High-throughput genome-wide phenotypic screening via immunomagnetic cell sorting

Barbara Mair [1,7], Peter M. Aldridge [2,7], Randy S. Atwal[2], David Philpott[3], Meng Zhang[2],
Sanna N. Masud[1,4], Mahmoud Labib [2], Amy H. Y. Tong[1], Edward H. Sargent [3], Stéphane Angers [2,5],
Jason Moffat [1,4,6]* and Shana O. Kelley [2,5]*

Genome-scale functional genetic screens are used to identify key genetic regulators of a phenotype of interest. However, the identification of genetic modifications that lead to a phenotypic change requires sorting large numbers of cells, which increases operational times and costs and limits cell viability. Here, we introduce immunomagnetic cell sorting facilitated by a micro-fluidic chip as a rapid and scalable high-throughput method for loss-of-function phenotypic screening using CRISPR–Cas9. We used the method to process an entire genome-wide screen containing more than 10^8 cells in less than 1 h—considerably surpassing the throughput achieved by fluorescence-activated cell sorting, the gold-standard technique for phenotypic cell sorting—while maintaining high levels of cell viability. We identified modulators of the display of CD47, which is a negative regulator of phagocytosis and an important cell-surface target for immuno-oncology drugs. The top hit of the screen, the glutaminyl cyclase QPCTL, was validated and shown to modify the N-terminal glutamine of CD47. The method presented could bridge the gap between fluorescence-activated cell sorting and less flexible yet higher-throughput systems such as magnetic-activated cell sorting.

Genome-scale genetic screens performed using CRISPR can be used to analyse determinants of cell viability and have enabled key studies of essential genes within the human genome[1]. These screens are also powerful tools for identification of regulators of phenotypes of interest, and can identify key modulators of the expression of therapeutically relevant proteins[2,3]. This type of functional phenotypic screening provides critical information for the design and development of new therapeutic strategies.

Genome-scale phenotypic genetic screens produce hundreds of millions of cells that can be studied to isolate rare altered phenotypes using antibodies targeting specific markers of interest or genetically encoded reporters. Phenotypic changes in genetically altered cells can then be detected by monitoring changes in levels of antibody binding or a change in reporter activity. Fluorescence-activated cell sorting (FACS) is the gold standard for sorting and isolation of antibody-labelled cells, but suffers from limited throughput for high-coverage genome-scale screening applications, resulting in reduced cell viability or requiring cell fixation owing to long sorting times. Furthermore, FACS requires advanced costly instrumentation and can induce perturbations in cellular metabolism and functions that can influence the results obtained[4,5]. Thus, although phenotype-based genetic screens[2,3,6–10] can facilitate the identification of regulators of therapeutically relevant proteoforms detected with labelled probes, they are less common compared with proliferation-based screens because of the challenges related to implementation. Rapid and robust selection approaches for the targeted sorting of live cells are required to realize the full potential of phenotype-based genome-scale screens for functional discovery and further annotation of the human genome.

Microfluidic approaches offer precise control of forces and flows at scales that are comparable to most cells, and have great potential to manipulate or fractionate CRISPR-edited cells. For example, membrane deformation induced in narrow microfluidic channels has been used to increase the permeability and, therefore, the efficiency of Cas9 delivery in cells that are difficult to transfect[11]. More recently, microfluidic sorting has been applied to a CRISPR loss-of-function screen, separating cells with increased deformability to potentially identify regulators of genes that promote a metastatic phenotype[12]. However, the major challenge of handling and sorting large collections of CRISPR-edited cells on the basis of phenotypic changes has not yet been addressed with a general solution and represents a major unmet need. We therefore set out to combine the advantages of precise microfluidic cell manipulation with high-throughput sorting capacities and genome-scale CRISPR technology to advance phenotypic screening towards targeted functional profiling.

Here we describe high-throughput microfluidic immunomagnetic cell sorting (MICS) combined with genome-wide CRISPR–Cas9 loss-of-function screening as an unbiased method for the identification of modifiers of protein or biomarker expression. Whereas FACS has limited throughput owing to the need to process and analyse cellular fluorescence for each cell sequentially, we designed a simultaneous sorting device with a network of magnetic guides that sense and deflect cells, which are labelled on the basis of their loading with antibody-labelled nanoparticles, as they travel through the device along with millions of other cells. On the basis of our previous work related to tumour-cell sorting[13], we developed a device that is able to preserve high levels of cell viability after

[1]Donnelly Centre, University of Toronto, Toronto, Ontario, Canada. [2]Leslie Dan Faculty of Pharmacy, University of Toronto, Toronto, Ontario, Canada. [3]Department of Electrical & Computer Engineering, University of Toronto, Toronto, Ontario, Canada. [4]Department of Molecular Genetics, University of Toronto, Toronto, Ontario, Canada. [5]Department of Biochemistry, University of Toronto, Toronto, Ontario, Canada. [6]Institute of Biomaterials and Biomedical Engineering, University of Toronto, Toronto, Ontario, Canada. [7]These authors contributed equally: Barbara Mair, Peter M. Aldridge.
*e-mail: j.moffat@utoronto.ca; shana.kelley@utoronto.ca

sorting, as well as high sorting efficiency and sufficient accuracy to separate phenotypically distinct populations.

As a proof-of-concept target of interest, we chose to screen for positive and negative regulators of CD47 cell surface display. CD47 is widely expressed across cell types and acts as a 'don't eat me' signal through inhibitory interactions with SIRPα, a protein expressed on macrophages and other myeloid cells that negatively regulates phagocytosis[14]. CD47 is highly expressed on various types of tumour cells, and blocking the interaction between CD47 and SIRPα has been explored as a strategy for cancer immunotherapy that has shown promising initial results for some cancer types[15–17]. MICS enabled the processing of a whole genome-wide screen with more than 10^8 live cells in less than 1 h, greatly exceeding the throughput that is feasible using FACS. We identified the glutaminyl cyclase QPCTL as a modifier of CD47, an enzyme that catalyses the cyclization of N-terminal glutamate and glutamine to pyrogluta-mate (pyro-Glu). We validated this interaction using both genetic and chemical perturbations of QPCTL, and developed a highly sensitive parallel reaction monitoring–mass spectrometry (PRM–MS) assay for direct and quantitative measurement of the pyro-Glu modification at the endogenous CD47 N-terminus. The work presented here demonstrates that scalable high-throughput cell sorting can be merged with genome-scale phenotypic CRISPR screening to elucidate bona fide modifiers of a marker of interest.

Results

To identify positive and negative regulators of a biomarker of interest, we designed a microfluidic chip that enables the collection of three subpopulations: a bulk (medium) population expressing baseline levels of the target biomarker, and two populations expressing either higher or lower levels (Fig. 1a). Cell sorting is facilitated by immunomagnetic labelling using antibodies coupled to magnetic particles. Magnetically labelled cells are directed to their respective outlets using ferromagnetic guides made of Metglas 2714A, a cobalt-based magnetic alloy, to induce a highly precise magnetic force that balances the drag force and buffer flow. We maximized throughput by modifying the width and height of the fluidic channel and the ratio of sample-to-buffer volume. The MICS chip contains two sets of deflection guides that are angled at 5° and 20° relative to the direction of flow to achieve an approximate 10%:80%:10% distribution of target cells in the low:medium:high outlets (Fig. 1b). At an optimized flow rate (Supplementary Fig. 1a), the chip enables processing of 3×10^7 cells h^{-1} per chip at sorting efficiencies of 73–92% (Supplementary Fig. 1b). In an arrayed setup using 30 parallel chips, for example, this device can achieve sorting capacities of close to 1 billion cells h^{-1} (Fig. 1c).

To benchmark the performance of MICS, we chose an antibody that targets the SIRPα binding site on CD47 (CC2C6)[18,19], a biologically relevant site that modulates inhibitory interactions with macrophages and is being explored as a strategy for cancer immunotherapy at present[15]. We transduced HAP1 cells—a near-haploid mammalian cell line widely used for functional genetic screens[20,21] that robustly expresses CD47[22,23]—with Cas9 and single-guide RNAs (sgRNAs) targeting CD47, and processed the mutants using MICS. In parallel, we performed FACS-based cell sorting. Detection of CD47^{low} cells was comparable between both methods, and recovery (~80%) and cell viability (~90%) after MICS were high, allowing for a secondary sort for further enrichment of CD47^{low} cells (Fig. 1d,e). Next, we mixed CD47 wild-type and knockout cells in defined ratios and again observed accurate recovery by flow cytometry and MICS (Fig. 1f).

HAP1 cells are small (~10 µm in diameter) and express CD47 at high levels (Supplementary Fig. 2a–c,e,f), thus providing a robust system for cell sorting based on protein expression. However, MICS can also be applied to larger cell types, such as LNCaP and PC3 (Supplementary Fig. 2a), and intracellular and surface markers expressed at lower density, such as EpCAM and vimentin (Supplementary Fig. 2b,c,e,f). At an optimized flow rate, MICS performs comparably to FACS for both markers in both cell lines (Supplementary Fig. 2d,g, Supplementary Table 5). These results demonstrate the flexibility of MICS as a customizable cell-sorting approach.

Having established MICS as a reliable sorting strategy, we then mutagenized HAP1 cells using the Toronto KnockOut (TKOv3) CRISPR library v.3.0[24–26] for a genome-scale proof-of-concept screen. The mutant pool was propagated for approximately 12 doublings and was then processed for 2 rounds of microfluidic sorting with recovery and expansion before re-sorting of rare mutants with altered CD47 levels (Fig. 2a). The abundance of sgRNAs in each of the sampled populations was determined by deep sequencing and compared between the enriched and unsorted cell populations to identify candidate regulators of CD47 expression. In total, we processed 3×10^8 target cells on parallelized MICS chips driven by multiple syringe pumps. This arrayed setup enabled bulk sample preparation, which added no extra processing time with increasing cell numbers and resulted in a net sorting time of around 1 h (Supplementary Table 1). For comparison, we performed a FACS-based screen using the same mutagenized cell pool, achieving—at best—sorting rates of 8×10^7 cells h^{-1}, which is only possible when using small resilient non-aggregating cell lines (Supplementary Table 1). We also compared MICS with magnetic-activated cell sorting (MACS; Supplementary Table 1), but this approach suffered from poor recovery of viable cells, and detected enrichment of CD47 sgRNAs in only one of three replicates (Supplementary Table 1). We did not pursue extensive optimization of this approach, leaving open the possibility that it could be made effective, but MACS is probably not amenable to separating cells with subtly different phenotypes.

We calculated normalized Z scores[27] for the enriched sgRNAs in the CD47^{high} (Supplementary Tables 2 and 3) and CD47^{low} populations (Fig. 2b,c, Supplementary Tables 3 and 4). CD47 was detected as a strong hit in the CD47^{low} population by both methods (Fig. 2b,c, Supplementary Tables 3 and 4). Notably, the three effective CD47 sgRNAs (Fig. 1d) were enriched in the CD47^{low} population in both MICS and FACS, whereas the ineffective sgRNA (4) was not (Supplementary Tables 2 and 3). In addition to CD47, four other top-ranked hits (<30% false-discovery rate (FDR)) overlapped between the FACS and MICS CD47^{low} screens, of which QPCTL showed the largest effect size (Fig. 2c,d), which is in agreement with two recent reports of genetic screens in different cell lines using the same antibody[28,29].

QPCTL encodes the glutaminyl-peptide cyclotransferase-like protein (also known as isoQC), a putative Golgi-resident enzyme and paralogue of the secreted glutaminyl-peptide cyclotransferase (QC, encoded by QPCT). Both enzymes catalyse the formation of N-terminal pyro-Glu through cyclization of glutamine and glutamate residues[30,31]. Interestingly, an N-terminal pyro-Glu has been detected by crystallographic analysis of the CD47 protein and was suggested to mediate the interaction with SIRPα[32], but was assumed to arise spontaneously[33].

