nature biotechnology

LETTERS

https://doi.org/10.1038/s41587-020-0466-7

High-throughput single-cell activity-based screening and sequencing of antibodies using droplet microfluidics

Annabelle Gérard[1,18], Adam Woolfe[1,17], Guillaume Mottet[2,17], Marcel Reichen[1,17], Carlos Castrillon[2,3,4,17], Vera Menrath[1,17], Sami Ellouze[1,17], Adeline Poitou[1,17], Raphaël Doineau[1,3,4,11,17], Luis Briseno-Roa[1,12], Pablo Canales-Herrerias[2,3,5], Pascaline Mary[6], Gregory Rose[6], Charina Ortega[6], Matthieu Delincé[6], Sosthene Essono[6], Bin Jia[7,13], Bruno Iannascoli[2], Odile Richard-Le Goff[2], Roshan Kumar[6], Samantha N. Stewart[6], Yannick Pousse[1], Bingqing Shen[1], Kevin Grosselin[1,4], Baptiste Saudemont[2,4], Antoine Sautel-Caillé[4], Alexei Godina[4], Scott McNamara[1], Klaus Eyer[5,14], Gaël A. Millot[8], Jean Baudry[5], Patrick England[9,10], Clément Nizak[4], Allan Jensen[1,15,18], Andrew D. Griffiths[4,18] ✉, Pierre Bruhns[2,18] ✉ and Colin Brenan[1,6,16,18] ✉

Mining the antibody repertoire of plasma cells and plasmablasts could enable the discovery of useful antibodies for therapeutic or research purposes[1]. We present a method for high-throughput, single-cell screening of IgG-secreting primary cells to characterize antibody binding to soluble and membrane-bound antigens. Celli*GO* is a droplet microfluidics system that combines high-throughput screening for IgG activity, using fluorescence-based in-droplet single-cell bioassays[2], with sequencing of paired antibody V genes, using in-droplet single-cell barcoded reverse transcription. We analyzed IgG repertoire diversity, clonal expansion and somatic hypermutation in cells from mice immunized with a vaccine target, a multifunctional enzyme or a membrane-bound cancer target. Immunization with these antigens yielded 100–1,000 IgG sequences per mouse. We generated 77 recombinant antibodies from the identified sequences and found that 93% recognized the soluble antigen and 14% the membrane antigen. The platform also allowed recovery of ~450–900 IgG sequences from ~2,200 IgG-secreting activated human memory B cells, activated ex vivo, demonstrating its versatility.

The immune system generates a vast repertoire of antibodies in response to infection or immunization that can potentially be explored for diagnostic, therapeutic or research applications. IgGs comprise the most abundant class of immunoglobulin in circulation. They protect from infections by blocking molecular interactions, and by inducing complement-dependent cytotoxicity and antibody-dependent cellular cytotoxicity. IgGs are matured during ongoing immune responses, and through this process improve their specificity and affinity for antigen. Their production can also be rapidly recalled following previously elicited protective immunity in response to pathogen or immunogen exposure[3]. The phenotypic diversity of the target-specific IgG repertoire in an individual underlies protection following vaccination or infection[4]. IgGs are secreted by circulating antibody-secreting cells (plasmablasts) and short-lived or long-lived tissue-resident antibody-secreting cells (plasma cells). Analysis of the activity and sequence of IgGs secreted by plasmablasts and plasma cells is hence of great interest both to investigate their pivotal role in the humoral immune response and to obtain antibodies for therapeutic or other applications[5].

Until recently, target-specific IgG discovery and repertoire analysis was based on B-cell immortalization using hybridoma technology or Epstein–Barr virus transformation, which are relatively inefficient, are biased toward dividing plasmablasts rather than fully mature plasma cells, and are limited to specific species. Direct PCR cloning of the V-gene sequences that encode antibody variable (V) regions, from plasma cells has been reported[6,7] but without prior knowledge of antigen specificity. Recently, next generation sequencing (NGS) approaches have been developed to probe the IgG repertoire more deeply[1,8], some even recovering the original pairing of heavy chain and light chain variable regions (V_H and V_L)[9–13]. Antigen-labeled IgG^+ memory B cells and plasmablasts have been isolated using

