate (M6P) receptor and CD222), as well as several genes encoding proteins required for endosomal trafficking and function (Fig. 4A). These genes included subunits of V-type ATPases (*ATP6V0D1*, *ATP6AP1*, *ATP6AP2*, *ATP6V1A*, *ATP6V1C1*), a chaperone required for V-ATPase assembly (*VMA21*), V-ATPase assembly factors (*TMEM199*, *CCDC115*), an accessory factor involved in endosomal acidification (*WDR7*), and proteins required for membrane trafficking (*VPS16*, *VPS18*) and fusion (*VPS39*). IGF2R is a known cargo receptor that transports M6P-modified proteins between the trans-Golgi network, endosomes, and plasma membrane and might therefore provide a mechanism for the known internalization and degradation of GABA<sub>B</sub> receptors through interactions with the GABBR2 subunit (Grampp et al. 2007). To investigate this further, we first showed using immunocytochemistry that both proteins were localized within the membranous compartments of cells with IGF2R localizing as discrete puncta in internal membranes that also contained GABA<sub>B</sub> receptors (Supplemental Fig. S6A). To further investigate the interaction between these two proteins, we first validated our results using individual gRNAs targeting *IGF2R*, which resulted in the loss of both IGF2R surface staining (Supplemental Fig. S6B) and binding of the GABBR2 ectodomain (Fig. 4B). It is known that cells treated with com-

pounds that increase lysosomal pH cause IGF2R to accumulate in endosomes with a consequent loss from the cell surface, providing an explanation for why genes required for endosomal function and trafficking were also enriched (Reaves and Banting 1994). Consistent with this, individual gRNAs targeting the *WDR7* gene, which is required for lysosomal acidification (Merkulova et al. 2015), resulted in the reduction of both IGF2R mAb staining (Supplemental Fig. S6B) and GABBR2 ectodomain binding (Fig. 4B). To demonstrate that the loss of GABBR2 binding to cells was due to the direct interaction with IGF2R, we first showed that GABBR2 ectodomain binding could be conferred to cells by transfecting them with an expression plasmid encoding an IGF2R-GFP fusion protein (Supplemental Fig. S6C) and then observed that the entire ectodomains of IGF2R expressed as a soluble beta-lactamase-tagged “prey” could be specifically captured by a biotinylated GABBR2 ectodomain “bait” (Fig. 4C). We further demonstrated that the interaction between IGF2R and GABBR2 was dependent on M6P-modified N-linked glycans by showing the interaction was abolished by either treating the GABBR2 ectodomain with PNGase F (Supplemental Fig. S6D) or adding soluble M6P (Fig. 4D). These results demonstrate that the



**Figure 4.** IGF2R interacts with GABBR2 in a mannose-6-phosphate (M6P)-dependent manner. (A) RRA-score rank-ordered genes identified from gRNA enrichment analysis from sorted mutant cells that had lost GABBR2 binding activity. Enriched genes encoded the IGF2R receptor and proteins involved in lysosome biology (also see Supplemental Data S3). (B) Binding of GABBR2 was quantified on HEK293 cells transduced with two gRNAs targeting different exons of *IGF2R* and one gRNA targeting *WDR7*. A near complete loss of binding was observed on *IGF2R*-targeted cells and a partial loss on *WDR7*-targeted cells; targeted cells were maintained as polyclonal lines. A representative experiment of three technical replicates is shown. (C) Direct binding between IGF2R and GABBR2 ectodomains. The biotinylated ectodomain of GABBR2 was immobilized as a “bait” on streptavidin-coated microtiter plates and tested for direct interactions using a beta-lactamase-tagged “prey” ectodomain of IGF2R. Binding was quantified using the beta-lactamase substrate nitrocefin, whose hydrolysis products absorb at 485 nm. Positive control was the CD200–CD200R1 interaction; control “prey” is an unrelated ectodomain, positive (+) represents total capture of all preys with an anti-prey antibody, and negative (–) represents a tag-only bait control. (D) The interaction between IGF2R and GABBR2 can be completely inhibited by 10 mM soluble M6P. The binding dependency on M6P was tested by adding serial dilutions of mannose, M6P, or buffer alone. Data points in C and D are mean  $\pm$  SEM,  $n = 3$ .

ectodomain of GABBR2 interacts directly with IGF2R in a M6P-dependent manner.

### IGF2R is a trafficking receptor for GABBR2

Based on what is already known about the GABA<sub>B</sub> receptor complex and IGF2R, we hypothesized that this interaction provided a possible molecular mechanism to explain how GABA<sub>B</sub> receptors are internalized from the cell surface to regulate their activity. To test this, we first targeted *IGF2R* in Cas9-expressing HEK293 cells and selected an IGF2R-deficient clonal line (Supplemental Fig. S7A). To determine how the cell biology of the GABA<sub>B</sub> receptor complex was altered in the absence of IGF2R, we transfected the cells with cDNAs encoding both subunits of the GABA<sub>B</sub> receptor and quantified the amount of GABA<sub>B</sub> receptor complex on the surface of unpermeabilized cells (Fig. 5A). We observed that the average amount of cell surface GABA<sub>B</sub> receptor complex was significantly higher in the *IGF2R*-KO cell line (Fig. 5B; Supplemental Fig. S7B,C), which is consistent with the model that GABA<sub>B</sub> receptor complex internalization is promoted by IGF2R. It has also been previously shown that GABA<sub>B</sub> receptor internalization depends on clathrin-mediated internalization as treatment of cells with hypertonic concentrations of sucrose or chlorpromazine, which have been shown to inhibit the formation of clathrin-coated pits, leads to a complete block of internalization of GABA<sub>B</sub> receptors from the surface of HEK293 cells (Grampp et al. 2007). To assess the influence of the clathrin-dependent pathway on the internalization of the GABA<sub>B</sub> receptors, we used our cell-based functional IGF2R–GABA<sub>B</sub> complex interaction assay. We observed that upon treatment of cells with sucrose, the average amount of cell surface GABA<sub>B</sub> receptor complex increased on a population of parental HEK293 cells relative to a control demonstrating that GABA<sub>B</sub> receptor trafficking depends on the clathrin pathway

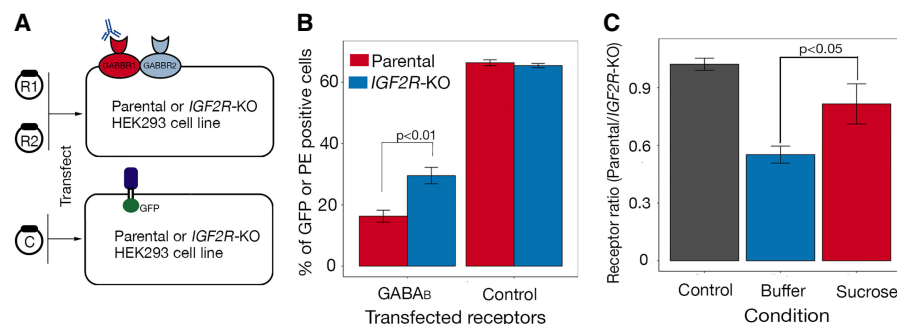
(Fig. 5C). Additionally, to investigate the effect of the agonist GABA on the interaction of the IGF2R–GABA<sub>B</sub> complex, we added GABA to both the parental and IGF2R-deficient cells expressing the GABA<sub>B</sub> receptor complex and quantified the IGF2R-dependent internalization of the complex using our cell-based assay. We observed that in the presence of GABA, there was no significant difference in the ratio of cell surface GABA<sub>B</sub> receptor complex between the parental and IGF2R-deficient cells, demonstrating that GABA agonists have no effect on the IGF2R–GABA<sub>B</sub> receptor complex interaction (Supplemental Fig. S8). These data demonstrate that the IGF2R-mediated internalization of the GABA<sub>B</sub> receptor complex is clathrin dependent, and provide evidence that IGF2R is involved in the internalization of GABA<sub>B</sub> receptors through direct interactions with the GABBR2 receptor subunit.