We confirmed reduced levels of CD47 pyro-Glu modification (CD47$^{pyro-Glu}$)—as determined by CC2C6 antibody binding—after transduction of sgRNAs targeting QPCTL by cell surface and intracellular flow cytometry, as well as immunofluorescence imaging in HAP1, HEK293T and KMS11 cell lines (Fig. 3a–c, Supplementary Figs. 3 and 4). Loss of QPCT did not affect CC2C6 binding to CD47$^{pyro-Glu}$, and QPCT and QPCTL double targeting did not further reduce the levels of CD47$^{pyro-Glu}$ (Fig. 3a,b). Using a different CD47 antibody (B6H12), we confirmed that overall CD47 protein expression and cell surface localization were not decreased after inactivation of QPCTL (Fig. 3b,c, Supplementary Fig. 4). To establish that the enzymatic activity of QPCTL is required for the observed loss of CC2C6 binding, we then tested two small-molecule inhibitors of QPCTL—SEN177[34] and PQ912[35–37]—that are currently in
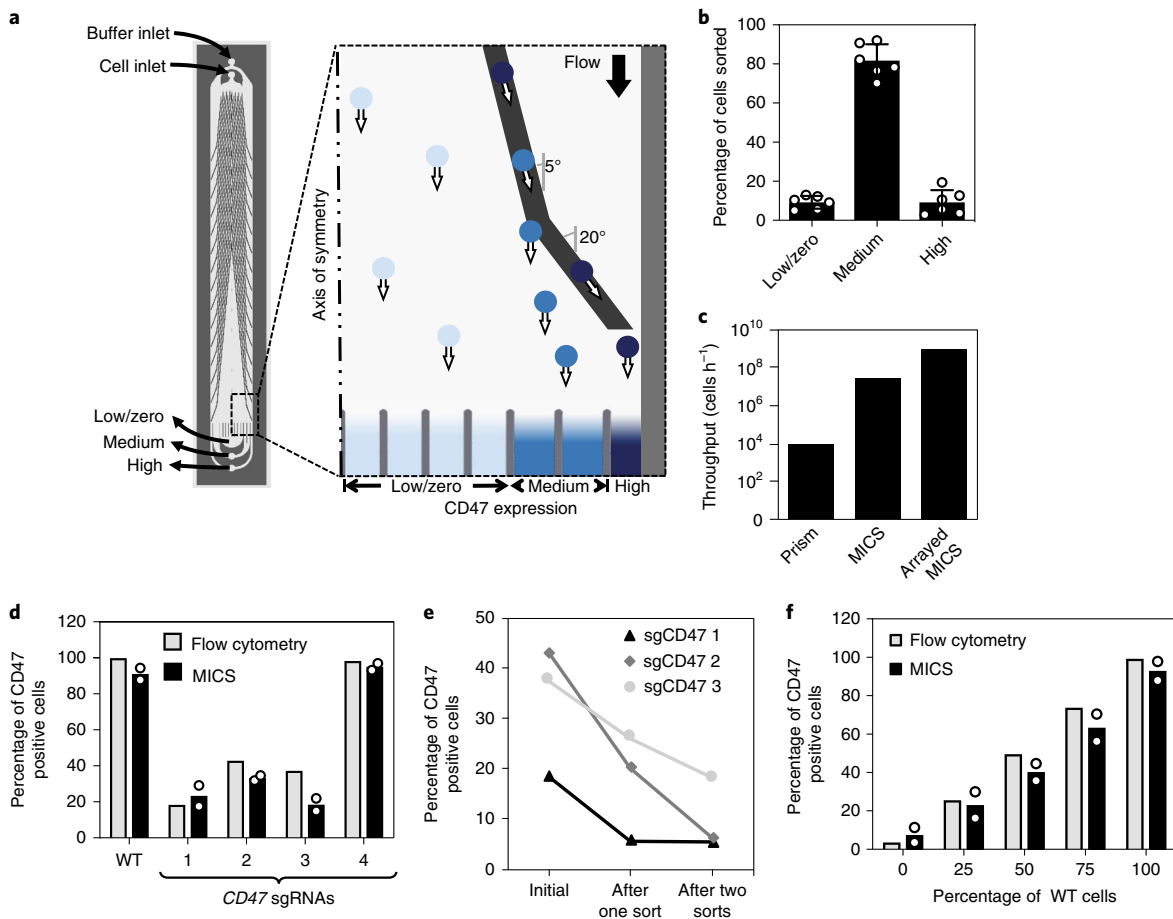
Fig. 1 | A microfluidic chip for high-throughput cell sorting. a, The MICS chip contains two sets of ferromagnetic deflection guides angled at 5° and 20° relative to the direction of flow. A neodymium magnet placed underneath the chip generates a near-uniform magnetic field, and the ferromagnetic guides generate local field amplifications. Magnetically labelled cells, which are subjected to magnetic and drag forces, follow the guides provided that the component of drag force acting perpendicular to the guides does not exceed the magnetic trapping force. **b**, Outlet profile of HAP1 cells sorted with the MICS chip, labelled with magnetic beads targeted to CD47 at a sample flow rate of 6 ml h⁻¹. Data are mean ± s.d. of $n = 3$ biological replicates and $n = 2$ technical replicates. **c**, Throughput of microfluidic chip sorting. Numbers are calculated on the basis of observed throughput with a Prism chip[13] and MICS for 1 (MICS) and 30 (arrayed MICS) chips. **d**, Comparison of the performance of MICS and flow cytometry. sgRNA 4 does not generate a knockout of CD47. MICS data are mean ± s.d. of $n = 2$ technical replicates. Flow cytometry was performed on an aliquot of the same cell pool at 14 d after transduction. CD47 positive are the combined CD47^med and CD47^high populations. WT, wild type. **e**, Sequential sorting of cells transduced with sgRNAs targeting CD47. CD47^low cells were expanded after the primary sort, re-sorted 6 d later, expanded and analysed by flow cytometry. **f**, Sorting defined mixtures of wild-type and two-sort-purified pooled CD47-knockout (sgRNA 1) cells. Data are mean ± s.d. of $n = 2$ technical replicates. Flow cytometry data for the same mixtures are shown for comparison.

clinical development for neurological diseases that involve pyro-Glu-mediated protein aggregation. As expected, we observed decreased CC2C6 binding, but not decreased B6H12 binding, after inhibitor treatment (Fig. 3b,d,e, Supplementary Figs. 4 and 5).

Thus far, we relied on indirect detection of the pyro-Glu modification using the CC2C6 antibody, which is assumed to bind specifically to CD47^pyro-Glu (refs. [18,19,28,29]). To directly prove that the loss of expression or inhibition of QPCTL causes a reduction in pyro-Glu levels on the endogenous CD47 N-terminus, we established a targeted quantitative MS assay for this particular peptide (Fig. 4a,b, Supplementary Figs. 6 and 7). Notably, we report the specific detection and direct quantification of endogenous pyro-Glu as a post-translational modification.

We hypothesized that QPCTL (and QPCT) act on more than the few reported substrates[38–40]. We performed an in silico prediction of mature human protein sequences, assuming that the major mechanism of exposing non-Met N-terminal amino acids is signal peptide cleavage. This approach revealed around 600 candidates with

predicted glutamine or glutamate residues at their N terminus (Fig. 4d). As expected, most candidates are either secreted or membrane proteins, and some have been annotated with pyro-Glu modification. Other than CD47, the list of high-confidence candidates contains the exocrine tissue-associated prolactin-inducible protein (PIP), as well as many chemokines and immunoglobulins including the known QPCTL substrates CCL2 and CX3CL1[39,40]. Another high-confidence candidate is the angiogenic ribonuclease angiogenin, the crystal structure of which confirmed the presence of an N-terminal pyro-Glu[41]. Finally, we also identified signalling proteins such as Frizzled-7 (*FZD7*) and G-protein-coupled receptor C6A (*GPRC6A*) among potential QPCTL targets, warranting further investigation of the role of QPCTL in post-translational protein modification.

Discussion

Here we demonstrate that MICS is a robust method for functional genetic screening. An arrayed setup of MICS chips can surpass the throughput of traditional FACS-based immunosorting technologies
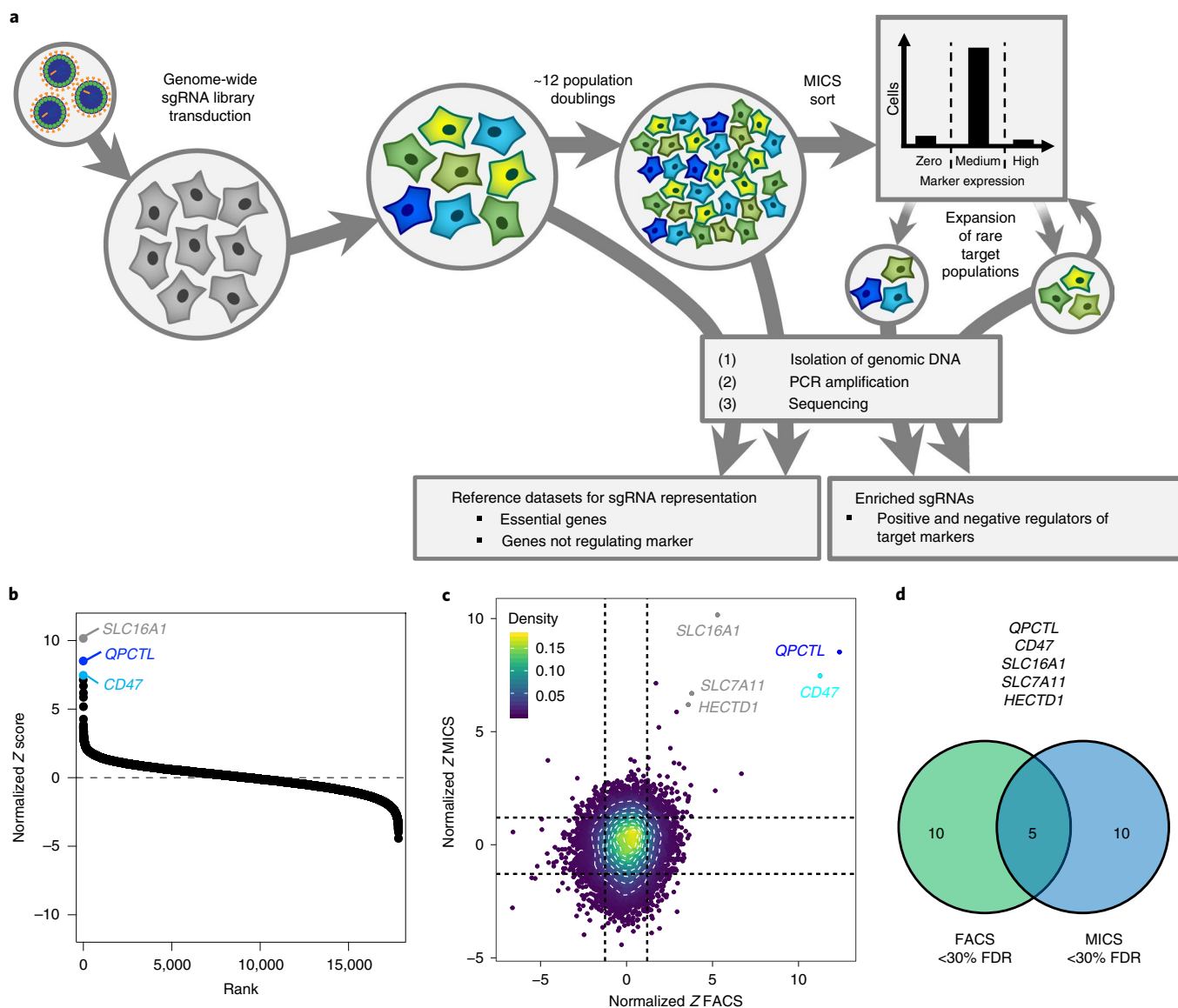
Fig. 2 | FACS-free MICS–CRISPR screen identifies QPCTL as a modifier of CD47. a, Schematic of CRISPR screen workflow with two rounds of MICS sorting. **b**, Genes targeted by sgRNAs detected in the MICS CD47^{low} screen, ranked by normalized Z score as calculated using the drugZ[27] algorithm. **c**, Comparison of MICS and FACS screen results for CD47^{low}. Normalized Z scores are displayed, coloured by density with contours indicated. The dashed lines indicate the top and bottom 90th percentile of the data. Overlapping hits at <30% FDR are highlighted. **d**, Overlap of hits from the MICS and FACS CD47^{low} screens at <30% FDR.

by an order of magnitude, which will expand the applicability of functional phenotypic screening, including to fragile cell types.