[1]HiFiBiO Therapeutics SAS, Paris, France. [2]Unit of Antibodies in Therapy and Pathology, Institut Pasteur, UMR1222 INSERM, Paris, France. [3]École Doctorale Frontières du Vivant (FdV), Centre de Recherches Interdisciplinaires, Paris, France. [4]Laboratoire de Biochimie (LBC), ESPCI Paris, PSL Research University, CNRS UMR8231 Chimie Biologie Innovation, Paris, France. [5]Laboratoire Colloïdes et Matériaux Divisés (LCMD), ESPCI Paris, PSL Research University, CNRS UMR8231 Chimie Biologie Innovation, Paris, France. [6]HiFiBiO Therapeutics Inc., Cambridge, MA, USA. [7]Pfizer, Cambridge, MA, USA. [8]Institut Pasteur, Hub Bioinformatique et Biostatistique, C3BI, USR 3756 IP CNRS, Paris, France. [9]Plateforme de Biophysique Moléculaire, Institut Pasteur, CNRS UMR3528, Paris, France. [10]Centre d'Innovation et de Technologie, Plateforme de Biophysique des Macromolecules et de leurs Interactions, Institut Pasteur, Paris, France. [11]Present address: Institute for Integrative Nanosciences, Leibniz IFW Dresden e.V., Dresden, Germany. [12]Present address: Medetia Pharma, Institut Imagine, Paris, France. [13]Present address: Abbvie Bioresearch Center, Worcester, MA, USA. [14]Present address: Laboratory for Functional Immune Repertoire Analysis, Institute of Pharmaceutical Sciences, D-CHAB, ETH Zürich, Zurich, Switzerland. [15]Present address: Lundbeck A/S, Valby, Denmark. [16]Present address: 1CellBiO Inc., Watertown, MA, USA. [17]These authors contributed equally: Adam Woolfe, Guillaume Mottet, Marcel Reichen, Carlos Castrillon, Vera Menrath, Sami Ellouze, Adeline Poitou, Raphaël Doineau. [18]These authors jointly supervised this work: Annabelle Gérard, Allan Jensen, Andrew D. Griffiths, Pierre Bruhns, Colin Brenan. ✉e-mail: andrew.griffiths@espci.fr; pierre.bruhns@pasteur.fr; c.brenan@1cell-bio.com

fluorescence-activated cell sorting (FACS) and have been sequenced[9,10], but antigen-specific plasma cells have not, given that they lack surface IgG expression. Phenotypic analysis of antibody-secreting cells has evolved from ELISpot[14] to high-throughput assays that are based on microwells or nanowells, or on picoliter or nanoliter droplets in microfluidic systems[15–17], but remains non-adapted or reliant on low-throughput manual micromanipulation for subsequent V_H–V_L sequencing. In vitro display technologies, notably phage and yeast display, have proved to be powerful tools to discover therapeutic antibodies[18]; however, screening is based on binding and rarely on function[18], and was, until recently[12,13], unable to preserve the cognate V_H–V_L pairing, limiting its utility for the analysis of immune responses. Yeast display has been shown to maintain cognate V_H–V_L pairing to provide antibodies with higher sensitivity and specificity than random pairing[13]; it involves multiple steps for library construction, antibody fragment display and multiple rounds of screening, and is restricted to soluble antigens[12,13]. There is, therefore, a need for a system that couples high-throughput single-cell phenotypic screening with paired V_H–V_L sequencing of IgG-secreting primary cells for both soluble and membrane-bound antigens in a flexible format that enables direct screening for functional activities[19].

We reported previously a droplet-based microfluidics technology in which up to 16,000 single IgG-secreting cells are compartmentalized in tens of thousands of 40-pl droplets and analyzed in two-dimensional droplet arrays using a fluorescence relocation-based immunoassay to measure IgG secretion rate and affinity[2]. This method, however, allows neither recovery of specific cells based on the properties of secreted IgGs (the phenotype), nor determination of the sequence of the secreted antibody genes (the genotype).

Here, we describe a droplet-based microfluidics technology, Celli*GO*, that enables high-throughput single-cell screening of millions of non-immortalized cells on the basis of the phenotypic properties of secreted IgGs, paired V_H–V_L sequencing of antibodies with the desired properties, bioinformatics analysis and rapid expression, and in vitro antibody characterization (Fig. 1a).