### Discussion

Here we have shown how it is possible to use a cell-based, genome-scale CRISPR-KO approach to identify genes encoding proteins required for cellular recognition. This genetic approach provides a valuable alternative to existing biochemical methods that must account for the largely insoluble nature of membrane-embedded receptors and the often highly transient nature of their extracellular interactions. This technique has advantages in that it not only is able to reveal the receptor protein that directly interacts with a presented ligand but can also identify other cell-intrinsic gene products that are required for presenting the receptor at the cell surface in an active conformation. In the experiments reported here, this included essential chaperones, actin-interacting proteins, vesicle trafficking adapters, enzymes required for specific glycan biosynthesis, and transcription factors. Genes that were repeatedly identified in different screens had known functional roles in general “house-keeping” receptor biology, including those

involved in the secretory and general glycosylation pathways. One important advantage of this approach is that it should be able to identify all gene products required for receptor interactions without the need to make any prior assumptions regarding the cell biology or biochemical nature of the receptor. The genetic nature of the approach may be especially useful to identify interactions involving membrane-embedded proteins that are biochemically very challenging to work with; for example, we show here how this approach can confidently identify the receptor for ERVW-1-RBD, a protein that spans the membrane 12 times and is therefore difficult to solubilize in an active conformation. We anticipate that this single method will be useful to explore the contribution of other molecules such as lipids and may reveal novel pathways for receptor trafficking and biology.

We found that we could additionally use this approach to investigate the contribution to extracellular interactions made by components of the cellular glycocalyx such as GAGs. Many secreted proteins are known to interact with



**Figure 5.** IGF2R traffics GABBR2 through a clathrin-dependent pathway. (A) Schematic of a flow cytometry-based GABA<sub>B</sub> cell surface expression assay. Parental and *IGF2R*-KO HEK293 cells were cotransfected with plasmids encoding both *GABBR1* (R1) and *GABBR2* (R2) full-length cDNAs and surface expression of GABA<sub>B</sub> receptors quantified by antibody staining and FACS. A plasmid containing a GFP-tagged *CD200* cDNA was transfected in parallel and the number of transfected cells quantified by GFP fluorescence. (B) The percentage of GABA<sub>B</sub> receptor-positive cells was significantly higher on the surface of *IGF2R*-KO cells compared with the parental line ( $P < 0.01$  using Welch's two sample *t*-test,  $n = 3$ ); control demonstrates both cell lines were transfected at comparable levels. Error bars,  $\pm$  SD,  $n = 3$  (cells were transfected in triplicate); a representative experiment of two independent experiments is shown. (C) The clathrin-pathway inhibitor sucrose was added to parental and *IGF2R*-deficient HEK293 cells expressing the GABA<sub>B</sub> receptor complex, and the cell surface receptor levels were quantified. The effect of sucrose on the IGF2R–GABA<sub>B</sub> receptor complex interaction was determined by calculating the ratio of positively stained cells on the parental cell line relative to the *IGF2R*-deficient cells and a buffer-only control. Sucrose treatment led to a significant increase ( $P < 0.05$  using Welch's two sample *t*-test) in cell surface GABA<sub>B</sub> receptor complex staining on the parental cell line, almost to the same level as on the *IGF2R*-deficient cells, demonstrating that GABA<sub>B</sub> receptor trafficking depends on the clathrin pathway. The control represents a transfection control to demonstrate parental and *IGF2R*-deficient cells were transfected with equal efficiency. Bars, mean  $\pm$  SD;  $n = 3$  (cells were transfected in triplicate).

GAGs, especially HS, and similar cell-based genetic screens to identify pathogen receptors have shown that they also play important roles in the interactions of bacteria (Rosmarin et al. 2012) and viruses (Marceau et al. 2016; Pillay et al. 2016) with host cells. The interactions are largely electrostatic with the brush-like negatively charged surface HS forming salt bridges with surface-exposed basic residues and are generally thought to provide a suitable scaffold to present ligands to host receptors in an appropriate manner by regulating their orientation and oligomerization and by establishing local concentration gradients (Yan and Lin 2009). The *P. falciparum* RH5 protein is a typical example of a secreted protein that is known to interact with both heparin-like molecules (Baum et al. 2009) and the BSG receptor (Crosnier et al. 2011), and we were able to clearly identify both in a single experiment using our approach. Further investigation showed that the contributions to RH5 binding by BSG and HS, at least in this context, were independent and could be experimentally separated. This finding enabled us to rapidly establish whether sulfated GAGs played a role in the binding of other ligands by testing binding on an *SLC35B2*-deficient cell line or preincubating the ligand with heparin.

We demonstrated the utility of this technique by identifying IGF2R as a binding partner for GABA<sub>B</sub> receptors. GABA<sub>B</sub> receptors are expressed abundantly in almost all types of neurons and glia throughout the central nervous system and mediate slow-acting control of neuron excitability by inhibiting neurotransmitter release. This expression pattern overlaps with that of IGF2R, which is also widely distributed throughout the nervous system with particular enrichment in cortical areas, the hippocampus, and cerebellum (Hawkes and Kar 2004). The regulation of the surface level of GABA<sub>B</sub> receptors by endocytosis is an important mechanism to attenuate signal strength and can be modelled in HEK293 cells, where GABA<sub>B</sub> receptors are rapidly and constitutively internalized by the clathrin-dependent pathway (Grampp et al. 2007) to early endosomes. Our finding that IGF2R can directly interact with the GABBR2 subunit of the GABA<sub>B</sub> receptor complex provides a mechanism for the internalization because IGF2R is itself constitutively endocytosed and trafficked to the endosome compartment through clathrin-mediated uptake via “YSKV” motifs in its cytoplasmic region (Jadot et al. 1992). This is also consistent with the regulatory role for the GABBR2 subunit in the uptake of the GABA<sub>B</sub> receptor complex (Hannan et al. 2011) and the fact that antibodies directed against the extracellular region, but not receptor agonists, can inhibit GABA<sub>B</sub> receptor endocytosis (Grampp et al. 2007). A similar role for IGF2R in down-regulating cell surface receptor levels has been shown for CD26 in activated T cells (Ikushima et al. 2000). This finding may now provide new opportunities to regulate GABA<sub>B</sub> receptor signaling to treat neurological conditions such as epilepsy and depression.

In summary, we have applied a genome-scale, cell-based genetic method based on CRISPR technology to identify genes that mediate cellular recognition processes. We believe that this approach, because it makes no prior assumptions regarding the biochemical nature of the cell surface receptor, will make novel and valuable insights into the molecular basis of cellular adhesion and signaling.

## Methods

### Recombinant protein expression and quantitation

All proteins were expressed in HEK293 cells maintained in FreeStyle media (Life Technologies) supplemented with 50 µg/mL

G418 and 0.1% Kolliphor P188 essentially as previously described (Kerr and Wright 2012). For routine culture,  $2.5 \times 10^7$  HEK293 cells were seeded in 500-mL Erlenmeyer culture flasks containing 100 mL culture media and cultured in a shaker set at 37°C, 5% CO<sub>2</sub>, 70% humidity, and 125 rpm. To maintain a logarithmic growth phase, cells were diluted into fresh media every 2–3 d when the cell density reached  $\sim 2 \times 10^6$ /mL. HEK293 cells were transiently transfected as previously described (Durocher et al. 2002); briefly, cells were prepared 48 h prior to transfection by seeding at a density of  $5 \times 10^5$  cells/mL, transfected, and incubated for 5 d before supernatants were harvested by removing cells by centrifugation at 3220g for 5 min and cell debris by filtering through a 0.2-µm filter. For the production of biotinylated proteins, the culture media was supplemented with 100 µM D-biotin (Bushell et al. 2008) and cotransfected with a plasmid encoding a secreted *Escherichia coli* BirA enzyme (Addgene plasmid no. 64395) as previously described (Kerr and Wright 2012). Supernatants containing pentameric recombinant proteins were stored at 4°C until further use, whereas the monomeric biotinylated his-tagged proteins were enriched on Ni<sup>2+</sup>-NTA Sepharose beads (Jena Biosciences) using gravity flow chromatography with polypropylene columns with a 50- to 100-µL bed volume (Qiagen). The beads were washed once with wash buffer (20 mM phosphate, 0.5 M NaCl, and 40 mM imidazole at pH 7.4) before the samples were eluted with 300–500 µL of elution buffer (20 mM phosphate, 0.5 M NaCl, and 400 mM imidazole at pH 7.4). Eluted proteins were separated by SDS-PAGE under reducing conditions and visualized with InstantBlue (Expedeon) to confirm their size and integrity following the manufacturer's recommendations. A list of the expression plasmids for human proteins used in this study can be found in Supplemental Table S2.