We have demonstrated that MICS can be used with different types of cell that vary in morphology, size and tissue of origin. Furthermore, we show live cell sorting using antibodies targeting different surface proteins that are expressed at varying levels across cell lines. Importantly, MICS is capable of faithfully sorting cells that display a wide range of expression densities of the marker of interest, that is, with large or small windows between the high, medium and low populations (Supplementary Fig. 2). MICS can also be used on fixed cells, enabling greater flexibility with respect to experimental timing and, more importantly, the use of intracellular markers (Supplementary Fig. 2), further expanding the range of potential applications.

In theory, our MICS approach is also suitable for identifying negative regulators of a biomarker of interest, such as in the CD47^{high} population of our screen (Supplementary Table 2), or negative regulators of markers with low expression levels. Identifying negative regulators of CD47^{high} would be of high interest—for instance, to modulate immunogenicity in regenerative medicine settings[42]. However, our screen was limited in identifying true negative modulators of CD47 levels owing to the use of the CC2C6 antibody and possibly the already high endogenous levels of CD47 in HAP1 cells.

Performing MICS with a new antibody or cell line requires some degree of optimization. In addition to labelling optimization, which is equally necessary for any antibody-based method, the MICS optimization steps mainly centre around the flow rate. This differs from FACS; however, depending on the cell and sorting type, some degree of parameter optimization is usually also required for FACS, making the workload comparable between both methods. In MICS, the flow rate modulates the throughput and, as the competing drag and magnetic forces acting on a cell dictate its
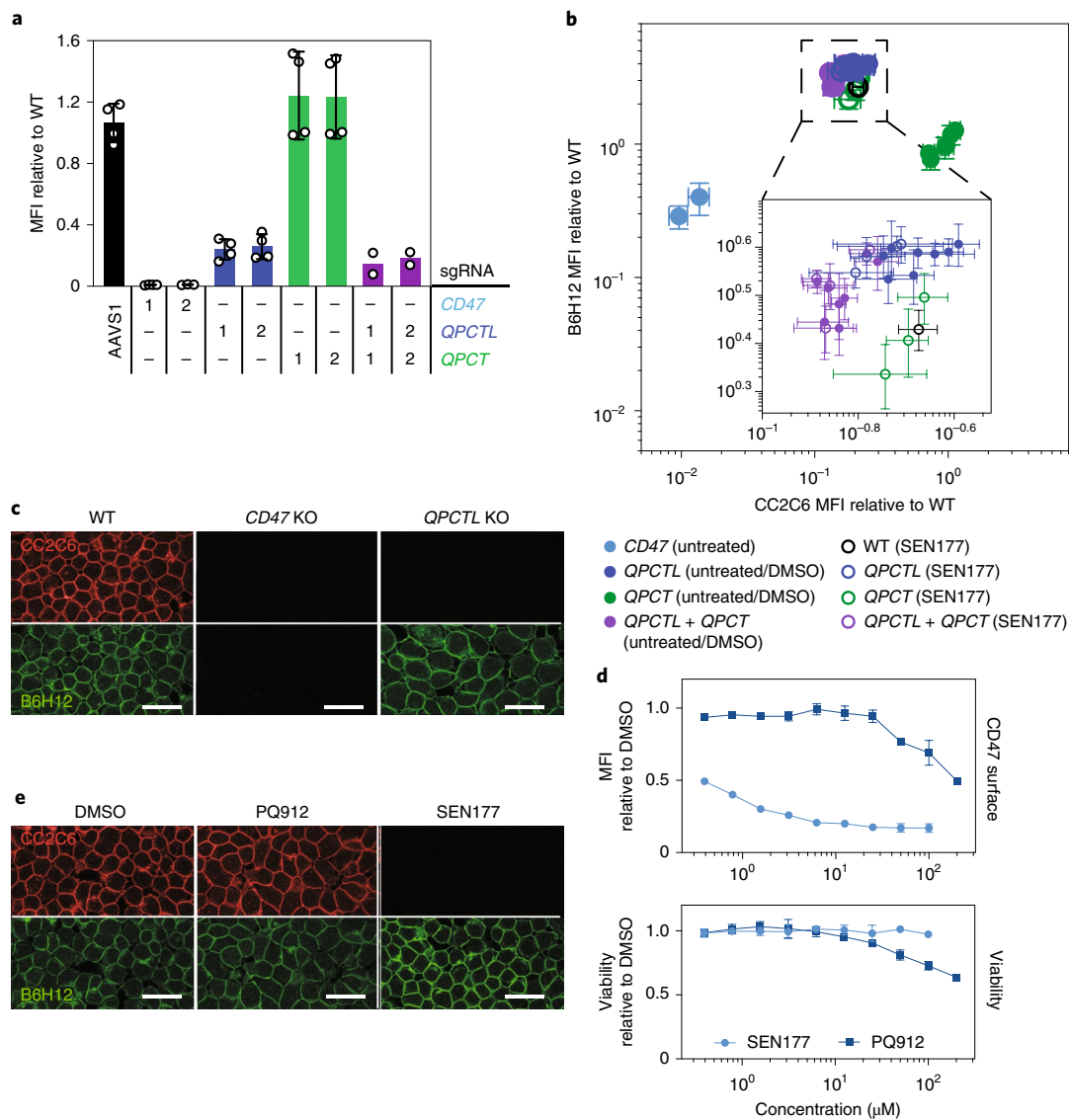
Fig. 3 | QPCTL regulates CD47[pyro-Glu] formation. a, Cell-surface flow cytometry (CC2C6 antibody) of HAP1-Cas9 cells transduced with sgRNAs targeting *CD47*, *QPCTL* and *QPCT* as indicated, 5–8 d after transduction. *AAVS1*-targeting sgRNA was used as a control. Data are mean ± s.d. of the relative median fluorescence intensity (MFI, to wild type) of $n \geq 2$ biological replicates from independent experiments. The double knockout is indicated in purple. **b**, Cell-surface flow cytometry (CC2C6 and B6H12 antibodies) of HAP1 single-cell knockout clones for *CD47*, *QPCT*, *QPCTL*, and *QPCT* + *QPCTL*. Cells were either left untreated or treated with SEN177 (25 μM) or DMSO for 48 h. Data representation is as in **a**; $n = 3$ biological replicates from independent experiments, relative to wild type (untreated) or wild type + DMSO (treated). Inset: close-up of *QPCTL* cluster. Individual data points represent independent single-cell-derived knockout lines ($n = 4$ for *QPCTL*, $n = 3$ for *QPCT*, $n = 4$ for *QPCTL* + *QPCT* and $n = 2$ for *CD47*) for each condition (untreated, DMSO and SEN177). **c**, Immunofluorescence staining (CC2C6, top (red); B6H12, bottom (green)) and confocal imaging of HAP1 single-cell knockout clones for *CD47* and *QPCTL*. **d**, Cell-surface flow cytometry of HAP1 cells and viability after SEN177 or PQ912 treatment for 72 h. Data are mean ± s.d. of MFI or viability relative to DMSO-treated cells from $n \geq 2$ biological replicates from independent experiments. **e**, Immunofluorescence images, as described in **c**, for SEN177- and PQ912-treated (both at 10 μM for 72 h) HAP1 cells. For **c** and **e**, scale bars, 20 μm.

outlet position, flow rate should be chosen carefully to achieve the desired fractionation. Altering the flow rate will not only influence throughput, but also the population distributions in the three outlets, analogous to the 'sort gates' in FACS (Supplementary Figs. 1a and 2). Owing to the fixed geometry of the MICS device, a certain desired population distribution may not be achievable. For example, adjusting the distribution to 20%:60%:20% in the three outlets (Supplementary Fig. 1a) would require a slower flow rate for the low outlet, but a higher rate for the high bin. In cases in which the marker of interest and/or the cell line of choice do not allow sorting in the desired outlet distribution, chip design adjustments can be

incorporated into the manufacturing workflow to accommodate such particular settings.

The flow rate adjustment is the only gating step in the MICS approach, making the method less flexible in this respect compared with FACS, as sequential gating on cell size and granularity, or live/dead exclusion and doublet discrimination is not directly possible. Similarly, in its current configuration, MICS is less amenable to multiplexing (that is, surveying multiple markers in a single run) compared with FACS. However, some degree of multiplexing would be conceivable despite having only a single phenotype (magnetic load) instead of multiple fluorophores. For example, one could imagine
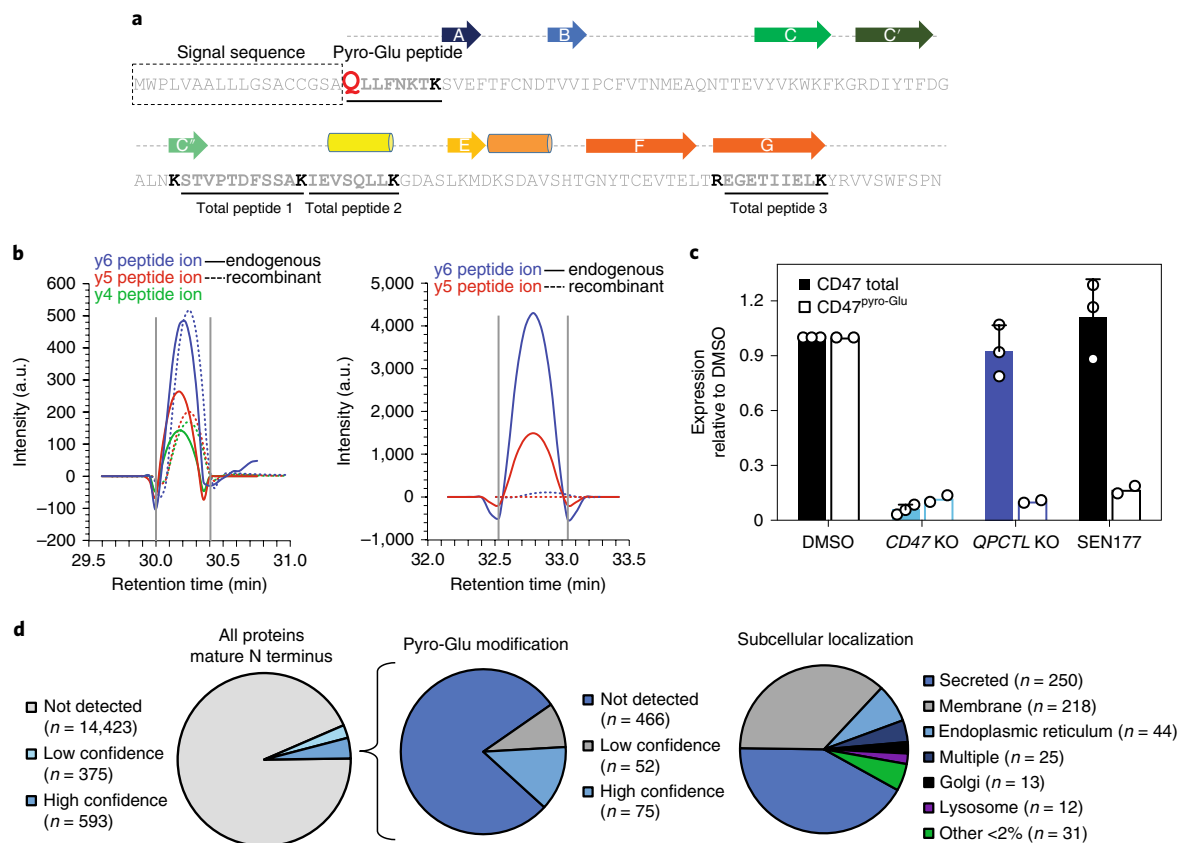
Fig. 4 | Direct detection of CD47^pyro-Glu by MS. a, Schematic of the secondary structure and the primary sequence of the IgSF domain of human CD47; cylinders indicate α-helices and arrows indicate β-sheets. The four CD47 tryptic peptides used for PRM assays are underlined. Lysine and arginine residues in bold indicate the tryptic peptide boundaries. The signal-peptide sequence is indicated by the dashed box. The N-terminal glutamine residue (substrate for pyro-Glu cyclization) is highlighted in red. **b**, Representative overlay of PRM transition chromatograms for total (peptide 3) and pyro-Glu peptides from **a**. The peptides are derived from the extracellular domain of purified recombinant human CD47 (dashed lines) and endogenous CD47 from HAP1 cells (solid lines). The colours indicate different peptide ions. **c**, Quantification of PRM–MS for endogenous CD47 expression and pyro-Glu modification quantification in HAP1 cells. Data are mean ± s.d. of two or three technical replicates. **d**, In silico analysis of potential QPCTL target proteins. The N-terminal amino acid residue of the protein was determined using signal-peptide prediction algorithms. Proteins with high-confidence prediction of N-terminal Glu and Gln residues were analysed for subcellular localization and previous annotation with pyro-Glu modifications. See Methods. Fig. 4a reproduced from ref. [32], Elsevier.