Microfluidic systems allow the generation of picolitre-volume droplets that contain single IgG-secreting cells (Fig. 1b, left), identification of droplets that contain IgGs with the desired activity and sorting of those identified droplets at up to 600 s^{-1} (Fig. 1b, middle). This highly flexible droplet-based system allows IgG-secreting cells to be screened on the basis of a diversity of assays, including binding to soluble antigens, cross-reactivity or specific binding (using multiple soluble antigens), binding to cell-surface targets[17] (on both bacterial and eukaryotic cells), inhibition of target activity[16], cellular internalization, opsonization and functional modulation of cellular signaling pathways (Supplementary Fig. 1).

Sorted cells are immediately re-compartmentalized in sub-nanoliter droplets together with a hydrogel bead grafted with barcoded primers[20], to generate droplets that containe a single cell and a single bead (Fig. 1b, right). After cell lysis and reverse transcription of V_H and V_L messenger RNAs in the droplets, the complimentary DNAs from each cell carry a unique barcode that allows cognate V_H and V_L pairs to be identified by NGS. This platform allows coupled phenotype–genotype repertoire characterization and informed selection of multiple target-specific antibody candidates within just 3.5 days of cell harvesting, followed by ~17 days for gene synthesis, cloning, production and validation of the phenotypic properties of the selected IgGs (Fig. 1c).

We applied the platform to characterize repertoires of IgG-secreting cells from mice immunized with two soluble protein antigens, the vaccine target tetanus toxoid (TT) and the multifunctional enzyme neuroleukin/glucose-6-phosphate isomerase (GPI) involved in cancer progression and metastasis[21,22], and with cells overexpressing an integral human membrane protein, tetraspanin-8 (TSPAN8), a potential cancer target involved in tumor proliferation and invasion[23].

We screened and sequenced IgG-secreting splenocytes from two mice immunized with TT (TT1 and TT2), three mice immunized with GPI (GPI1, GPI2 and GPI3) and two mice immunized with TSPAN8-expressing mouse M300.19 cells (TSPAN8-1 and TSPAN8-2), or from a pool of splenocytes from five TT-immunized mice (TT5-pool) (see Methods). Splenocytes were harvested from each mouse and B-cell lineages were enriched 4–6-fold by depletion, resulting in 0.8–10% $B220^{-/int}/CD138^{+}$ plasmablasts and plasma cells (Methods; Supplementary Table 1). From 1.5 to 11 million enriched cells were compartmentalized with bioassay reagents in 4–11 million monodisperse 40-pl aqueous droplets, produced at 2,100 droplets per second by hydrodynamic flow focusing[24] with an inert fluorinated carrier oil (Fig. 1b, left).

To detect cells that were secreting IgGs binding soluble antigens (TT and GPI), splenocytes from TT- or GPI-immunized mice were compartmentalized in droplets following a Poisson distribution with a mean, λ, of 0.2–0.4 cells per droplet to minimize compartmentalization of more than one cell per droplet. After incubating for 60 min at 37 °C to allow antibody secretion, antigen binding was detected using a fluorescent sandwich immunoassay, in which secreted antibodies were captured onto paramagnetic colloidal nanoparticles coated with an anti-mouse kappa light chain (Igκ) nanobody (V_HH) in each droplet[2]. Upon application of a magnetic field, the ~1,300 nanoparticles in each droplet formed an elongated aggregate (termed a beadline). Each droplet also contained Alexa Fluor 647-labeled Fc-specific $F(ab')_2$-anti-mouse IgG that relocated onto the beadline if the secreted antibody was an IgG, and Alexa Fluor 488-labeled antigen that relocated to the beadline if the IgG had affinity for the antigen, but not otherwise (Fig. 2a).