Biotinylated protein expression levels were quantified by ELISA as previously described (Kerr and Wright 2012). Briefly, proteins were captured on 96-well streptavidin-coated plates (NUNC) for 1 h before adding 10 µg/mL primary antibody recognizing the rat Cd4d3+4 tag (mouse anti-rat Cd4d3+4, clone OX68, Bio-Rad [MCA1022R]) common to all recombinant proteins for another hour. Plates were washed 3× in PBS/0.1% Tween-20 (PBST) before adding 100 µL of an anti-mouse alkaline phosphatase conjugate (Sigma, A3562) at 0.2 µg/mL. Plates were washed 3× PBST and 1× PBS before adding 100 µL *p*-nitrophenyl phosphate (Sigma 104 alkaline phosphatase substrate) at 1 mg/mL in diethanolamine buffer (10% diethanolamine [v/v], 0.5 mM MgCl<sub>2</sub>, 10 mM NaN<sub>3</sub> at pH 9.8). Absorbance readings were taken 15–20 min post substrate addition at 405 nm on a PHERAstar plus (BMG Laboratories). Beta-lactamase-tagged proteins were quantified using nitrocefin hydrolysis as previously described (Bushell et al. 2008). The proteins were normalized to enzyme activity corresponding to  $\sim 1$  nmol/min, which corresponds to complete hydrolysis of 14.5 nmol nitrocefin in  $\sim 15$  min.

### Plate-based direct protein interaction assay

A biotinylated “bait” protein consisting of the entire ectodomain of GABBR2 and controls were first immobilized in a well of a streptavidin-coated 96-well microtiter plate (NUNC) at a concentration that saturated the biotin binding capacity of the well and probed for direct interactions with the entire ectodomain of IGF2R expressed as a beta-lactamase-tagged “prey.” The plate was washed 2× in PBST, after which normalized beta-lactamase-tagged “prey” (IGF2R and controls) proteins were added to the wells for 1 h. Following another wash step (2× with PBST and a final wash with only PBS), 100 µL of 125 µg/mL nitrocefin was added, and prey capture was quantified by measuring the absorbance of nitrocefin hydrolysis products at 485 nm on a PHERAstar plus

(BMG Laboratories). Biotinylated Cd4d3+4 tag alone was used as a negative control bait, and a biotinylated anti-Cd4 mAb (anti-prey) used as a positive control as required.

Where soluble monosaccharides were used in blocking experiments, prey proteins were first incubated with a range of concentrations (10–0.04 mM) of M6P or mannose for 1 h, prior to incubation with bait proteins. To remove N-linked glycans from soluble recombinant GABBR2, 1500 U of PNGase F (New England Biolabs) was added to 10 µg of GABBR2 and incubated for a duration ranging from 1–16 h at 37°C.

### Cell culture of human cell lines

NCI-SNU-1, HEL, and SK-N-SH cells were cultured in RPMI 1640 (Life Technologies) supplemented with 10% heat-inactivated (50°C for 20 min) FBS, 1 mM sodium pyruvate (Life Technologies), 10 mM D-glucose (Sigma), and penicillin–streptomycin at 37°C with 5% CO<sub>2</sub>. HEK293 cells were cultured in FreeStyle media (Life Technologies) supplemented with 1% heat-inactivated FBS in shaker flasks in a shaking incubator set at 125 rpm, 37°C, 70% humidity, and 5% CO<sub>2</sub>. To maintain a logarithmic growth phase, NCI-SNU-1, HEL, and HEK-293 cells were diluted into fresh media every 2–3 d once the cell density reached  $\sim 1 \times 10^6$ /mL. SK-N-SH cells were grown in a monolayer, and the culture medium was changed every 3 d. The cells were passaged once they reached  $\sim 80\%$  confluence. NCI-SNU-1 and SK-N-SH cells were obtained from the Sanger Institute Cancer Cell Line Panel ([https://cancer.sanger.ac.uk/cell\\_lines](https://cancer.sanger.ac.uk/cell_lines)). HEK293 cells were obtained from Dr. Yves Durocher (Durocher et al. 2002). HEL cells were purchased from Deutsche Sammlung von Mikroorganismen und Zellkulturen (DSMZ). All cell lines were tested and found to be mycoplasma free.

### Cell binding assay with recombinant proteins and mAbs

#### *Binding assay using biotinylated monomeric proteins*

To increase binding avidity, biotinylated monomeric Cd4d3+4-tagged proteins were multimerized around streptavidin–phycoerythrin (PE). To ensure all biotin binding sites on the streptavidin were occupied and to minimize the presence of excess monomer, serial dilutions of biotinylated protein samples were titrated against a fixed concentration of streptavidin–PE (100 µL of 0.1 µg/mL) for 20 min at room temperature before transferring to a streptavidin-coated plate and assaying for the capture of any excess biotinylated Cd4d3+4-tagged proteins using the OX68 ELISA. The minimal dilution at which all biotinylated Cd4d3+4-tagged protein was captured was subsequently used to create tetramers. A 10× tetramer staining solution was prepared using 4 µg/mL streptavidin–PE and the appropriate biotinylated protein dilution by incubating for 30 min at room temperature and then diluted to 1× and 100 µL added to  $0.5\text{--}1 \times 10^6$  cells in an U-bottomed 96-well microtiter plates and incubated for 1 h at room temperature. Where the anti-BSG antibody was used in blocking experiments, cells were first incubated with 10 µg/mL antibody (or isotype-matched control) for 1 h prior to incubation with the RH5–streptavidin–PE complex. Cells were washed once with wash buffer (PBS with Ca<sup>2+</sup>/Mg<sup>2+</sup> [HyClone, Sigma] supplemented with 1% BSA) and analyzed by flow cytometry.

#### *Binding assay using pentameric protein supernatants*

The 3xFLAG and beta-lactamase-tagged pentameric proteins were quantified directly from supernatants and normalized to  $\sim 1$  nmol/min using the beta-lactamase enzyme activity. Next, 100 µL of diluted proteins was added to  $0.5\text{--}1 \times 10^6$  cells in a U-bottomed 96-well microtiter plates for 1 h at room temperature. Following

a wash with wash buffer, 100 µL PE-conjugated anti-FLAG antibody (0.5 µg/mL, Abcam ab72469) was added to the samples and incubated for 1 h. The cells were again washed once in wash buffer and analyzed by flow cytometry.

#### *Binding assay with mAbs*

For antibody staining of cell surface proteins, 50 µL of 1 µg/mL primary antibody was incubated with  $1 \times 10^6$  cells in 96-well U bottom plates. The cells were washed after 1 h of primary antibody incubation, after which 100 µL of an appropriate secondary antibody, also conjugated to PE, was used at 0.1 µg/mL. For IGF2R staining, anti-IGF2R mAb-clone-2G11 (Abcam, ab2733) was used. For staining of GABA<sub>B</sub> receptors, anti-GABA B receptor 1 antibody (Abcam, ab55051) was used.