pooling multiple magnetic bead-coupled antibodies and sorting cells on the basis of overall magnetic load to rapidly enrich for 'all positive' or 'all negative' cells. A secondary deconvolution sort could subsequently be performed by FACS on the enriched populations. An alternative for multiplexed magnetic sorting could be a modified cellular indexing of transcriptomes and epitopes (CITE)-seq setup[43] with antibodies that, in addition to a magnetic label, carry a DNA barcode for a sequenceable readout. Multiplexing in similar systems has also been demonstrated by exploiting the differences in magnetic force that can be applied to cellular targets bound with magnetic particles of varying sizes[44] or by sequential sorting and labelling using aptamers[45].

Taken together, we believe that MICS can bridge the gap between high-flexibility high-control systems such as FACS, which sort cells at a lower throughput, and inflexible binary high-throughput systems like MACS. In MICS, compromising on the flexibility and single-cell data acquisition achievable by FACS comes at the benefit of considerably increased sorting capacity and more flexible, non-binary gating compared with MACS.

In practical terms, core facilities of major research institutions are typically well equipped with FACS sorters and experienced operators. However, their availability—especially for long sort durations as required for genome-scale screens—might become limiting

due to other users requiring access and sorts inevitably extending outside regular working hours. As instruments and operators are usually billed per hour, the costs of running such screens on facility-operated FACS sorters are considerable, particularly when phenotypic screening is an integral part of a research group's workflow. Although our MICS technology is not available as a plug-and-play device at present, we are working towards making it more user-friendly and accessible. As a first step, we provide a stepwise protocol that details all of the necessary materials, reagents and procedures required for manufacturing and operating MICS chips (Supplementary Information; the protocol has also been deposited in the Protocol Exchange[46]). This should enable a user with microfabrication experience to produce and operate the MICS chips used in this study.

We applied MICS to a genome-wide CRISPR–Cas9 loss-of-function screen, probing genetic regulators of CD47 and yielding overlapping hits with a parallel FACS-based screen. Most notably, we identified and validated QPCTL as a regulator of CD47^pyro-Glu, corroborating recent findings that QPCTL is a potential modifier of CD47-targeted cancer immunotherapy[28,29]. Supporting this role, high expression of *QPCTL* has been shown to be a poor prognostic indicator for patients with renal cancer and leukaemia[29,47]. More detailed in vivo studies will be required to establish whether the

modulation of QPCTL activity is an actionable strategy to boost macrophage-based cancer immunotherapy efficiency.

Interestingly, our screens also identified *SLC16A1* and *SLC7A11*, which have been implicated in free pyro-Glu metabolism as transporters of pyro-Glu itself[18] or glutamine, respectively. Furthermore, our in silico prediction suggests that QPCTL could have additional substrate proteins, such as angiogenin, in which an N-terminal pyro-Glu has been detected[41]. As for CD47, this modification has been assumed to occur spontaneously, and has been implicated in modulating the catalytic activity and cytotoxicity of angiogenin, as well as other RNase A homologues with N-terminal pyro-Glu[49–51]. This, as well as other examples, suggests that the substrates of QPCTL (and QPCT) could be involved in widespread biological, physiological and pathological processes, warranting further investigation of these understudied enzymes.

In summary, MICS is a robust, flexible, parallelizable and customizable high-throughput cell-sorting method that is particularly suitable for genome-scale screening applications.

Methods

MICS device fabrication. Strips of Metglas 2714A were obtained from Metglas and epoxy bonded (Loctite M-31CL, McMaster-Carr) onto 100 mm soda lime glass wafers (550 μm thick, University Wafer) and left to cure for 24 h. Excess epoxy was removed with acetone. The metallic surface was then primed with MCC 80/20 (Microchem) before spin coating with S1811 positive photoresist (Microchem). The positive resist was photolithographically patterned, and then the exposed Metglas 2714A was etched using a mixture of 3.6% HCl (Sigma), 14.3% H_2O_2 (Sigma) and 82.1% H_2O. After stripping the remaining photoresist and priming the surface with OmniCoat to improve adhesion, the ferromagnetic guides were encapsulated with a layer of SU-8 3010 (Microchem), and then microfluidic channel features were patterned with a 100 μm layer of SU-8 3050 (Microchem). Each chip was then capped with cured PDMS with holes cored for all of the inlet and outlet ports, following an APTES treatment (Sigma)[52]. To minimize friction and limit cell adhesion to the chip surfaces, every chip was treated with a solution of 1% w/v pluronic F108 (BASF) in deionized H_2O for a minimum of 12 h as described previously[53].

HAP1 cells. HAP1 cells were obtained from Horizon (clone C631; sex, male with lost Y chromosome; RRID: CVCL_Y019) and HAP1-Cas9 cells were generated as described previously[24]. For screening, HAP1 cells were cultured in minimal DMEM without sodium pyruvate, with 3,700 mg l⁻¹ sodium bicarbonate, 1.982 g l⁻¹ glucose and 0.161 g l⁻¹ L-glutamine (Wisent Bioproducts) with 10% FBS (Gibco) and 1% penicillin–streptomycin (ThermoFisher). For all of the other experiments, cells were cultured in standard medium (IMDM (Gibco) supplemented with 10% FBS (Gibco) and 1% penicillin–streptomycin (ThermoFisher)). The different medium conditions do not alter CD47 levels or modification (Supplementary Fig. 4a). Cells were cultured at 37 °C and 5% CO_2 in humidified incubators, were free of mycoplasma and routinely tested using the MycoAlert Detection Kit (Lonza).

Other cell lines. HEK293T cells (CRL-3216; sex, female; RRID: CVCL_0063) were obtained from ATCC and maintained in DMEM (Gibco) with high glucose, L-glutamine and sodium pyruvate, supplemented with 10% FBS (Gibco) and 1% penicillin–streptomycin (ThermoFisher). KMS11 (sex, female; RRID: CVCL_2989) were a gift from the Toronto Recombinant Antibody Centre (University of Toronto) and were maintained in RPMI 1640 with L-glutamine (Gibco), supplemented with 10% FBS (Gibco) and 1% penicillin–streptomycin (ThermoFisher). LNCaP cells were obtained from ATCC (clone FGC, CRL-1740; sex, male; RRID: CVCL_1379) and were cultured in RPMI 1640 with L-glutamine (Gibco), supplemented with 10% FBS (Gibco) and 1% penicillin–streptomycin (ThermoFisher). PC3 cells were obtained from ATCC (CRL-1435; sex, male; RRID: CVCL_0035) and cultured in F-12K medium (ATCC), supplemented with 10% FBS (Gibco) and 1% penicillin–streptomycin (ThermoFisher). Cells were cultured at 37 °C and 5% CO_2 in humidified incubators, were authenticated at regular intervals by STR profiling at the Centre for Applied Genomics (TCAG) of the Hospital for Sick Children (SickKids) in Toronto, were free of mycoplasma and were routinely tested using the MycoAlert Detection Kit (Lonza).

CRISPR sgRNA lentivirus production. The pLCKO-TKOv3 plasmid library lentivirus was produced as previously described[26]. In brief, HEK293T cells were seeded at a density of 9×10^6 cells per 15 cm plate and incubated overnight, after which cells were transfected with a mixture of psPAX2 (4.8 μg; Addgene, 12260), pMDG.2 (3.2 μg; Addgene, 12259), TKOv3 plasmid library (8 μg) and X-tremeGENE 9 (48 μl; Roche) in Opti-MEM (Gibco). Then, 24 h after transfection, the medium was changed to DMEM with 1% BSA (Sigma) and 1% penicillin–streptomycin (Gibco). Virus-containing medium was collected 48 h

after transfection, centrifuged at 1,500 r.p.m. for 5 min and stored at −80 °C. Functional titres were determined by virus titration on HAP1 cells. Subsequently, 24 h after infection, the medium was replaced with puromycin-containing medium (1 μg ml⁻¹) and the cells were incubated for 48 h. The multiplicity of infection (MOI) was determined 72 h after infection by comparing survival of infected cells with infected unselected and uninfected selected control cells. Lentivirus for individual sgRNA constructs was produced on a smaller scale; HEK293T cells were seeded at a density of 0.5×10^6 per 6-well in low-antibiotic growth medium (DMEM with 10% FBS (Gibco), 0.1% penicillin–streptomycin) and incubated overnight. Cells were transfected with a mixture of psPAX2 (1,800 ng), pMDG.2 (200 ng), sgRNAs in pLCKO or pLCV2[24,25] and X-tremeGENE 9 (12 μl) in Opti-MEM. Then, 24 h after transfection, the medium was changed to serum-free high-BSA growth medium as above. Virus-containing medium was collected 48 h after transfection, centrifuged at 1,500 r.p.m. for 5 min and stored at −80 °C. Functional titres in cells for validation experiments were determined by virus titration.