The spatial distribution of fluorescence in the droplets was analyzed by re-injecting them into a second microfluidic chip where each droplet was scanned with superimposed laser lines and epifluorescence was detected using photomultiplier tubes. Secreted IgG and antigen binding were determined from red and green fluorescence localization to the beadline, respectively (Methods; Fig. 2a–c and Supplementary Fig. 2), and droplets that contained a cell of interest were sorted by fluorescence-activated dielectrophoretic sorting (Methods; Fig. 2d and Supplementary Fig. 3). Compared to approaches that rely on the co-encapsulation of a single cell and a single microparticle[25], the use of beadlines that comprise many nanoparticles increases the binding capacity, eliminates the double Poisson statistics that govern single cell–single microparticle encapsulation to increase the number of cells assayed, drastically lowers the number of droplets that contain no particles, and lowers the risk that the fluorescent signals are not in the focal plane of the detection system[2]. Additionally, droplet scanning with laser lines parallel to beadlines optimizes detection sensitivity as relocalized fluorophores will be maximally excited.

To detect cells that secrete IgGs bound to cells that overexpress human TSPAN8, calcein AM-labeled M300.19 reporter cells expressing TSPAN8 (that is, the same cell line used for mouse immunizations) were co-compartmentalized in 80-pl droplets with calcein violet-labeled B-cell-enriched splenocytes with a mean, λ, of 1.6 reporter cells and 1 splenocyte per droplet (Fig. 2e) and were incubated for 60 min at 37 °C. Alexa Fluor 647-labeled Fc-specific $F(ab')_2$-anti-mouse IgG relocated onto the green fluorescent reporter cell if the secreted IgG bound the antigen, but not otherwise (Fig. 2e–g). The distribution of fluorescence in the droplets was analyzed as above and droplets that contained a cell of interest were sorted by fluorescence-activated dielectrophoretic sorting (Methods; Fig. 2h and Supplementary Fig. 4).

IgG secretion was detected in droplets from 2.7% and 2.8% of the compartmentalized cells from B-cell-enriched splenocytes from TT-immunized and GPI-immunized mice, respectively. On average, for the TT-immunized mice ~47% of IgG-secreting cells showed detectable TT binding (58% for TT5-pool) and for the