#### *Binding assay with transiently transfected cells*

Human IGF2R was expressed on the surface of transfected cells using an expression construct in which its cytoplasmic region was replaced by eGFP, as previously described (Sollner and Wright 2009). NCI-SNU-1 cells, which do not have detectable levels of plasma membrane IGF2R expression, were transiently transfected with either IGF2R-TM-eGFP or CD200R1-TM-eGFP as a control and probed for binding interactions either with GABBR2 ectodomain presented as a tetramer around streptavidin–PE or with an anti-IGF2R mAb (ab2733). Human GABA<sub>B</sub> receptor heterodimer was expressed on the surface of parental and IGF2R-KO HEK293 cell lines by transiently cotransfecting both GABBR1 and GABBR2 cDNA constructs using Lipofectamine 2000 (Invitrogen) according to the manufacturer's protocol. CD200-TM-eGFP was transfected into the cells as a transfection control. The surface level of GABA<sub>B</sub> receptors was analyzed 3 d after transfection by surface-staining using a GABAB receptor 1 antibody (Abcam, ab55051). Where appropriate, the clathrin-pathway inhibitor sucrose was added at 450 µM. All flow cytometry was performed on a Becton-Dickinson LSRI Fortessa flow cytometer, collecting between 10,000 to 30,000 events; live cells were gated using forward and side scatter. PE was excited at a wavelength of 561 nm and emission detected using a 582/15 band pass filter; BFP was excited at 405 nm and the emission detected using a 450/50 band pass filter; GFP was excited at 488 nm and the emission detected using 530/30 band pass filter. Analysis was performed using FlowJo software (Treestar).

#### *Immunofluorescence of SK-N-SH cells*

SK-N-SH cells grown as a monolayer in coverslips were fixed in 4% formalin. Fixed cells were permeabilized in PBS/Triton X-100 0.1% and blocked in PBS/2% BSA. Cells were stained with primary antibodies: 1:400 anti-IGF2R (rabbit monoclonal, Abcam 124767) and 1:250 anti-GABBR1 antibody (mouse monoclonal, Abcam ab55051) overnight at 4°C. After three washes in the wash buffer (PBS/2% BSA), the cells were incubated with secondary antibodies (goat anti-mouse IgG [H+L] AlexaFluor 488 [Molecular Probes A-11001], donkey anti-rabbit IgG [H+L] AlexaFluor 647 [Abcam 150075]) diluted 1:500 for 1 h at room temperature. The coverslips were washed twice in wash buffer and three times in PBS and mounted on a microscope glass in SlowFade with DAPI (Molecular Probes) before images were captured with Leica SP5/DM6000 confocal microscope.

#### *Lentiviral vectors*

The Human Improved Genome-Wide Knockout CRISPR Library v1 (Addgene no. 67989), lentiviral Cas9 reporter plasmids pKLV2-

U6gRNA5(gGFP)-PGKBFP2AGFP-W (Addgene no. 67980) and pKLV2-U6gRNA5(Empty)-PGKBFP2AGFP-W (Addgene no. 67979), lentiviral vector expressing Cas9 fused with the blasticidin-resistant gene at the C terminus pKLV2-EF1a-Cas9Bsd-W (Addgene no. 68343), and lentiviral CRISPR gRNA expression vector pKLV2-U6gRNA5(BbsI)-PGKpuro2ABFP-W (Addgene no. 67974) were used in this study. All gene-specific gRNAs were cloned into BbsI site of the gRNA expression vector as described previously (Tzelepis et al. 2016).

### Preparation of lentivirus and transductions

Lentivirus were prepared by transfecting HEK293-FT cells as previously described (Koike-Yusa et al. 2014). All Cas9-expressing human cell lines were selected following transduction of cells with lentivirus prepared from the pKLV2-EF1a-Cas9Bsd-W plasmid (Tzelepis et al. 2016). Polybrene (8  $\mu$ g/mL) was added for the transduction of NCI-SNU-1 and HEL cells. Cells were selected using 20  $\mu$ g/mL blasticidin for the HEK293 and NCI-SNU-1 cell lines and 10  $\mu$ g/mL for HEL cells 2 d following transduction. We found that polyclonal Cas9-expressing lines contained a variable fraction (up to 30%) that were refractory to gene editing, which affected the sensitivity of our screens. We therefore always established clonal high Cas9 activity cell lines by sorting individual blasticidin-resistant cells into wells of 96-well plates (MoFlo XDP), which were further expanded and tested for Cas9 activity using the GFP-BFP system (Tzelepis et al. 2016). In brief, cells were transduced with lentivirus encoding GFP, BFP, and a gRNA targeting *GFP* (pKLV2-U6gRNA5(gGFP)-PGKBFP2AGFP-W) or the same construct with an “empty” gRNA (pKLV2-U6gRNA5(Empty)-PGKBFP2AGFP-W) as a negative control. High-activity Cas9 stable cell lines were selected by examining the ratio of BFP only to GFP-BFP double-positive cells transduced by the two lentiviruses. These clonal cell lines were expanded and further tested by targeting an endogenous gene encoding the BSG cell surface protein using lentivirus prepared using a plasmid encoding puromycin, BFP, and a gRNA targeting *BSG* [pKLV2-U6gRNA5(BbsI-gBSG)-PGKpuro2ABFP-W]. The surface expression of BSG was quantified by flow cytometry using an anti-BSG mAb (MEM-M6/6, Abcam ab119114) 8 d post transduction to validate high Cas9 efficiency.

### Lentiviral transduction of HEK293 cells with genome-scale gRNA library

The CRISPR-KO library we used in this study contains 90,709 gRNAs targeting 18,009 human genes at a coverage of five gRNAs per gene (Tzelepis et al. 2016). The backbone vector contains the optimized gRNA scaffold, which can maximize genome editing efficiency. A genome-scale “knock-out” library of HEK293-Cas9 cells was produced by transducing  $3 \times 10^7$  cells such that ~30% of the total cell population were transduced to increase the chances that each cell just received a single gRNA. The transduced (BFP-positive) cells were harvested 3 d after transduction using a cell sorter (MoFlo XDP), and libraries containing at least  $5 \times 10^6$  cells were selected. The libraries were cultured in media containing 2  $\mu$ g/mL puromycin to remove the contaminated nontransduced cells, and at every passage, at least 10 $\times$  the initial library (starting cell number on day three) for each library was seeded into new culture flasks. Phenotyping screens for cell surface binding events were carried out between 9 and 16 d post transduction.

### Lentiviral transduction of NCI-SNU-1 and HEL cells with genome-scale gRNA library

A spinoculation protocol was used to transduce HEL and NCI-SNU-1 cell lines. Two milliliters of  $5 \times 10^6$  cells/mL was aliquoted in 8 $\times$ 15-mL Falcon tubes and mixed with lentivirus together

with 8  $\mu$ g/mL polybrene and incubated at room temperature for 30 min followed by centrifugation for 100 min at 800g at 32°C. The supernatant was removed, and the cells from each Falcon tube were resuspended in 50 mL culture media. As with HEK293 cells, cells were sorted on day 3 post transduction to generate control and sample libraries and were grown further in media supplemented with 1  $\mu$ g/mL puromycin.

### Lentiviral transduction of human cells with individual gRNAs

For the validation of individual target genes, the corresponding gRNAs from the CRISPR library (Supplemental Table S3) were cloned into the BbsI site of pKLV2-U6gRNA5(BbsI)-PGKpuro2ABFP-W, and lentivirus was produced as described before. Cells ( $1 \times 10^6$  cells/well in six-well plate) were transduced with the lentivirus for at least 24 h. Polybrene (8  $\mu$ g/mL) was added during transduction of HEL and NCI-SNU-1 cell lines. Fresh media containing puromycin (2  $\mu$ g/mL for HEK293 cells and 1  $\mu$ g/mL for all other cell lines) was added to the cells after transduction and cultured for a further 8 d before use in binding assays.