Generation of *CD47*, *QPCT* and *QPCTL* knockout cells. For transduction experiments using cell pools, single-stranded sgRNA oligos were annealed using T4 polynucleotide kinase in T4 ligation buffer (both NEB) and ligated into digested (BsmBI; NEB), phosphatase-treated (rSAP or CIP; NEB) and gel-purified modified pLCKO (for HAP1-Cas9) or pLCV2 (for HAP1, HEK293T and KMS11) backbones[24,25] using T4 DNA ligase. All of the plasmids were verified by Sanger sequencing and virus was prepared and titred as described above. HAP1, HAP1-Cas9, HEK293T or KMS11 cells were infected with target and control gRNAs (MOI < 1) in the presence of 8 μg ml⁻¹ polybrene. Then, 24 h after infection, the medium was replaced with fresh medium containing puromycin (1 μg ml⁻¹ for HAP1 and HAP1-Cas9, 1–2 μg ml⁻¹ for HEK293T and 0.5 μg ml⁻¹ for KMS11) and cells were incubated for 48 h. Cells were further cultured in selection-free medium as indicated for individual experiments (typically T_3–T_6) and were passaged every 3–4 d. CD47 levels or modification did not increase or decrease with prolonged passaging up to T_{12} in HAP1-Cas9 cells (Supplementary Fig. 4a). For infections with two sgRNAs (HAP1-Cas9 only), the second sgRNA was cloned into pLCKO hygro (Moffat laboratory). Selection was carried out in medium containing puromycin (1 μg ml⁻¹) and hygromycin (800 ug ml⁻¹) for the first 48 h, then in hygromycin-only medium for a further 4–5 d. For single-cell knockout clones (HAP1 only), sgRNAs were cloned into modified PX459v2.0 (Moffat laboratory, Addgene, 62988, 1 kb stuffer sequence added). HAP1 cells were transfected using Lipofectamine 3000 (3.75 μl and 2.3 μg plasmid per 6-well plate), selected with puromycin for 48 h as above and seeded in limiting dilutions (1 cell per 100 μl per well of a 96-well plate) 3 d later. Single-cell colonies were expanded, cryo-banked (in standard medium with 10% DMSO and a total of 20% FBS) and analysed for mutations as described below. An sgRNA targeting the *AAVS1* locus was used as negative control. The sgRNA sequences are provided in Supplementary Table 6.

CRISPR editing analysis. For analysing indels after CRISPR-mediated editing, genomic DNA was isolated using the QIAamp DNA Blood Mini kit (Qiagen) or the Extracta DNA Prep kit (Quantabio). sgRNA target regions (100–200 bp upstream and 500 bp downstream of sgRNA sequence) were amplified by touchdown PCR (−0.5 °C per cycle from 72 °C to 60 °C plus 10 additional cycles at 59 °C) from 50–100 ng input DNA, PCR products were purified using the PureLink Quick PCR purification kit (Invitrogen) if necessary and analysed by Sanger sequencing. Sequencing data were analysed to determine the percentage of edited sequences contained in the sample using TIDE[54] and a KO-score using ICE[55] (Synthego). Non-edited cells amplified with the same primers were used as control samples. A representative analysis of double-sgRNA infected samples from two independent experiments yielded the following editing results for *QPCT* and *QPCTL* target loci (by sgRNA; TIDE mean percentage edited ± s.d./ICE KO-score ± s.d. for respective on-target locus): QPCTL_sg1, 58.2 ± 16.2%/NA; QPCTL_sg2, 33.23 ± 32.96%/47 ± 0%; QPCT_sg1, 33.15 ± 36.51%/56 ± 0%; QPCT_sg2, 80.2 ± 19.85%/76.4 ± 2.42%; QPCT_sg3, 68.18 ± 4.29%/36.33 ± 2.08%. For assessing background or off-target rates, cells transduced with sgRNAs targeting AAVS1 and CD47 were used (mean ± s.d. over all *QPCT* and *QPCTL* target loci), 14.4 ± 14.95%/1.33 ± 0.58%. For single-cell-derived clones, TIDE and ICE were used for screening purposes, and mutations were independently confirmed by PCR, Sanger sequencing and sequence alignment as described. The primers are provided in Supplementary Table 7.

MICS sorting. HAP1 cells were detached using 0.125% trypsin, washed once in PBS with 10% FBS and resuspended at a concentration of 1×10^7 cells ml⁻¹ in a solution of Hank's balanced salt solution (HBSS) supplemented with 2% BSA. Cells were labelled for CD47 expression with biotin anti-human CD47 antibody (clone CC2C6; BioLegend, 323104; RRID: AB_756134) at 0.0625 μg per 100 μl. Excess antibody was removed by washing twice in HBSS with 2% BSA. Cells were resuspended at 1×10^8 cells ml⁻¹, and anti-Biotin MicroBeads UltraPure were added at a 20% concentration by volume (Miltenyi Biotec, 130-105-637) and incubated at room temperature for 30 min. LNCaP cells were trypsinized as described above, washed in PBS and resuspended in HBSS with 2% BSA at a concentration of $4–6 \times 10^6$ cells ml⁻¹. Cells were labelled for EpCAM expression with anti-EpCAM microbeads (Milteny Biotec, 130-061-101), added at a 20% concentration by

volume and incubated at room temperature for 30 min. PC3 cells were trypsinized as described above, washed and resuspended in 100 µl of PBS ($4-6 \times 10^6$ cells ml^{-1}). Cells were fixed and permeabilized by adding 90% (v/v) cold methanol to the cell pellet and incubating on ice for 15 min. Permeabilized cells were centrifuged, washed twice in PBS with 2% BSA, then labelled using a biotin-conjugated vimentin antibody (polyclonal; R&D Systems, BAF2105; RRID: AB_2288538) at a concentration of 4 µg ml^{-1} for 30 min at 4 °C. Following three washes in HBSS with 2% BSA, cells were incubated with anti-Biotin MicroBeads UltraPure (Miltenyi Biotec, 130-105-637) as described above.

HAP1 cells were sorted by MICS at a concentration of around 5×10^6 cells ml^{-1} in HBSS with 2% BSA and 5 mM EDSS; the initial cell concentration was measured using a Countess Automated Cell Counter (Invitrogen). Other cell lines were sorted at 5×10^4 cells ml^{-1} in HBSS with 2% BSA and 5 mM EDTA. Syringe pumps (Fusion 200, Chemyx), operating in withdrawal mode, were used to drive flow in the MICS chips. Custom 3D-printed mounting hardware enabled up to five stacks of three syringes (containing 20 ml, 10 ml and 3 ml syringes; Becton Dickinson) to be driven by the same pump. The different cross-sectional areas of the syringes were used to generate different flow rates, corresponding to the width of the low, medium and high outlet channels. Two inlet reservoirs, one containing the sample (cell) solution and one containing a flow focusing buffer stream (HBSS with 2% BSA and 5 mM EDSS or EDTA), were connected to the two inlets of the MICS chip. The pump flow rate was chosen such that the sample flow rate was 6 ml h^{-1} (with a total flow rate of 12 ml h^{-1}) unless indicated otherwise. The sorted samples were collected in their respective syringes, and the volume of solution collected in each syringe was measured by weight. A small fraction (100 µl) of each sample was collected for cell counting to determine cell concentration after sorting. Each sample was stained with 1 µl of Syto24 Green nucleic acid stain (ThermoFisher Scientific) and incubated for 15 min at room temperature. Each sample was then loaded onto a 10-chambered microscope slide (Quick Read 3805, Globe Scientific) and cells were counted under fluorescence excitation using a custom counting macro (Supplementary Information) and a Nikon TI Eclipse microscope. For Supplementary Fig. 2, accurate counting of cells at the outlet of the chip was performed by adding 1 µl of Syto24 Green nucleic acid stain (ThermoFisher Scientific) per million cells during incubation with microbeads and automated counting on a Nikon Eclipse Ti fluorescence microscope. The sort efficiency was calculated by dividing the number of cells collected in the medium and high outlets by the total number of cells collected (low, medium and high). Recovery efficiency was defined as the percentage of input cells that were recovered in collected outlet populations. For the screen, the fractions collected from the low/zero (21%, 24% and 30% of sorted cells for replicates A, B and C, respectively) and high outlets (10%, 18% and 15%) were collected in 15 ml falcon tubes on ice. Cells were then centrifuged, plated in minimal medium and cultured for 6 d before secondary sorting (for the CD47^{low} fraction from primary sort: 57%, 50% and 35% in low/zero; 1%, 1%, 2% in high). Sorted cells (low/zero and high) were pelleted for genomic DNA extraction. See Supplementary Table 1 for input cell numbers, recovery and throughput.

Pooled genome-wide CRISPR screens in HAP1 cells. CRISPR screens in stable HAP1-Cas9 cells were performed essentially as previously described[24,25]. In brief, 150×10^6 cells were infected with the TKOv3 lentiviral library[26] at an MOI of around 0.3 (more than 400-fold coverage of the library after selection with puromycin). Medium was changed 24 h after infection to puromycin-containing medium (1 µg ml^{-1}). Then, 72 h after infection, 100×10^6 puromycin-selected cells were cryo-banked, 90×10^6 cells were split into three replicates of 30×10^6 cells, passaged every 3–4 d and maintained at 400-fold coverage. Subsequently, 30×10^6 cells were collected for genomic DNA extraction at T_0 after selection and at every passage until day 12 after selection (T_{12}), when sorting was performed. The unsorted T_{12} sample was used as a reference. Genomic DNA extraction, library preparation and sequencing were performed as described below. For the FACS and MACS screens, 90×10^6 cryo-banked T_0 cells were taken in culture, cultured until T_{12} as above and sorted as described below. The unsorted T_{12} sample was used as reference. All of the cell populations tested negative for mycoplasma before and after sorting.

FACS and MACS sorting. Cells were detached using 0.125% trypsin, counted and $2 \times 30 \times 10^6$ aliquots were pelleted for genomic DNA extraction. The remaining population was split in half for staining and sorting by FACS and MACS ($90–100 \times 10^6$ cells per replicate). For FACS sorting, cells were washed once in PBS and once in flow buffer (PBS with 2% BSA) and stained with anti-human CD47-APC antibodies (clone CC2C6; BioLegend, 323123/4; RRID: AB_2716202, AB_2716203) in flow buffer (20 µl antibody per 40×10^6 cells ml^{-1}) for 1 h rotating at 4 °C in the dark. Cells were washed three times with sort buffer (PBS with 1 mM EDTA, 25 mM HEPES pH 7 and 1% BSA), resuspended in sort buffer at 40×10^6 cells ml^{-1}, filtered through a 40 µm sieve and stained with 7-aminoactinomycin D (7-AAD; BioLegend; 50 µl per 40×10^6 cells ml^{-1}). Small aliquots were taken as single-stain controls and stained as above. Sorting was performed using a BD FACS Aria IIIu: 4 laser (405/488/561/633) 15 parameter cuvette-based sorter with injection of approximately 40×10^6 cells h^{-1}. The top and bottom 15% (gated on a CD47 histogram of viable 7-AAD-negative cells) were collected in 15 ml falcon tubes with

minimal medium supplemented with 50% FCS on ice. Cells were then centrifuged, resuspended in minimal medium, plated and cultured for 6 d. CD47^{low} and CD47^{high} cells were then detached, counted, stained and sorted again as described above (approximately $90–120 \times 10^6$ cells per replicate per fraction). Sorting gates were set using unsorted mutagenized cells that had been cultured in parallel. For MACS sorting, a MACS LS column mounted on a MidiMACS (Miltenyi Biotec) was used for sorting. Cells were labelled with magnetic nanobeads targeted to CD47 as described above for MICS, suspended in HBSS supplemented with 2% BSA and sorted through the column. The negative fraction was collected in 15 ml falcon tubes with minimal medium supplemented with 50% FCS on ice. Cells were then centrifuged, resuspended in minimal medium, plated and cultured for 7–21 d. Only one replicate could be processed for a secondary sort owing to poor recovery and cell viability. Cells were pelleted after the secondary sort and gDNA extraction, library preparation, sequencing and data analysis were performed as described above. All of the cell populations tested negative for mycoplasma before and after sorting. Input cell numbers, recovery and throughput are provided in Supplementary Table 1.