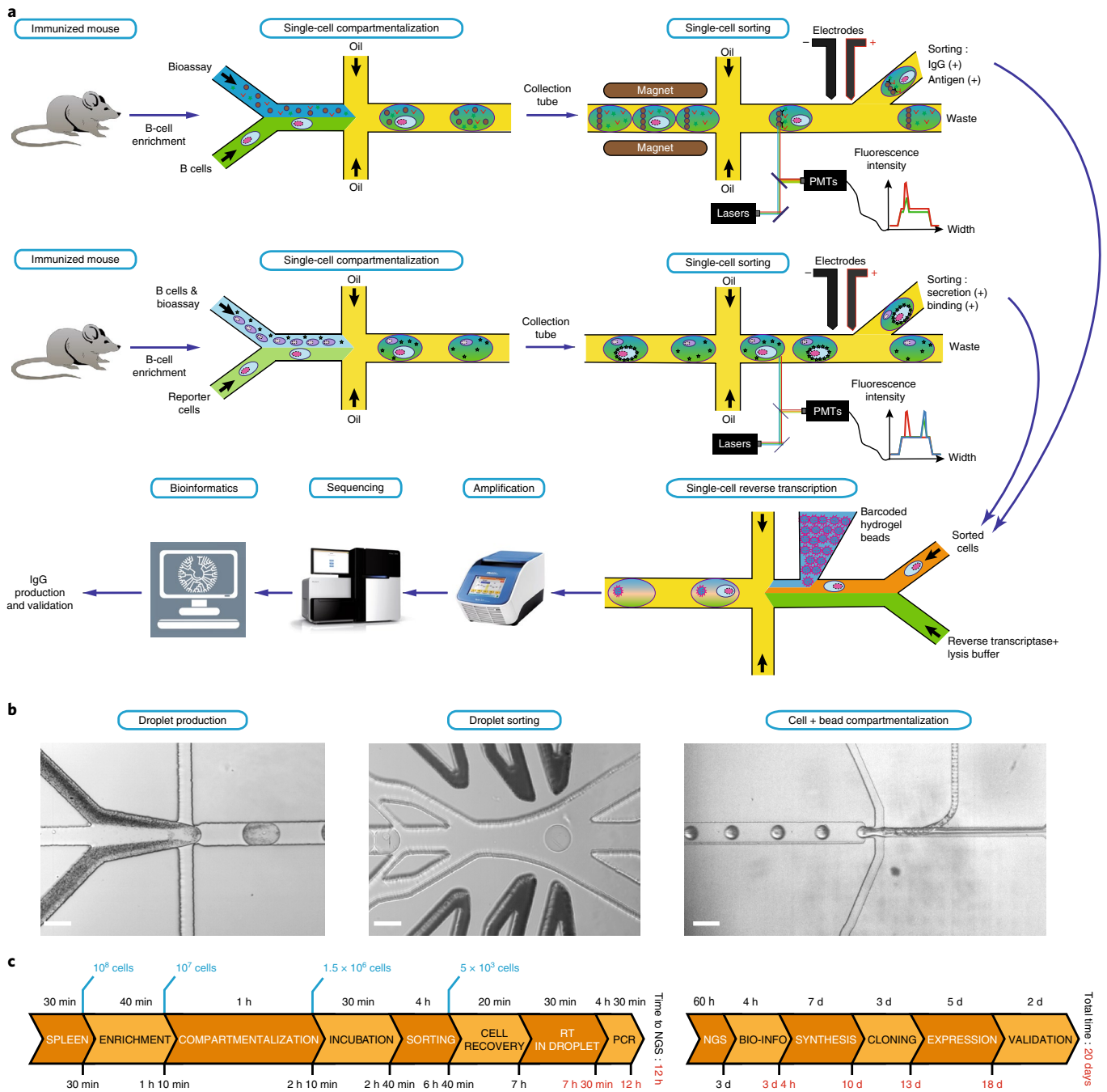


Fig. 1 | The Celli*GO* process. a, Schematic of sorting on the basis of IgG binding to the soluble antigens TT and GPI (top panel) or to the membrane antigen TSPAN8 (middle panel), followed by paired V_H–V_L sequencing of the sorted cells (bottom panel). See text for details. **b**, Images of microfluidic operations. From left to right: droplet production, droplet sorting and co-compartmentalization of single cells with single beads in droplets (see also Supplementary Movie). Scale bars, left and middle 100 µm, right 40 µm. **c**, Celli*GO* timeline. The total time for the entire process from cell harvesting to antibody validation takes 20 d, including only 12 h from cell harvesting to NGS. PMTs, photomultiplier tubes; RT, reverse transcription.

GPI-immunized mice ~28% of IgG-secreting cells showed GPI-binding, which is consistent with the 6-fold higher mean serum titer (measured as EC50, the half maximal effective concentration) in TT-immunized mice (Supplementary Table 1). Very low frequencies (<0.002%) of antigen-binding IgGs were detected in splenocytes from mice that were immunized with a different antigen, naive mice, or B-cell-deficient ($RAG^{-/-}$) mice (Supplementary Fig. 5).

From ~3,500 to ~22,000 droplets that contained secreted IgG with detectable binding to TT, GPI or TSPAN8-overexpressing cells were sorted per run (Methods; Supplementary Table 1). The entire compartmentalization, bioassay and sorting process took typically 5.5 h (Fig. 1c), which contributed to the high viability (79% ± 2%) of cells recovered from sorted droplets (Supplementary Fig. 6a). Recovered cells demonstrated plasma cell or blast morphology; ~56% of sorted cells from TT sorts secreted IgG (by ELISpot) and most, if not all, of these cells secreted antigen-specific IgG (Supplementary Fig. 6b,c).

To initiate barcoded single-cell reverse transcription, sorted cells were re-compartmentalized for ~10 min at ~250 s^{-1} in ~100,000 ~1 nl droplets together with lysis buffer, reverse transcriptase and

~160 pl of hydrogel beads (Fig. 1b, right). Each hydrogel bead carried ~10^9 primers for V_H and V_L reverse transcription, tagged with a unique barcode generated by split-and-pool synthesis (Methods; Supplementary Fig. 7). The injection of close-packed deformable beads[26] (Supplementary Movie) resulted in >79% of droplets that contained a single barcoded hydrogel bead and <2.5% that contained more than one bead, and led to compartmentalization of ~70% of droplets that contained a single cell together with a single bead (Supplementary Fig. 8). In the droplet the cells were lysed and the barcoded primers, released from the beads by photocleavage, were used as primers to reverse transcribe the V_H and V_L mRNA (Fig. 3a). The pooled barcoded cDNAs recovered from the droplets were amplified and sequenced using Illumina MiSeq (2×300 bp) paired-end sequencing to identify cognate V_H–V_L pairs that carried the same barcode (Supplementary Fig. 9).