### Cell surface phenotyping, selection, and amplification of selected gRNAs

To identify the genes required for cell surface recognition, the binding of the soluble ligand (mAb or recombinant protein) of interest was first quantified by flow cytometry to a small panel of cell lines, and the cell line exhibiting the highest level of binding was typically selected for genome-wide screening. Genome-scale knock-out libraries were phenotyped by cell surface staining using flow cytometry between 9 and 16 d post transduction. The mutant library was divided into two parts: At least  $5 \times 10^7$  cells from the mutant library were collected as “control” population for later analysis, whereas  $5 \times 10^7$ – $15 \times 10^7$  cells from the library were stained with appropriate reagent (recombinant protein or antibody) using the binding assay protocol as described above with minor modifications: Cells ( $5 \times 10^6$  cells/mL) were stained in 15-mL Falcon tubes with gentle rotation (6 rpm), the stained cells were then analyzed using an XDP flow sorter, and the BFP<sup>+</sup>/PE<sup>+</sup> cells were collected. The percentage of the total library population that was collected in each screen varied between 0.2% and 2.3% and are listed in Supplemental Table S4. We observed that stringent sorting thresholds resulted in the identification of few genes that had strong effects on binding loss, and increasing this threshold enabled the identification of additional genes that presumably had weaker effect sizes on the day of screening. We determined that collecting between 300,000 and 500,000 cells at a 0.5%–1% stringency threshold was generally an appropriate parameter that permitted the identification of genes with weaker effect sizes. All genetic screens except those performed using anti-CD59 mAb and TIGIT recombinant protein in this study were carried out once.

### Amplification of gRNAs, sequencing, and enrichment analysis

Genomic DNA extraction, PCR enrichment, and Illumina sequencing of gRNAs from both control and sorted samples were carried out as described previously (Koike-Yusa et al. 2014), except in sorted samples where the sorted cell number was less than 100,000. In that case, a cell lysate protocol was used to isolate gRNAs prior to PCR enrichment. Cell lysates were prepared from sorted cells by first aliquoting cells in a 96-well PCR plate (10,000 cells/well) and boiling the samples with 25  $\mu$ L water for 10 min at 95°C. Next, 5  $\mu$ L of 2 mg/mL freshly diluted Proteinase K was added to each sample and incubated for 1 h at 56°C and then heated for 10 min at 95°C to inactivate the protease.

The gRNAs were then amplified using 10- $\mu$ L cell lysates per PCR reaction.

For all samples, 19-bp single-end sequencing was performed using a custom sequencing primer 5'-TCTTCCGATCTCTTGTTGGAAGGACGAAACACCG-3'. The read count for each gRNA and gene level enrichment analysis was carried out using the MAGeCK statistical package (version 0.5.5) (Li et al. 2014) by comparing the read counts from the sorted population with those from the control population. Pathway analysis was also carried out using MAGeCK software with default settings with KEGG annotated pathways. All further analysis was carried out using R (R Core Team 2017). For all the target receptors identified, the ranks of the individual gRNAs targeting the related gene are listed in Supplemental Table S5.

### Quantifying gRNA representation in mutant cell libraries to ensure maintenance of library complexity

We observed that it was necessary to culture the cells for between 9 and 16 d after transduction to allow sufficient time for gene targeting and consequent loss of the encoded protein. During this time, gRNAs can be lost from the library for biological reasons (e.g., the gRNA targets a gene essential for cell survival) or for technical reasons (such as the need to passage the library because of cell growth that may create population restriction points reducing gRNA library representation). To ensure that gRNA representation is maintained using our cellular KO library preparation protocol, we quantified the gRNA abundance in at least  $5 \times 10^7$  cells on different days after transduction from two independent NCI-SNU-1-Cas9 libraries and one library of both HEK293-Cas9 and HEL-Cas9 cells. We observed that gRNA abundances were highly correlated between the biological replicates of NCI-SNU-1-Cas9 libraries, and similarly high correlations were observed among the three different cell line libraries at equivalent time points, demonstrating that the procedure we used to make genome-scale mutant libraries was reproducible (Supplemental Fig. S9). The correlation between the different mutant cell line libraries and the original plasmid population was lower, and so the most appropriate control for gRNA enrichment analysis was the cell line on that day rather than the population of gRNAs in the original plasmid population.

To further compare the gRNA abundances, the number of sequence reads for each gRNA between the original plasmid pool and each cell library was quantified. In the plasmid library, while a small fraction of gRNAs was under- or overrepresented, >82% of the gRNAs were uniformly distributed, with only an eightfold difference in abundance between the 10th and 90th percentiles. The cell-based mutant libraries created from this plasmid library also showed a uniform coverage, although a small drop in the overall representation of the gRNA library was observed in all cell lines as expected, and this decrease was more pronounced when comparing days 9–16 post transduction. It is likely that these depleted gRNAs target genes that are essential for cell growth in culture. To establish this, we performed a gene-level negative selection enrichment analysis to identify genes that were depleted in the mutant library compared with the plasmid library in all 3 d in the NCI-SNU-1 and HEK293 cell lines and on day 14 for HEL cells. As a quality control, we initially analyzed genes required for ribosome biosynthesis (annotations from KEGG-Ribosome), which are known to be essential and are often robustly identified in similar negative selection screens (Wang et al. 2014; Hart et al. 2015). Reassuringly, the majority of gRNAs targeting genes required for ribosome biosynthesis were among the most significantly depleted (FDR < 10%) across all 3 d in all three cell lines (Supplemental Fig. S10). Next, we attempted to identify which pathways were defined by the genes targeted by the enriched gRNAs using KEGG annotat-

ed pathways (Kanehisa et al. 2017). We observed that the pathways identified were in biological processes described as being essential, including those required for the spliceosome, cell cycle, purine and pyrimidine biosynthesis, and both DNA and RNA polymerases (Supplemental Data S4). These results provided further confidence that the cellular mutant libraries generated using our protocol retained their gRNA complexity and could be used for genome-scale screening.

### Data access

All reads from the genome-wide screening experiments in this study have been submitted to the European Nucleotide Archive (<http://www.ebi.ac.uk/ena>) under accession number ERP104831.

### Acknowledgments

We thank the Cytometry Core facility: Bee Ling Ng, Jennifer Graham, Sam Thompson, and Christopher Hall for help with FACS; Mandy Sanders for help with sequencing sample submission and Sanger Institute Core Sequencing facility for sequencing; and Kirsten Dundas for supervisory help. This work was supported by the Wellcome Trust grant 206194.

**Author contributions:** S.S., S.J.B., K.Y., and G.J.W. conceived and designed the study. S.S. performed most experiments and the data analysis. S.J.B. contributed to assay development and the RH5 experiments. A.C.M.C. contributed assistance with cell staining. S.S. and G.J.W. wrote the manuscript.