Genomic DNA extraction and Illumina sequencing. Genomic DNA was extracted from screen cell pellets using the Wizard Genomic DNA Purification kit (Promega). Sequencing libraries were prepared by amplifying sgRNA inserts by a two-step PCR reaction using primers that include Illumina TruSeq adapters with i5 and i7 indices. The resulting libraries were subsequently sequenced on an Illumina HiSeq2500 (RRID: SCR_016383) as previously described[25]. Each read was completed with standard primers for dual indexing with Rapid Run V1 reagents. The first 20 cycles of sequencing were dark cycles or base additions without imaging. The actual 26 bp read begins after the dark cycles and contains two index reads, reading the i7 first, followed by i5 sequences. The primers used for sequencing are provided in Supplementary Table 8.

Screen data processing and quality control. Sample reads were trimmed by locating the first 8 bp of the anchors used in the barcoding primers and extracting the flanking 20 bp after the anchor was found. We allowed for a 2 bp mismatch for the anchor search. After trimming, a quality-control alignment was performed using Bowtie v.0.12.8 (allowing for maximum of 2 mismatches, ignoring qualities). For each sample, all of the available reads were combined from different sequencing runs if applicable, aligned using Bowtie as described above and sgRNAs were tallied. Read counts for all of the samples in a screen were combined in a matrix (the percentage of recovered sgRNAs in each sample is provided below; recovered is defined as ≥1 raw read) and normalized by dividing each read count by the sum of all of the read counts in the sample and then multiplying by the expected read number (10 million). Fold change was calculated to a reference sample (T_{12} unsorted). The calculated fold changes were then used to generate normalized Z scores using drugZ (v.1.1.0.2)[27].

MICS: T_0 99.4%, T_{12} A_unsorted 98%, T_{12} B_unsorted 98.2%, T_{12} C_unsorted 86.3%, A_CD47^{high} 2.9%, B_CD47^{high} 2.2%, C_CD47^{high} 8.6%, A_CD47^{low} 86.1%, B_CD47^{low} 71.3, C_CD47^{low} 7.1%.

FACS: T_0 99.4%, T_{12} A_unsorted 96.5%, T_{12} B_unsorted 97.2%, T_{12} C_unsorted 97.6%, A_CD47^{high} 88.3%, B_CD47^{high} 91.9%, C_CD47^{high} 92.7%, A_CD47^{low} 53%, B_CD47^{low} 92.6%, C_CD47^{low} 93.8%.

MACS: T_0 99.4%, T_{12} A_unsorted 96.5%, T_{12} B_unsorted 97.2%, T_{12} C_unsorted 97.6%, A_CD47$^-$ 1.7%, B_CD47$^-$ 0.9%, C_CD47$^-$ 1.6%.

Flow cytometry. Cells were dissociated with 0.125% trypsin and washed once in flow buffer (PBS with 2% BSA). For cell-surface analysis, antibody staining was carried out in flow buffer for 30 min on ice at 4 °C in the dark. For intracellular CD47, cells were first fixed with 4% PFA (Electron Microscopy Sciences) in PBS for 10 min on ice, followed by three washes and permeabilization using flow buffer with 0.1% Triton X-100 (Sigma) for 5 min at room temperature. For vimentin, cells were fixed and permeabilized by adding 90% (v/v) cold methanol to the cell pellet and incubating on ice for 15 min. Cells were washed twice with flow buffer, followed by staining for at least 30 min on ice at 4 °C in the dark. Stained cells were washed three times with flow buffer and 7-AAD viability dye (3–5 µl, BioLegend) was added before quantification (live cells only). The following antibodies were used for these studies: anti-human CD47-APC (1 µl per 10^6 cells in 100 µl; clone CC2C6; BioLegend, 323123/4; RRID: AB_2716202, AB_2716203), anti-human CD47-FITC (5 µl per 10^6 cells in 100 µl; clone B6H12; eBioscience, 11-0479-4/12; RRID: AB_2043842, AB_2043843), anti-human vimentin-AlexaFluor488 (10 µl per 10^6 cells in 100 µl; polyclonal; R&D Systems, IC8104G; RRID: NA), anti-human EpCAM-FITC (5 µl per 10^6 cells in 100 µl; clone 9C4; BioLegend, 324203; RRID: AB_756077). Stained cells were quantified using an LSRII flow cytometer (BD Biosciences) or an iQue Screener PLUS (IntelliCyt), and data were analysed using FlowJo software (RRID: SCR_008520). MFI was defined as median fluorescence across the population, and was generally displayed relative to wild-type cells in the same experiment. The gating strategy is outlined in Supplementary Fig. 8.

Drug treatments. SEN177 (Sigma) and PQ912 (DC Chemicals) were dissolved in DMSO at a concentration of 50 mM and added to cell medium at the indicated concentrations either during plating or 24 h after plating. The cells were incubated

for 12–72 h as indicated (without refreshing the drugs), and the same volume of DMSO was used as control.

Immunofluorescence microscopy. Cells were seeded onto poly-D-lysine-coated 8-well micro-slides (ibidi). The next day, spent media was removed and the cells were washed twice with PBS and fixed with 100% ice-cold methanol for 10 min at −20 °C. After two more washes with PBS, cells were incubated with a blocking solution (5% FBS with 0.05% Tween-20 in PBS) at room temperature for 1 h. Anti-human CD47-APC (clone CC2C6; BioLegend, 323123/4; RRID: AB_2716202, AB_2716203) and anti-human CD47-FITC (clone B6H12; eBioscience, 11-0479-4/12; RRID: AB_2043842, AB_2043843) primary antibodies were added at 1:50 for B6H12-FITC and 1:250 for CC2C6-APC in antibody dilution solution (1% FBS with 0.05% Tween-20 in PBS) and stained over night at 4 °C. Hoechst 33342 (ThermoFisher Scientific) was added at a 1:5,000 dilution in PBS and cells were stained for 10 min. Finally, cells were washed twice in PBS and imaged using a Leica SP5 700 confocal microscope (Zeiss). For mean fluorescence intensity measurements, Hoechst staining intensity was used for automated image segmentation (ImageJ custom macro; Supplementary Information) for operator independent unbiased selection of primary regions of interest masks (in the green channel) and the mean fluorescence intensity was measured from the APC and/or FITC channel. In total, nine random fields of view were imaged per condition. Mean fluorescence intensity from wild-type untreated HAP1 cells was used to normalize all of the tested conditions to generate a relative CD47 expression metric.

Affinity precipitation and LC–MS/MS sample preparation. To prepare samples for affinity precipitation and liquid chromatography–tandem MS (LC–MS/MS), cells were collected at 80–90% confluency from 10 cm dishes by scraping and the cell pellets were washed twice in ice-cold PBS. For the preparation of total cell lysate, cells were lysed in RIPA buffer (ThermoFisher Scientific; supplemented with HALT protease and phosphatase inhibitors (ThermoFisher Scientific)) by incubation for 1 h at 4 °C with gentle rocking followed by 3 bursts of sonication for 5 s at 10% amplitude. After centrifugation at 13,000 g for 30 min at 4 °C, the supernatant was collected. Total protein concentrations were measured using the BCA Protein Assay kit (ThermoFisher Scientific). For immunoprecipitation of endogenous CD47, either clone B6H12.2 (for total protein; ThermoFisher Scientific, MA5-11895; RRID: AB_11009368) or biotinylated CC2C6 (for pyro-Glu modified protein; BioLegend, 323104; RRID: AB_756134) anti-CD47 antibodies were used with Protein-G Dynabeads or Biotin-binder Dynabeads, respectively, following the manufacturer's instructions (ThermoFisher Scientific). For the quantitative detection of N-terminal CD47$^{pyro-Glu}$ modification, CD47 immunoprecipitates were eluted in 100 mM ammonium bicarbonate (pH 7.5) and digested with trypsin (ThermoFisher Scientific) overnight at 37 °C to generate tryptic peptides. Peptides were desalted using PepClean C18 spin columns following the manufacturer's instructions (ThermoFisher Scientific).

High-performance liquid chromatography. EASY-nLC 1200 (ThermoFisher Scientific) was coupled to the Q-Exactive MS (ThermoFisher Scientific) for peptide separation and detection. An EASY-Spray column (2 µm, 100 Å, 75 µm × 50 cm; ThermoFisher Scientific) was used for compound separation. Mobile phase A (0.1% formic acid in H$_2$O) and mobile phase B (0.1% formic acid in acetonitrile) were used with the following gradient: 0 min, 5% B; 50 min, 35% B; 55 min, 100% B; 60 min, 100% B at a flow rate of 225 nl min^{-1}.

PRM–MS assays. Samples were analysed using a Q-Exactive HF quadrupole Orbitrap mass spectrometer (ThermoFisher Scientific). Six peptides were monitored through a PRM acquisition composed of one MS1 scan followed by six targeted MS/MS scans in high-energy collision dissociation with cycle times of 2.7 s. To generate CD47-specific peptides suitable for the PRM assay, we purchased the purified extracellular domain of human CD47 (G&P Biosciences) and used trypsin digestion as described above to generate tryptic peptide-specific PRM transition profiles. The recombinant purified CD47 extracellular domain was expected to contain very little pyro-Glu modification (arising primarily due to spontaneous conversion). Acquisition was performed in positive ion mode for all seven peptides. Owing to a low number of target peptides, data acquisition was performed in an unscheduled MS/MS assay but retention times were noted.

MS data analysis. PRM data acquired by LC–MS/MS were imported into Skyline for peak extraction and peak area calculation for each peptide. The top three fragments for each ion were used for quantification. For quantitative comparison across samples, DMSO control samples were used to generate a relative expression metric for both total CD47 protein expression and pyro-Glu modification across the multiple conditions tested.

In silico prediction of QPCTL targets. To predict potential candidates for N-terminal pyro-Glu modification, knowledge of the first amino acid of the mature N terminus (that is, Gln or Glu) is required. Many secretory or membrane proteins as well as Golgi- and endoplasmic reticulum-resident proteins contain signal peptides, which are proteolytically removed, revealing the mature N

terminus. As signal-peptide processing is a major (but not the only) proteolytic maturation event, we decided to use signal-peptide prediction tools to derive an approximation of a mature human proteome. FASTA sequences of human proteins were downloaded from UniProt (filters: evidence at protein level and reviewed) and were used as input for SignalP[56] (using SecretSanta[57]), signalHSMM[58] and Phobius[59]. All of the tools were run from within R. The following parameters were used for SignalP: v.4.1, organism 'euk', run_mode 'starter', sensitive TRUE. For signalHSMM, a signal-peptide probability cut-off of >0.45 was used. If a signal peptide was detected, the first residue after the predicted cleavage site was used as the new mature N-terminal amino acid, otherwise the original N terminus was used. Proteins were grouped according to prediction confidence for putative Gln or Glu N termini (by 0 tools = not detected, 1 = low confidence, 2/3 = high confidence). We then sub-classified the high-confidence group according to subcellular localization (downloaded from UniProt followed by manual formatting) and previously detected pyro-Glu modification (downloaded from UniProt, PDB, dbptm at http://dbptm.mbc.nctu.edu.tw/ and TopFIND 3.0[60] at http://clipserve.clip.ubc.ca/topfind/ ntermini; by 0 sources = not detected, 1 = low confidence, 2/3/4 = high confidence). Note that annotation with pyro-Glu in PDB does not necessarily predict direct modification due to the presences of multiple entities in a crystal structure. The 67/86, 53/71, 9/24 and 49/84 pyro-Glu annotations are in the high-confidence Q/E N-terminal group for UniProt, dbptm, PDB and TopFIND, respectively. Localizations occurring at less than 2% of the total were grouped together under other.