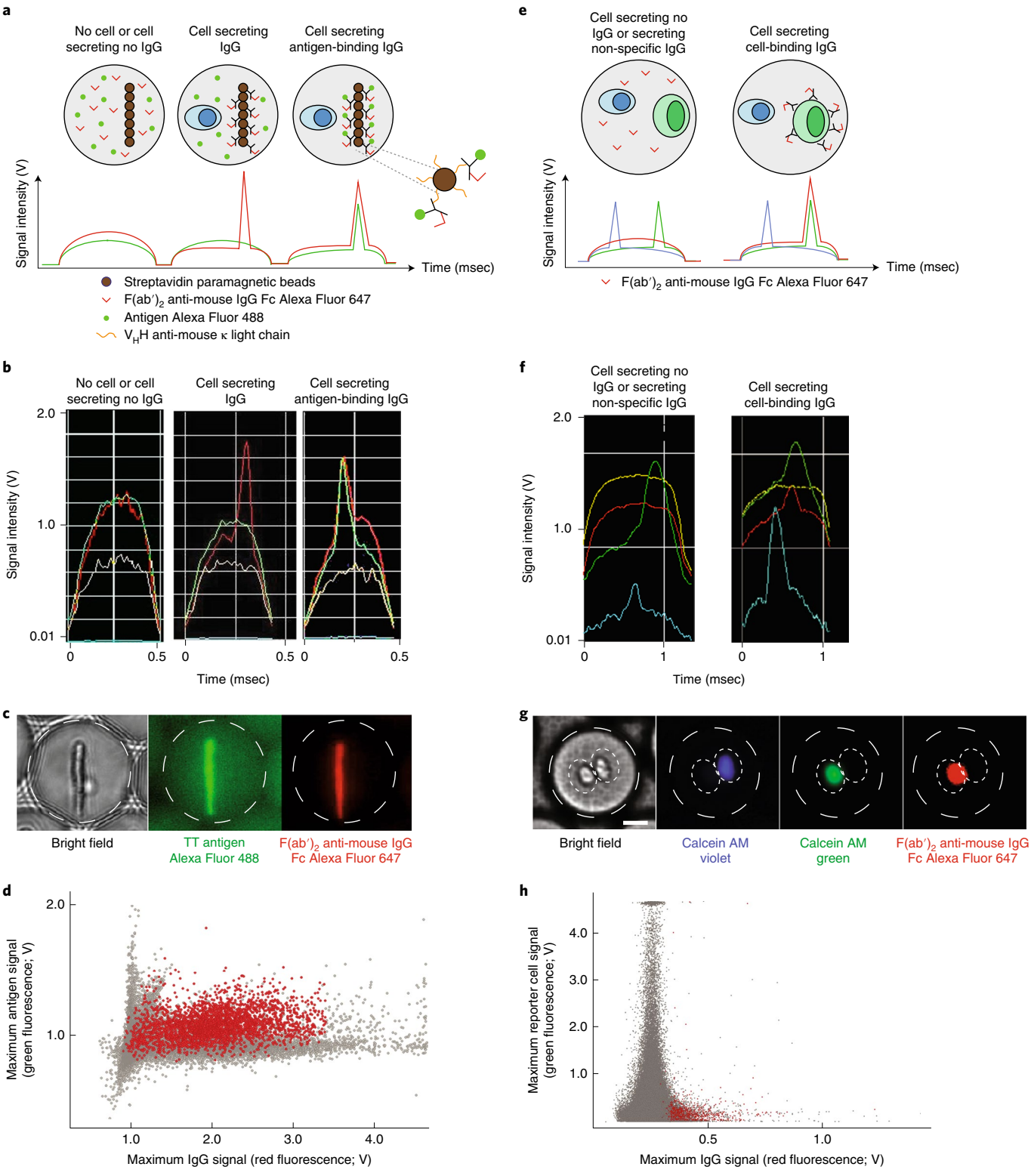
Sequencing data were processed computationally to identify cell-specific barcodes and to identify cognate V_H–V_L pairs (one pair per barcode) and were corrected for errors introduced by PCR or sequencing by creation of consensus sequences (see Methods; Supplementary Note 1 and Supplementary Figs. 10,11). Non-redundant V_H–V_L pairs were clustered into IgG clonotypes that derive from an ancestral recombination event and have expanded and undergone affinity maturation (see Methods). The total number of V_H–V_L pairs, unique IgGs and to a lesser extent, the total number of clonotypes correlate positively with serum IgG titer (EC50) in TT and GPI-immunized mice (Supplementary Fig. 12 and Supplementary Table 1). For each antigen sort, we saw a high diversity of IgGs covering multiple V-gene families (Fig. 3b,c and Supplementary Figs. 13–15). Matrix statistical comparison of V-gene usage in heavy and light chains separately (Methods) identified significant differences ($P<0.05$) not only between samples from mice that were immunized against different antigens, but also between mice that were immunized with the same antigen, particularly in the case of TT (Supplementary Table 2) compared to GPI or TSPAN8. This suggests higher stochasticity in the immune response of individual mice to TT compared to GPI or TSPAN8. Analysis of the combined V_H–V_L gene usage showed that the difference (distance, D) was almost maximal (that is, close to 1) with no or rare identification of common V_H–V_κ pairs between mice that were immunized with different antigens, and was lower—with frequent identification of common V_H–V_L pairs—between mice that were immunized with the same antigen (Supplementary Figs. 16,17). Multidimensional scaling of the distance matrix indicates an antigen-dependent clustering of the combined V_H–V_L usages, without cluster overlaps in the first two dimensions (Supplementary Fig. 18). Moreover, within IgG clonotypes, construction of phylogenetic trees of natively paired V_H and V_L genes revealed evidence of co-evolution during affinity maturation (exemplified for GPI sorts in Fig. 3d,e). An average of 95, 72 and 35 IgG clonotypes were identified for TT-, GPI- and TSPAN8-immunized mice, respectively (per 1.5×10^6 encapsulated cells; Supplementary Table 1). The ratio of V_H–V_L pairs in multi-member clonotypes compared to singletons, the average size of clonally expanded V_H–V_L clusters, and the fraction of highly expanded clonotypes (containing more than ten members), were all highest for TT, intermediate for GPI and lowest for TSPAN8-immunized mice (Fig. 3f and Supplementary Fig. 19a,b). This is consistent with greater clonal expansion being linked to stronger and more specific immune responses but may also reflect differences in the screening bioassays (for example, the use of purified antigen versus transfected reporter cells).