### References

- Baum J, Chen L, Healer J, Lopatnicki S, Boyle M, Triglia T, Ehlgren F, Ralph SA, Beeson JG, Cowman AF. 2009. Reticulocyte-binding protein homologue 5—an essential adhesin involved in invasion of human erythrocytes by *Plasmodium falciparum*. *Int J Parasitol* **39**: 371–380.
- Benke D. 2013. GABA<sub>B</sub> receptor trafficking and interacting proteins: targets for the development of highly specific therapeutic strategies to treat neurological disorders? *Biochem Pharmacol* **86**: 1525–1530.
- Brakebusch C, Fassler R. 2003. The integrin–actin connection, an eternal love affair. *EMBO J* **22**: 2324–2333.
- Burr ML, Spabier CE, Chan YC, Williamson JC, Woods K, Beavis PA, Lam EYN, Henderson MA, Bell CC, Stolzenburg S, et al. 2017. CMTM6 maintains the expression of PD-L1 and regulates anti-tumour immunity. *Nature* **549**: 101–105.
- Bushell KM, Sollner C, Schuster-Boeckler B, Bateman A, Wright GJ. 2008. Large-scale screening for novel low-affinity extracellular protein interactions. *Genome Res* **18**: 622–630.
- Crosnier C, Bustamante LY, Bartholdson SJ, Bei AK, Theron M, Uchikawa M, Mboup S, Ndir O, Kwiatkowski DP, Duraisingh MT, et al. 2011. Basigin is a receptor essential for erythrocyte invasion by *Plasmodium falciparum*. *Nature* **480**: 534–537.
- Durocher Y, Perret S, Kamen A. 2002. High-level and high-throughput recombinant protein production by transient transfection of suspension-growing human 293-EBNA1 cells. *Nucleic Acids Res* **30**: E9.
- Esko JD, Kimata K, Lindahl U. 2009. Proteoglycans and sulfated glycosaminoglycans. In *Essentials of glycobiology*, 2nd ed. (ed. Varki A, et al.), Chapter 16. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
- Grampp T, Sauter K, Markovic B, Benke D. 2007.  $\gamma$ -Aminobutyric acid type B receptors are constitutively internalized via the clathrin-dependent pathway and targeted to lysosomes for degradation. *J Biol Chem* **282**: 24157–24165.
- Hannan S, Wilkins ME, Dehghani-Tafti E, Thomas P, Baddeley SM, Smart TG. 2011.  $\gamma$ -Aminobutyric acid type B (GABA<sub>B</sub>) receptor internalization is regulated by the R2 subunit. *J Biol Chem* **286**: 24324–24335.
- Hart T, Chandrashekar M, Aregger M, Steinhart Z, Brown KR, MacLeod G, Mis M, Zimmermann M, Fradet-Turcotte A, Sun S, et al. 2015. High-resolution CRISPR screens reveal fitness genes and genotype-specific cancer liabilities. *Cell* **163**: 1515–1526.
- Hawkes C, Kar S. 2004. The insulin-like growth factor-II/mannose-6-phosphate receptor: structure, distribution and function in the central nervous system. *Brain Res Brain Res Rev* **44**: 117–140.

- Ikushima H, Munakata Y, Ishii T, Iwata S, Terashima M, Tanaka H, Schlossman SF, Morimoto C. 2000. Internalization of CD26 by mannose 6-phosphate/insulin-like growth factor II receptor contributes to T cell activation. *Proc Natl Acad Sci* **97**: 8439–8444.
- Jadot M, Canfield WM, Gregory W, Kornfeld S. 1992. Characterization of the signal for rapid internalization of the bovine mannose 6-phosphate/insulin-like growth factor-II receptor. *J Biol Chem* **267**: 11069–11077.
- Kamiyama S, Suda T, Ueda R, Suzuki M, Okubo R, Kikuchi N, Chiba Y, Goto S, Toyoda H, Saigo K, et al. 2003. Molecular cloning and identification of 3'-phosphoadenosine 5'-phosphosulfate transporter. *J Biol Chem* **278**: 25958–25963.
- Kanehisa M, Furumichi M, Tanabe M, Sato Y, Morishima K. 2017. KEGG: new perspectives on genomes, pathways, diseases and drugs. *Nucleic Acids Res* **45**: D353–D361.
- Kerr JS, Wright GJ. 2012. Avidity-based extracellular interaction screening (AVEXIS) for the scalable detection of low-affinity extracellular receptor-ligand interactions. *J Vis Exp* (61): e3881.
- Kirk P, Wilson MC, Heddle C, Brown MH, Barclay AN, Halestrap AP. 2000. CD147 is tightly associated with lactate transporters MCT1 and MCT4 and facilitates their cell surface expression. *EMBO J* **19**: 3896–3904.
- Koike-Yusa H, Li Y, Tan EP, Velasco-Herrera Mdel C, Yusa K. 2014. Genome-wide recessive genetic screening in mammalian cells with a lentiviral CRISPR-guide RNA library. *Nat Biotechnol* **32**: 267–273.
- Li W, Xu H, Xiao T, Cong L, Love MI, Zhang F, Irizarry RA, Liu JS, Brown M, Liu XS. 2014. MAGeCK enables robust identification of essential genes from genome-scale CRISPR/Cas9 knockout screens. *Genome Biol* **15**: 554.
- Ma H, Dang Y, Wu Y, Jia G, Anaya E, Zhang J, Abraham S, Choi JG, Shi G, Qi L, et al. 2015. A CRISPR-based screen identifies genes essential for West-Nile-virus-induced cell death. *Cell Rep* **12**: 673–683.
- Marceau CD, Puschnik AS, Majzoub K, Ooi YS, Brewer SM, Fuchs G, Swaminathan K, Mata MA, Elias JE, Sarnow P, et al. 2016. Genetic dissection of *Flaviviridae* host factors through genome-scale CRISPR screens. *Nature* **535**: 159–163.
- Merkulova M, Paunescu TG, Azroyan A, Marshansky V, Breton S, Brown D. 2015. Mapping the H<sup>+</sup> (V)-ATPase interactome: identification of proteins involved in trafficking, folding, assembly and phosphorylation. *Sci Rep* **5**: 14827.
- Ozkan E, Carrillo RA, Eastman CL, Weiszmann R, Waghray D, Johnson KG, Zinn K, Celniker SE, Garcia KC. 2013. An extracellular interactome of immunoglobulin and LRR proteins reveals receptor-ligand networks. *Cell* **154**: 228–239.
- Parnas O, Jovanovic M, Eisenhaure TM, Herbst RH, Dixit A, Ye CJ, Przybylski D, Platt RJ, Tirosh I, Sanjana NE, et al. 2015. A genome-wide CRISPR screen in primary immune cells to dissect regulatory networks. *Cell* **162**: 675–686.
- Pillay S, Meyer NL, Puschnik AS, Davulcu O, Diep J, Ishikawa Y, Jae LT, Wosen JE, Nagamine CM, Chapman MS, et al. 2016. An essential receptor for adeno-associated virus infection. *Nature* **530**: 108–112.
- R Core Team. 2017. *R: a language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria. <https://www.R-project.org/>.
- Reaves B, Banting G. 1994. Vacuolar ATPase inactivation blocks recycling to the trans-Golgi network from the plasma membrane. *FEBS Lett* **345**: 61–66.
- Rosmarin DM, Carette JE, Olive AJ, Starnbach MN, Brummelkamp TR, Ploegh HL. 2012. Attachment of *Chlamydia trachomatis* L2 to host cells requires sulfation. *Proc Natl Acad Sci* **109**: 10059–10064.
- Shalem O, Sanjana NE, Hartenian E, Shi X, Scott DA, Mikkelsen T, Heckl D, Ebert BL, Root DE, Doench JG, et al. 2014. Genome-scale CRISPR-Cas9 knockout screening in human cells. *Science* **343**: 84–87.
- Shalem O, Sanjana NE, Zhang F. 2015. High-throughput functional genomics using CRISPR-Cas9. *Nat Rev Genet* **16**: 299–311.
- Sollner C, Wright GJ. 2009. A cell surface interaction network of neural leucine-rich repeat receptors. *Genome Biol* **10**: R99.
- Tzelepis K, Koike-Yusa H, De Braekeleer E, Li Y, Metzakopian E, Dovey OM, Mupo A, Grinkevich V, Li M, Mazan M, et al. 2016. A CRISPR dropout screen identifies genetic vulnerabilities and therapeutic targets in acute myeloid leukemia. *Cell Rep* **17**: 1193–1205.
- Visser JJ, Cheng Y, Perry SC, Chastain AB, Parsa B, Masri SS, Ray TA, Kay JN, Wojtowicz WM. 2015. An extracellular biochemical screen reveals that FLRTs and Unc5s mediate neuronal subtype recognition in the retina. *eLife* **4**: e08149.
- Wang T, Wei JJ, Sabatini DM, Lander ES. 2014. Genetic screens in human cells using the CRISPR-Cas9 system. *Science* **343**: 80–84.
- Weiner GJ. 2015. Building better monoclonal antibody-based therapeutics. *Nat Rev Cancer* **15**: 361–370.
- Wright GJ. 2009. Signal initiation in biological systems: the properties and detection of transient extracellular protein interactions. *Mol Biosyst* **5**: 1405–1412.
- Wright GJ, Martin S, Bushell KM, Sollner C. 2010. High-throughput identification of transient extracellular protein interactions. *Biochem Soc Trans* **38**: 919–922.
- Wright KE, Hjerrild KA, Bartlett J, Douglas AD, Jin J, Brown RE, Illingworth JJ, Ashfield R, Clemmensen SB, de Jongh WA, et al. 2014. Structure of malaria invasion protein RH5 with erythrocyte basigin and blocking antibodies. *Nature* **515**: 427–430.
- Yan D, Lin X. 2009. Shaping morphogen gradients by proteoglycans. *Cold Spring Harb Perspect Biol* **1**: a002493.
- Zhang R, Miner JJ, Gorman MJ, Rausch K, Ramage H, White JP, Zuiani A, Zhang P, Fernandez E, Zhang Q, et al. 2016. A CRISPR screen defines a signal peptide processing pathway required by flaviviruses. *Nature* **535**: 164–168.
- Zotova A, Zotov I, Filatov A, Mazurov D. 2016. Determining antigen specificity of a monoclonal antibody using genome-scale CRISPR-Cas9 knockout library. *J Immunol Methods* **439**: 8–14.