Reporting Summary. Further information on research design is available in the Nature Research Reporting Summary linked to this article.

Data availability
The main data supporting the results in this study are available within the paper and the Supplementary Information. Supplementary Tables 2–4 contain raw read counts, normalized read counts and normalized Z scores for all of the screens. Unprocessed sequencing files are available from the corresponding authors on reasonable request.

Code availability
The ImageJ custom macro used for automated image segmentation is provided in the Supplementary Information.

References
1. Sharma, S. & Petsalaki, E. Application of CRISPR-Cas9 based genome-wide screening approaches to study cellular signalling mechanisms. *Int. J. Mol. Sci.* **19**, 933 (2018).
2. Burr, M. L. et al. CMTM6 maintains the expression of PD-L1 and regulates anti-tumour immunity. *Nature* **549**, 101–105 (2017).
3. Mezzadra, R. et al. Identification of CMTM6 and CMTM4 as PD-L1 protein regulators. *Nature* **549**, 106–110 (2017).
4. Binek, A. et al. Flow cytometry has a significant impact on the cellular metabolome. *J. Proteome Res.* **18**, 169–181 (2019).
5. Llufrio, E. M., Wang, L., Naser, F. J. & Patti, G. J. Sorting cells alters their redox state and cellular metabolome. *Redox Biol.* **16**, 381–387 (2018).
6. Brockmann, M. et al. Genetic wiring maps of single-cell protein states reveal an off-switch for GPCR signalling. *Nature* **546**, 307–311 (2017).
7. Wroblewska, A. et al. Protein barcodes enable high-dimensional single-cell CRISPR screens. *Cell* **175**, 1141–1155 (2018).
8. de Groot, R., Lüthi, J., Lindsay, H., Holtackers, R. & Pelkmans, L. Large-scale image-based profiling of single-cell phenotypes in arrayed CRISPR-Cas9 gene perturbation screens. *Mol. Syst. Biol.* **14**, e8064 (2018).
9. Haney, M. S. et al. Identification of phagocytosis regulators using magnetic genome-wide CRISPR screens. *Nat. Genet.* **50**, 1716–1727 (2018).
10. Parnas, O. et al. A genome-wide CRISPR screen in primary immune cells to dissect regulatory networks. *Cell* **162**, 675–686 (2015).
11. Han, X. et al. CRISPR-Cas9 delivery to hard-to-transfect cells via membrane deformation. *Sci. Adv.* **1**, e1500454 (2015).
12. Han, X. et al. Microfluidic cell deformability assay for rapid and efficient kinase screening with the CRISPR-Cas9 system. *Angew. Chem. Int. Edn* **55**, 8561–8565 (2016).
13. Aldridge, P. M. et al. Prismatic deflection of live tumor cells and cell clusters. *ACS Nano* **12**, 12692–12700 (2018).
14. Matlung, H. L., Szilagyi, K., Barclay, N. A. & van den Berg, T. K. The CD47-SIRPα signaling axis as an innate immune checkpoint in cancer. *Immunol. Rev.* **276**, 145–164 (2017).
15. Weiskopf, K. Cancer immunotherapy targeting the CD47/SIRPα axis. *Eur. J. Cancer* **76**, 100–109 (2017).
16. Advani, R. et al. CD47 blockade by Hu5F9-G4 and rituximab in non-Hodgkin's lymphoma. *N. Engl. J. Med.* **379**, 1711–1721 (2018).

17. Kong, F. et al. CD47: a potential immunotherapy target for eliminating cancer cells. *Clin. Transl. Oncol.* **18**, 1051–1055 (2016).
18. Seiffert, M. et al. Human signal-regulatory protein is expressed on normal, but not on subsets of leukemic myeloid cells and mediates cellular adhesion involving its counterreceptor CD47. *Blood* **94**, 3633–3643 (1999).
19. Leclair, P. et al. CD47-ligation induced cell death in T-acute lymphoblastic leukemia. *Cell Death Dis.* **9**, 544 (2018).
20. Carette, J. E. et al. Ebola virus entry requires the cholesterol transporter Niemann–Pick C1. *Nature* **477**, 340–343 (2011).
21. Bürckstümmer, T. et al. A reversible gene trap collection empowers haploid genetics in human cells. *Nat. Methods* **10**, 965–971 (2013).
22. Lee, S.-E. et al. Proteogenomic analysis to identify missing proteins from haploid cell lines. *Proteomics* **18**, e1700386 (2018).
23. Paulo, J. A. & Gygi, S. P. Isobaric tag-based protein profiling of a nicotine-treated alpha7 nicotinic receptor-null human haploid cell line. *Proteomics* **18**, e1700475 (2018).
24. Hart, T. et al. High-resolution CRISPR screens reveal fitness genes and genotype-specific cancer liabilities. *Cell* **163**, 1515–1526 (2015).
25. Hart, T. et al. Evaluation and design of genome-wide CRISPR/SpCas9 knockout screens. *G3 (Bethesda)* **7**, 2719–2727 (2017).
26. Mair, B. et al. Essential gene profiles for human pluripotent stem cells identify uncharacterized genes and substrate dependencies. *Cell Rep.* **27**, 599–615 (2019).
27. Colic, M. et al. Identifying chemogenetic interactions from CRISPR knockout screens with drugZ. *Genome Med.* **11**, 52 (2019).
28. Logtenberg, M. E. W. et al. Glutaminyl cyclase is an enzymatic modifier of the CD47–SIRPα axis and a target for cancer immunotherapy. *Nat. Med.* **25**, 612–619 (2019).
29. Wu, Z. et al. Identification of glutaminyl cyclase isoenzyme isoQC as a regulator of SIRPα-CD47 axis. *Cell Res.* **29**, 502–505 (2019).
30. Cynis, H. et al. Isolation of an isoenzyme of human glutaminyl cyclase: retention in the Golgi complex suggests involvement in the protein maturation machinery. *J. Mol. Biol.* **379**, 966–980 (2008).
31. Stephan, A. et al. Mammalian glutaminyl cyclases and their isoenzymes have identical enzymatic characteristics. *FEBS J.* **276**, 6522–6536 (2009).
32. Hatherley, D. et al. Paired receptor specificity explained by structures of signal regulatory proteins alone and complexed with CD47. *Mol. Cell* **31**, 266–277 (2008).
33. Ho, C. C. M. et al. "Velcro" engineering of high affinity CD47 ectodomain as signal regulatory protein α (SIRPα) antagonists that enhance antibody-dependent cellular phagocytosis. *J. Biol. Chem.* **290**, 12650–12663 (2015).
34. Pozzi, C., Di Pisa, F., Benvenuti, M. & Mangani, S. The structure of the human glutaminyl cyclase-SEN177 complex indicates routes for developing new potent inhibitors as possible agents for the treatment of neurological disorders. *J. Biol. Inorg. Chem.* **23**, 1219–1226 (2018).
35. Ramsbeck, D. et al. Structure-activity relationships of benzimidazole-based glutaminyl cyclase inhibitors featuring a heteroaryl scaffold. *J. Med. Chem.* **56**, 6613–6625 (2013).
36. Lues, I. et al. A phase 1 study to evaluate the safety and pharmacokinetics of PQ912, a glutaminyl cyclase inhibitor, in healthy subjects. *Alzheimers Dement.* **1**, 182–195 (2015).
37. Hoffmann, T. et al. Glutaminyl cyclase inhibitor PQ912 improves cognition in mouse models of Alzheimer's disease—studies on relation to effective target occupancy. *J. Pharmacol. Exp. Ther.* **362**, 119–130 (2017).
38. Kumar, A. & Bachhawat, A. K. Pyroglutamic acid: throwing light on a lightly studied metabolite. *Curr. Sci.* **102**, 288–297 (2012).
39. Kehlen, A. et al. N-terminal pyroglutamate formation in CX3CL1 is essential for its full biologic activity. *Biosci. Rep.* **37**, BSR20170712 (2017).
40. Cynis, H. et al. The isoenzyme of glutaminyl cyclase is an important regulator of monocyte infiltration under inflammatory conditions. *EMBO Mol. Med.* **3**, 545–558 (2011).
41. Leonidas, D. D. et al. Refined crystal structures of native human angiogenin and two active site variants: implications for the unique functional properties of an enzyme involved in neovascularisation during tumour growth. *J. Mol. Biol.* **285**, 1209–1233 (1999).
42. Deuse, T. et al. Hypoimmunogenic derivatives of induced pluripotent stem cells evade immune rejection in fully immunocompetent allogeneic recipients. *Nat. Biotechnol.* **37**, 252–258 (2019).
43. Stoeckius, M. et al. Simultaneous epitope and transcriptome measurement in single cells. *Nat. Methods* **14**, 865–868 (2017).
44. Adams, J. D., Kim, U. & Soh, H. T. Multitarget magnetic activated cell sorter. *Proc. Natl Acad. Sci. USA* **105**, 18165–18170 (2008).
45. Labib, M. et al. Aptamer and antisense-mediated two-dimensional isolation of specific cancer cell subpopulations. *J. Am. Chem. Soc.* **138**, 2476–2479 (2016).
46. Philpott, D. et al. High-throughput microfluidic cell sorting platform (MICS). *Prot. Exch.* https://doi.org/10.21203/rs.2.10282/v1 (2019).
47. Uhlen, M. et al. A pathology atlas of the human cancer transcriptome. *Science* **357**, eaan2507 (2017).
48. Sasaki, S., Futagi, Y., Kobayashi, M., Ogura, J. & Iseki, K. Functional characterization of 5-oxoproline transport via SLC16A1/MCT1. *J. Biol. Chem.* **290**, 2303–2311 (2015).
49. Boix, E. et al. Role of the N terminus in RNase A homologues: differences in catalytic activity, ribonuclease inhibitor interaction and cytotoxicity. *J. Mol. Biol.* **257**, 992–1007 (1996).
50. Liao, Y.-D. et al. The structural integrity exerted by N-terminal pyroglutamate is crucial for the cytotoxicity of frog ribonuclease from *Rana pipiens*. *Nucleic Acids Res.* **31**, 5247–5255 (2003).
51. La Mendola, D. et al. Copper binding to naturally occurring, lactam form of angiogenin differs from that to recombinant protein, affecting their activity. *Metallomics* **8**, 118–124 (2016).
52. Ren, Y. et al. A simple and reliable PDMS and SU-8 irreversible bonding method and its application on a microfluidic-MEA device for neuroscience research. *Micromachines* **6**, 1923–1934 (2015).
53. Luk, V. N., Mo, G. C. & Wheeler, A. R. Pluronic additives: a solution to sticky problems in digital microfluidics. *Langmuir* **24**, 6382–6389 (2008).
54. Brinkman, E. K., Chen, T., Amendola, M. & van Steensel, B. Easy quantitative assessment of genome editing by sequence trace decomposition. *Nucleic Acids Res.* **42**, e168 (2014).
55. Hsiau, T. et al. Inference of CRISPR edits from Sanger trace data. Preprint at *bioRxiv* https://doi.org/10.1101/251082 (2019).
56. Nielsen, H. in *Protein Function Prediction: Methods in Molecular Biology* Vol. 1611 (ed. Kihara, D.) 59–73 (Springer, 2017).
57. Gogleva, A., Drost, H.-G. & Schornack, S. SecretSanta: flexible pipelines for functional secretome prediction. *Bioinformatics* **34**, 2295–2296 (2018).
58. Burdukiewicz, M., Sobczyk, P., Chilimoniuk, J., Gagat, P. & Mackiewicz, P. Prediction of signal peptides in proteins from malaria parasites. *Int. J. Mol. Sci.* **19**, 3709 (2018).
59. Käll, L., Krogh, A. & Sonnhammer, E. L. L. A combined transmembrane topology and signal peptide prediction method. *J. Mol. Biol.* **338**, 1027–1036 (2004).
60. Fortelny, N., Yang, S., Pavlidis, P., Lange, P. F. & Overall, C. M. Proteome TopFIND 3.0 with TopFINDer and PathFINDer: database and analysis tools for the association of protein termini to pre- and post-translational events. *Nucleic Acids Res.* **43**, D290–D297 (2015).