Screening of cells that were pooled from five TT-immunized mice identified 219 IgG clonotypes; that is, a greater than 2-fold increase in diversity compared to a single TT-immunized mouse, which suggests that a pooling strategy may be useful to maximize antibody candidate diversity per sort (Supplementary Table 1). For both soluble antigens, ≥96–99% of the non-redundant IgG sequences had undergone somatic hypermutation (SHM), compared with only 65% of such sequences for the membrane antigen; this may be due to lower efficiency of cell-based immunization and/or higher false positive rates (Supplementary Table 1). Across sorts, amino acid changes were more abundant in V_H proteins than in V_L proteins (Supplementary Figs. 19c,20), and were enriched in complementarity determining regions compared to framework regions (Supplementary Fig. 21), as has been reported for other antigenic specificities[27]. We found no significant differences in the extent of SHMs between expanded and singleton clonotypes in TT-immunized mice ($P > 0.05$; two-sided Welch's test), but found significantly higher numbers of SHMs in expanded clonotypes with GPI ($P < 2.2\times10^{-16}$; two-sided Welch's test) and to a lesser extent TSPAN8 ($P = 0.038$; two-sided Welch's test) (Supplementary Fig. 19d), which suggests that affinity maturation is a feature of clonally expanded cells, albeit dependent on the antigen.