Received October 17, 2017; accepted in revised form June 15, 2018.



## Genome-scale identification of cellular pathways required for cell surface recognition

Sumana Sharma, S. Josefin Bartholdson, Amalie C.M. Couch, et al.

*Genome Res.* 2018 28: 1372-1382 originally published online June 18, 2018

Access the most recent version at doi:[10.1101/gr.231183.117](https://doi.org/10.1101/gr.231183.117)

---

### Supplemental Material

<http://genome.cshlp.org/content/suppl/2018/08/10/gr.231183.117.DC1>

### References

This article cites 39 articles, 13 of which can be accessed free at:  
<http://genome.cshlp.org/content/28/9/1372.full.html#ref-list-1>

### Open Access

Freely available online through the *Genome Research* Open Access option.

### Creative Commons License

This article, published in *Genome Research*, is available under a Creative Commons License (Attribution 4.0 International), as described at <http://creativecommons.org/licenses/by/4.0/>.

### Email Alerting Service

Receive free email alerts when new articles cite this article - sign up in the box at the top right corner of the article or [click here](#).

---

---

To subscribe to *Genome Research* go to:  
<http://genome.cshlp.org/subscriptions>

---

## ARTICLE AND VIDEO LICENSE AGREEMENT - UK

Title of Article:	Cell surface receptor identification using genome-scale CRISPR/Cas9 genetic screens
Author(s):	Sumana Sharma and Gavin J Wright

Item 1: The Author elects to have the Materials be made available (as described at <http://www.jove.com/publish>) via:



Standard Access



Open Access

Item 2: Please select one of the following items:



The Author is **NOT** a United States government employee.



The Author is a United States government employee and the Materials were prepared in the course of his or her duties as a United States government employee.



The Author is a United States government employee but the Materials were NOT prepared in the course of his or her duties as a United States government employee.

### ARTICLE AND VIDEO LICENSE AGREEMENT

1. **Defined Terms.** As used in this Article and Video License Agreement, the following terms shall have the following meanings: "**Agreement**" means this Article and Video License Agreement; "**Article**" means the article specified on the last page of this Agreement, including any associated materials such as texts, figures, tables, artwork, abstracts, or summaries contained therein; "**Author**" means the author who is a signatory to this Agreement; "**Collective Work**" means a work, such as a periodical issue, anthology or encyclopedia, in which the Materials in their entirety in unmodified form, along with a number of other contributions, constituting separate and independent works in themselves, are assembled into a collective whole; "**CRC License**" means the Creative Commons Attribution 3.0 Agreement (also known as CC-BY), the terms and conditions of which can be found at: <http://creativecommons.org/licenses/by/3.0/us/legalcode>; "**CRC NonCommercial License**" means the Creative Commons Attribution-NonCommercial 3.0 Agreement (also known as CC-BY-NC), the terms and conditions of which can be found at: <http://creativecommons.org/licenses/by-nc/3.0/legalcode>; "**Derivative Work**" means a work based upon the Materials or upon the Materials and other pre-existing works, such as a translation, musical arrangement, dramatization, fictionalization, motion picture version, sound recording, art reproduction, abridgment, condensation, or any other form in which the Materials may be recast, transformed, or adapted; "**Institution**" means the institution, listed on the last page of this Agreement, by which the Author was employed at the time of the creation of the Materials; "**JoVE**" means MyJoVE Corporation, a Massachusetts corporation and the publisher of The Journal of Visualized Experiments; "**Materials**" means the Article and / or the Video; "**Parties**" means the Author and JoVE; "**Video**" means any video(s) made by the Author, alone or in conjunction with any other parties, or by JoVE or its

affiliates or agents, individually or in collaboration with the Author or any other parties, incorporating all or any portion of the Article, and in which the Author may or may not appear.

2. **Background.** The Author, who is the author of the Article, in order to ensure the dissemination and protection of the Article, desires to have the JoVE publish the Article and create and transmit videos based on the Article. In furtherance of such goals, the Parties desire to memorialize in this Agreement the respective rights of each Party in and to the Article and the Video.

3. **Grant of Rights in Article.** In consideration of JoVE agreeing to publish the Article, the Author hereby grants to JoVE, subject to **Sections 4 and 7** below, the exclusive, royalty-free, perpetual (for the full term of copyright in the Article, including any extensions thereto) license (a) to publish, reproduce, distribute, display and store the Article in all forms, formats and media whether now known or hereafter developed (including without limitation in print, digital and electronic form) throughout the world, (b) to translate the Article into other languages, create adaptations, summaries or extracts of the Article or other Derivative Works (including, without limitation, the Video) or Collective Works based on all or any portion of the Article and exercise all of the rights set forth in (a) above in such translations, adaptations, summaries, extracts, Derivative Works or Collective Works and (c) to license others to do any or all of the above. The foregoing rights may be exercised in all media and formats, whether now known or hereafter devised, and include the right to make such modifications as are technically necessary to exercise the rights in other media and formats. If the "Open Access" box has been checked in **Item 1** above, JoVE and the Author hereby grant to the public all such rights in the Article as provided in, but subject to all limitations and requirements set forth in, the CRC License. If the "Standard Access" box

612542.6 For questions, please contact us at [submissions@jove.com](mailto:submissions@jove.com) or +1.617.945.9051.

## ARTICLE AND VIDEO LICENSE AGREEMENT - UK

has been checked in **Item 1** above, JoVE and the Author hereby grant to the public all such rights in the Article as provided in, but subject to all limitations and requirements set forth in, the CRC NonCommercial License.

4. **Retention of Rights in Article.** Notwithstanding the exclusive license granted to JoVE in **Section 3** above, the Author shall, with respect to the Article, retain the non-exclusive right to use all or part of the Article for the non-commercial purpose of giving lectures, presentations or teaching classes, and to post a copy of the Article on the Institution's website or the Author's personal website, in each case provided that a link to the Article on the JoVE website is provided and notice of JoVE's copyright in the Article is included. All non-copyright intellectual property rights in and to the Article, such as patent rights, shall remain with the Author.

5. **Grant of Rights in Video - Standard Access.** This **Section 5** applies if the "Standard Access" box has been checked in **Item 1** above or if no box has been checked in **Item 1** above. In consideration of JoVE agreeing to produce, display or otherwise assist with the Video, the Author hereby acknowledges and agrees that, subject to **Section 7** below, JoVE is and shall be the sole and exclusive owner of all rights of any nature, including, without limitation, all copyrights, in and to the Video. To the extent that, by law, the Author is deemed, now or at any time in the future, to have any rights of any nature in or to the Video, the Author hereby disclaims all such rights and transfers all such rights to JoVE.