Acknowledgements

Author contributions

B.M. and P.M.A. performed most of the experiments and analysed data with help from R.S.A., D.P., M.L. and S.N.M. M.Z. and R.S.A. developed the MS assay. A.H.Y.T. helped with screen sequencing and data analysis. B.M., P.M.A., R.S.A., J.M. and S.O.K. wrote the manuscript. B.M., P.M.A., E.H.S., S.A., J.M. and S.O.K. designed the study. S.A., J.M. and S.O.K. supervised the study.

Competing interests

The authors declare no competing interests.

Additional information

nature research

Corresponding author(s): Shana O. Kelley, Jason Moffat

Last updated by author(s): May 3, 2019

Reporting Summary

Nature Research wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Research policies, see Authors & Referees and the Editorial Policy Checklist.

Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

n/a	Confirmed	
☐	☒	The exact sample size (*n*) for each experimental group/condition, given as a discrete number and unit of measurement
☐	☒	A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
☒	☐	The statistical test(s) used AND whether they are one- or two-sided *Only common tests should be described solely by name; describe more complex techniques in the Methods section.*
☒	☐	A description of all covariates tested
☒	☐	A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
☐	☒	A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
☒	☐	For null hypothesis testing, the test statistic (e.g. *F*, *t*, *r*) with confidence intervals, effect sizes, degrees of freedom and *P* value noted *Give P values as exact values whenever suitable.*
☒	☐	For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
☒	☐	For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
☒	☐	Estimates of effect sizes (e.g. Cohen's *d*, Pearson's *r*), indicating how they were calculated

Our web collection on statistics for biologists contains articles on many of the points above.

Software and code

Policy information about availability of computer code

Data collection	Chip design: AutoCAD 2015, SolidWorks 2016 Sorting: Nikon Elements IF: Zeiss ZEN Black MS: ThermoScientific Xcalibur 4.0 Flow cytometry: FACSDIVA and ForeCyt (see below)
Data analysis	Screen data analysis: BOWTIE v0.12.8, drugZ v1.1.0.2 (see methods and references for details) IF: FIJI 2.0.0 (ImageJ v1.52) custom macro, available upon request Editing analysis: TIDE v2.0.1 and ICE v1.2 (see methods and references) MS: Skyline 4.2 Signal peptide prediction: SignalP 4.0 (via SecretSanta5), signalHSMM and Phobius (see methods and references) General: Excel 2013 v15.0.5127.1000, Excel 2016 v16.16.9, GraphPad Prism v8.0.2 (159) and v8.1.2 (227) R: v3.4.4, relevant packages used: viridis_0.5.1, seqinr_3.4-5, devtools_2.0.1, SecretSanta_0.99.0, ggrepel_0.8.0, stringr_1.4.0, MASS_7.3-51.3, dplyr_0.7.8, ggplot2_3.1.1, ragp _0.3.0.0001, signalHsmm_1.5

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors/reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Research guidelines for submitting code & software for further information.

Data

Policy information about availability of data

All manuscripts must include a data availability statement. This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A list of figures that have associated raw data
- A description of any restrictions on data availability

The main data supporting the results in this study are available within the paper and its Supplementary Information. Supplementary Tables 2–4 contain raw read counts, normalized read counts and normZ scores for all screens. Unprocessed sequencing files are available on request.

Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

☒ Life sciences ☐ Behavioural & social sciences ☐ Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see nature.com/documents/nr-reporting-summary-flat.pdf

Life sciences study design

All studies must disclose on these points even when the disclosure is negative.

Sample size	No sample-size calculations were performed, and sample sizes were arbitrarily chosen according to conventions in the field. For small-scale experiments, the number of replicates exceeds at least 2 biological replicates (that is, independent experiments) and/or at least 3 technical replicates (that is, repeated measurements of the same original sample). For screens, the initial mutagenized cell pool was split into 3 replicates post-selection, and processed independently in all downstream steps.
Data exclusions	No data were excluded from any experiments and from any of the figures, except for Supplementary Fig. 2d, where only one representative experiment is shown (other experiments were performed with different internal controls, and showed the same trend).
Replication	All technical chip optimization and QPCTL validation experiments were reproducible. All replicates are shown (except for individual chips when technical replicates failed for technical reasons such as bubbles in the device). For validation of additional screen hits, results were inconsistent and had smaller effect sizes and were therefore not included in the manuscript.
Randomization	No animal or human subjects were involved in this study, and therefore randomization was not performed.
Blinding	No sample allocation to groups was performed, as blinding was not relevant to this study.

Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

Materials & experimental systems

n/a	Involved in the study
☐	☒ Antibodies
☐	☒ Eukaryotic cell lines
☒	☐ Palaeontology
☒	☐ Animals and other organisms
☒	☐ Human research participants
☒	☐ Clinical data

Methods

n/a	Involved in the study
☒	☐ ChIP-seq
☐	☒ Flow cytometry
☒	☐ MRI-based neuroimaging

Antibodies

Antibodies used	Dilutions and conditions of use are provided in Methods. – biotinylated anti-human CD47 antibody (clone CC2C6, BioLegend, Cat.# 323104, RRID: AB_756134, multiple lots) – anti-human CD47-FITC (clone B6H12, eBioscience, Cat. # 11-0479-4/12, RRID: AB_2043842/3, multiple lots) – anti-human CD47-APC (clone CC2C6, BioLegend, Cat. #323123/4, RRID: AB_2716202/3, multiple lots) – anti-human CD47 (clone B6H12.2, ThermoFisher Scientific, Cat. #MA5-11895, RRID: AB_11009368, lot #UA2704833) – anti-human vimentin-AlexaFluor488 (polyclonal, R&D Systems, Cat. #IC8104G, RRID: NA, lot #ADRK0113051) – biotinylated anti-human vimentin (polyclonal, R&D Systems, Cat. #BAF2105, RRID: AB_2288538, lot #CBFJ0815121)

	– anti-human EpCAM MicroBeads (clone proprietory, Miltenyi Biotec, Cat. #130-061-101, RRID: NA, lot #5170630346) – anti-human EpCAM-FITC (clone 9C4, BioLegend, Cat. #324203, RRID: AB_756077, lot #B229481)
Validation	– Biotinylated and APC-conjugated anti-human CD47 antibodies (clone CC2C6) are validated for flow cytometry by the supplier, for N-terminal CD47 binding by Seiffert et al., Blood (1999) (and additional publications) and for flow cytometry, immunoprecipitation, magnetic immunolabeling and immunofluorescence in this study through use of control and KO cells. – Anti-human CD47-FITC (clone B6H12) is validated for flow cytometry by the supplier (and additional publications) and in this study for flow cytometry and immunofluorescence through use of control and KO cells. – Anti-human CD47 (clone B6H12.2) is validated for immunofluorescence and immunoprecipitation by the supplier (and additional publications) and in this study through use of control and KO cells. – Anti-human vimentin-AlexaFluor488 (polyclonal) is validated for immunocytochemistry by the supplier (flow cytometry in this study). – Biotinylated anti-human vimentin (polyclonal) is validated for Western Blot by the supplier and for immunocytochemistry by Lee et al., Theranostics (2018) (magnetic immunolabeling in this study). – Anti-human EpCAM MicroBeads are validated for magnetic immunolabeling by the supplier (and additional publications) and in this study through siRNA-mediated knock-down. – Anti-human EpCAM-FITC (clone 9C4) is validated for flow cytometry, immunofluorescence and immunohistochemistry by the supplier (and additional publications) and for flow cytometry in this study through siRNA-mediated knock-down.

Eukaryotic cell lines

Policy information about cell lines

Cell line source(s)	HAP1 cells were obtained from Horizon (clone C631, sex: male with lost Y chromosome, RRID: CVCL_Y019) and HAP1-Cas9 cells were generated as described in Hart el al., Cell (2015). Generation of CD47, QPCT and QPCTL KO clones is described in Methods. HEK293T (CRL-3216, sex: female, RRID: CVCL_0063) were obtained from ATCC . KMS11 (sex: female, RRID: CVCL_2989) were a gift from the Toronto Recombinant Antibody Centre (TRAC; University of Toronto). LNCaP cells were obtained from ATCC (clone FGC, CRL-1740, sex: male, RRID: CVCL_1379). PC3 cells were obtained from ATCC (CRL-1435, sex: male, RRID: CVCL_0035).
Authentication	All cell lines were authenticated by STR profiling at the Centre for Applied Genomics (TCAG) at the Hospital for Sick Children (SickKids) in Toronto. HAP1 cells were also whole-genome sequenced.
Mycoplasma contamination	All cell lines were routinely tested and confirmed negative for mycoplasma contamination.
Commonly misidentified lines (See ICLAC register)	None of the cell lines used in this study is listed as commonly misidentified.

Flow Cytometry

Plots

Confirm that:

☒ The axis labels state the marker and fluorochrome used (e.g. CD4-FITC).

☒ The axis scales are clearly visible. Include numbers along axes only for bottom left plot of group (a 'group' is an analysis of identical markers).

☒ All plots are contour plots with outliers or pseudocolor plots.

☒ A numerical value for number of cells or percentage (with statistics) is provided.

Methodology

Sample preparation	Cell lines (HAP1, HEK293, KMS11, LnCAP, R22v1, PC3) were dissociated with 0.125% trypsin and washed once in once in Flow buffer (PBS with 2% BSA) or PBS. For cell-surface analysis, antibody staining was carried out in Flow buffer for 30 min on ice at 4 degC in the dark. For intracellular CD47, cells were first fixed with 4% PFA (Electron Microscopy Sciences) in PBS for 10 min on ice, followed by permeabilization using Flow buffer with 0.1% TritonX-100 (Sigma) for 5 min at room temperature. For vimentin, cells were fixed and permeablized by adding 90% (v/v) cold methanol to the cell pellet and incubating on ice for 15 min. Cells were washed twice with Flow buffer, followed by staining for 30min on ice at 4 degC in the dark. Stained cells were washed thrice with Flow buffer and 7-AAD viability dye (3-5µl, BioLegend) was added before quantification (live cells only).
Instrument	LSR II: 3 laser (488/640/405) configuration (BD Biosciences) or iQue Screener PLUS Blue-Red configuration (IntelliCyt)
Software	Data were acquired by using FACSDIVA software v8.0-8.1 (BD Biosciences) or ForeCyt software v6.2.6752 (IntelliCyt) with automated compensation (performed independently for each experiment with single-stained samples following the software assistant) and analyzed by using FlowJo software (RRID:SCR_008520) v10.4.-v10.6.0 (TreeStar).
Cell population abundance	For the FACS screen, gates were set such that post-sort populations comprised approximately the top and bottom 15% of the CD47-APC histograms (see Supplementary Fig. 8). Purity was checked by measuring the sorted fractions using the same settings, and confirmed the expected shift of the median in all 'bottom' and 'top' fractions.

Gating strategy | The gating strategy is described in Supplementary Fig. 8.

☒ Tick this box to confirm that a figure exemplifying the gating strategy is provided in the Supplementary Information.