6. **Grant of Rights in Video - Open Access.** This **Section 6** applies only if the "Open Access" box has been checked in **Item 1** above. In consideration of JoVE agreeing to produce, display or otherwise assist with the Video, the Author hereby grants to JoVE, subject to **Section 7** below, the exclusive, royalty-free, perpetual (for the full term of copyright in the Article, including any extensions thereto) license (a) to publish, reproduce, distribute, display and store the Video in all forms, formats and media whether now known or hereafter developed (including without limitation in print, digital and electronic form) throughout the world, (b) to translate the Video into other languages, create adaptations, summaries or extracts of the Video or other Derivative Works or Collective Works based on all or any portion of the Video and exercise all of the rights set forth in (a) above in such translations, adaptations, summaries, extracts, Derivative Works or Collective Works and (c) to license others to do any or all of the above. The foregoing rights may be exercised in all media and formats, whether now known or hereafter devised, and include the right to make such modifications as are technically necessary to exercise the rights in other media and formats.

7. **Government Employees.** If the Author is a United States government employee and the Article was prepared in the course of his or her duties as a United States government employee, as indicated in **Item 2** above, and any of the licenses or grants granted by the Author hereunder exceed the scope of the 17 U.S.C. 403, then the rights granted hereunder shall be limited to the maximum rights permitted under such statute. In such case, all provisions contained herein that are not in conflict with

such statute shall remain in full force and effect, and all provisions contained herein that do so conflict shall be deemed to be amended so as to provide to JoVE the maximum rights permissible within such statute.

8. **Protection of the work.** The Author(s) authorize JoVE to take steps in the Author(s) name and on their behalf if JoVE believes some third party could be infringing or might infringe the copyright of either the Author's Article and/or Video.

9. **Likeness, Privacy, Personality.** The Author hereby grants JoVE the right to use the Author's name, voice, likeness, picture, photograph, image, biography and performance in any way, commercial or otherwise, in connection with the Materials and the sale, promotion and distribution thereof. The Author hereby waives any and all rights he or she may have, relating to his or her appearance in the Video or otherwise relating to the Materials, under all applicable privacy, likeness, personality or similar laws.

10. **Author Warranties.** The Author represents and warrants that the Article is original, that it has not been published, that the copyright interest is owned by the Author (or, if more than one author is listed at the beginning of this Agreement, by such authors collectively) and has not been assigned, licensed, or otherwise transferred to any other party. The Author represents and warrants that the author(s) listed at the top of this Agreement are the only authors of the Materials. If more than one author is listed at the top of this Agreement and if any such author has not entered into a separate Article and Video License Agreement with JoVE relating to the Materials, the Author represents and warrants that the Author has been authorized by each of the other such authors to execute this Agreement on his or her behalf and to bind him or her with respect to the terms of this Agreement as if each of them had been a party hereto as an Author. The Author warrants that the use, reproduction, distribution, public or private performance or display, and/or modification of all or any portion of the Materials does not and will not violate, infringe and/or misappropriate the patent, trademark, intellectual property or other rights of any third party. The Author represents and warrants that it has and will continue to comply with all government, institutional and other regulations, including, without limitation all institutional, laboratory, hospital, ethical, human and animal treatment, privacy, and all other rules, regulations, laws, procedures or guidelines, applicable to the Materials, and that all research involving human and animal subjects has been approved by the Author's relevant institutional review board.

11. **JoVE Discretion.** If the Author requests the assistance of JoVE in producing the Video in the Author's facility, the Author shall ensure that the presence of JoVE employees, agents or independent contractors is in accordance with the relevant regulations of the Author's institution. If more than one author is listed at the beginning of this Agreement, JoVE may, in its sole discretion, elect not take any action with respect to the Article until such time as it has received complete, executed Article and Video License Agreements from each such author. JoVE reserves the right, in its absolute and sole

## ARTICLE AND VIDEO LICENSE AGREEMENT - UK

discretion and without giving any reason therefore, to accept or decline any work submitted to JoVE. JoVE and its employees, agents and independent contractors shall have full, unfettered access to the facilities of the Author or of the Author's institution as necessary to make the Video, whether actually published or not. JoVE has sole discretion as to the method of making and publishing the Materials, including, without limitation, to all decisions regarding editing, lighting, filming, timing of publication, if any, length, quality, content and the like.

**12. Indemnification.** The Author agrees to indemnify JoVE and/or its successors and assigns from and against any and all claims, costs, and expenses, including attorney's fees, arising out of any breach of any warranty or other representations contained herein. The Author further agrees to indemnify and hold harmless JoVE from and against any and all claims, costs, and expenses, including attorney's fees, resulting from the breach by the Author of any representation or warranty contained herein or from allegations or instances of violation of intellectual property rights, damage to the Author's or the Author's institution's facilities, fraud, libel, defamation, research, equipment, experiments, property damage, personal injury, violations of institutional, laboratory, hospital, ethical, human and animal treatment, privacy or other rules, regulations, laws, procedures or guidelines, liabilities and other losses or damages related in any way to the submission of work to JoVE, making of videos by JoVE, or publication in JoVE or elsewhere by JoVE. The Author shall be responsible for, and shall hold JoVE harmless from, damages caused by lack of sterilization, lack of cleanliness or by contamination due to the making of a video by JoVE its employees, agents or independent contractors. All sterilization, cleanliness or


decontamination procedures shall be solely the responsibility of the Author and shall be undertaken at the Author's expense. All indemnifications provided herein shall include JoVE's attorney's fees and costs related to said losses or damages. Such indemnification and holding harmless shall include such losses or damages incurred by, or in connection with, acts or omissions of JoVE, its employees, agents or independent contractors.

**13. Fees.** To cover the cost incurred for publication, JoVE must receive payment before production and publication of the Materials. Payment is due in 21 days of invoice. Should the Materials not be published due to an editorial or production decision, these funds will be returned to the Author. Withdrawal by the Author of any submitted Materials after final peer review approval will result in a US\$1,200 fee to cover pre-production expenses incurred by JoVE. If payment is not received by the completion of filming, production and publication of the Materials will be suspended until payment is received.

**14. Transfer, Governing Law.** This Agreement may be assigned by JoVE and shall inure to the benefits of any of JoVE's successors and assignees. This Agreement shall be governed and construed by the internal laws of the Commonwealth of Massachusetts without giving effect to any conflict of law provision thereunder. This Agreement may be executed in counterparts, each of which shall be deemed an original, but all of which together shall be deemed to be one and the same agreement. A signed copy of this Agreement delivered by facsimile, e-mail or other means of electronic transmission shall be deemed to have the same legal effect as delivery of an original signed copy of this Agreement.

A signed copy of this document must be sent with all new submissions. Only one Agreement is required per submission.

### CORRESPONDING AUTHOR

Name:	Sumana Sharma		
Department:	Cell Signalling Laboratory		
Institution:	Wellcome Sanger Institute, EMBL-EBI		
Title:	Dr		
Signature:		Date:	09/19/2019



Please submit a **signed** and **dated** copy of this license by one of the following three methods:

1. Upload an electronic version on the JoVE submission site
2. Fax the document to +1.866.381.2236
3. Mail the document to JoVE / Attn: JoVE Editorial / 1 Alewife Center #200 / Cambridge, MA 02140

# Signature Certificate

Document Ref.: BYV7P-WQDRX-CRFRD-RT2VF

Document signed by:

	<p><b>Sumana Sharma</b></p> <p>Verified E-mail: sumana@ebi.ac.uk</p> <p>IP: 193.62.194.248      Date: 19 Sep 2019 12:42:29 UTC</p>	<p><i>Sumana Sharma</i></p> 
---	--	---

Document completed by all parties on:  
19 Sep 2019 12:42:29 UTC

Page 1 of 1



Signed with PandaDoc.com

PandaDoc is the document platform that boosts your company's revenue by accelerating the way it transacts.