To demonstrate antigen-binding activity, putative TT-, GPI- and TSPAN8-binding IgGs from one or more representative sorts were cloned and expressed as human IgG1, κ isotypes (see Methods; Supplementary Table 3). From the TT sorts, 42 V_H–V_L pairs were expressed. Almost all of these (39/42; ~93%) bound to TT by enzyme-linked immunosorbent assay (ELISA). The EC50 for TT was measured for 27 of these, and ranged from 0.1 nM to 300 nM (median 0.8 nM). From the GPI sorts, nearly all (13/14; ~93%) of the V_H–V_L pairs expressed had measurable affinity for GPI by surface plasmon resonance, which ranged from 0.2 nM to 26 nM (median 0.4 nM)

Fig. 2 | Phenotypic screening and sorting of IgG-secreting cells. a,**e**, Binding assay in droplets. The droplets are scanned one-by-one as they pass the laser line, producing a 1.7 ms time trace of fluorescence signals in four colors (blue, green, orange and red) for each droplet. The distribution of the fluorescent signal of the soluble antigen (**a**) or the reporter cell expressing the membrane antigen (**e**) (green fluorescence), and of the secreted IgG (detected by red fluorescent F(ab')$_2$ anti-IgG Fc) are monitored in real time for each droplet at ~ 600 Hz. A schematic of the time trace of both signals is shown. **a**, from left to right, for a drop that contains no IgG-secreting cell (both green and red signals are uniformly distributed across the drop), a drop that contains a cell secreting an IgG that does not bind the antigen (the secreted IgG is captured on the beads and the red fluorescent F(ab')$_2$ anti-IgG Fc is relocated to the bead line) and a drop that contains a cell secreting an antigen-binding IgG (the secreted IgG is captured on the beads and both the red fluorescent F(ab')$_2$ anti-IgG Fc and the green fluorescent antigen are relocated to the beadline)**. e**, left, for a drop that contains a cell secreting an IgG that does not bind the reporter cell and right, a drop that contains a cell secreting an IgG that binds the reporter cell (the secreted IgG is captured on the green fluorescence reporter cell, the red fluorescent F(ab')$_2$ anti-IgG Fc is relocated to the reporter cell and the red and green peaks overlap). **b**,**f**, Experimental time traces recorded for droplets analyzed at ~ 600 Hz and corresponding to the examples schematized in panels **a** and **e** for the TT antigen (**b**) and TSPAN8$^+$ M300.19 cells (**f**). Droplets are sorted if they display a green peak (green line, relocated antigen (**b**) or TSPAN8$^+$ M300.19 cells (**f**)), and a red peak (red line, relocated antibody) that co-localize. Blue fluorescence (blue line) and orange fluorescence (orange line) signals are used to identify droplets containing purified control antibodies (**b**) or calcein violet-labeled splenocytes and control antibodies (**f**), respectively (see Methods). **c**,**g**, Bright field and epifluorescence micrographs of a sorted droplet containing a splenocyte secreting a TT-binding IgG (**c**) or a TSPAN8$^+$ M300.19 cell-binding IgG (**g**)). **d**,**h**, Plots of maximum green versus maximum red fluorescence in each droplet containing cells from a TT-immunized mouse (**d**) or a TSPAN8-immunized mouse (**h**). The sorted droplets are indicated as red dots. Panels **b**, **c**, **f** are representative of 5, 3 and 10 experiments, respectively. See Supplementary Fig. 2 for further details of the assay and signal analysis and Supplementary Figs. 3,4 for the full gating strategy.

(Fig. 3g). No correlation between the number of mutations and the affinity was observed, as has been described for antibody responses following infection[28,29], though four expressed germline V_H–V_L pairs (no somatic mutations in either chain) showed no detectable antigen binding (Supplementary Table 3). From the TSPAN8 sorts, 7 out of 21 (33%) V_H–V_L pairs demonstrated higher binding to TSPAN8-transfected cells than to non-transfected cells, but among these, only 3 (14%) bound strongly enough to allow the affinity for TSPAN8 to be determined using flow cytometry, with affinities of 374 nM, 397 nM and 3,640 nM (Fig. 3h, Supplementary Fig. 22 and Supplementary Table 3). The lower frequency of IgGs that showed detectable antigen binding to TSPAN8 compared to TT and GPI may be due to the immunization strategy, as the murine TSPAN8-expressing M300.19 cells that were used for immunizations in BALB/c mice originate from NIH Swiss outbred mice and should be immunogenic for other membrane antigens; the fact that the immunization and the in-droplet



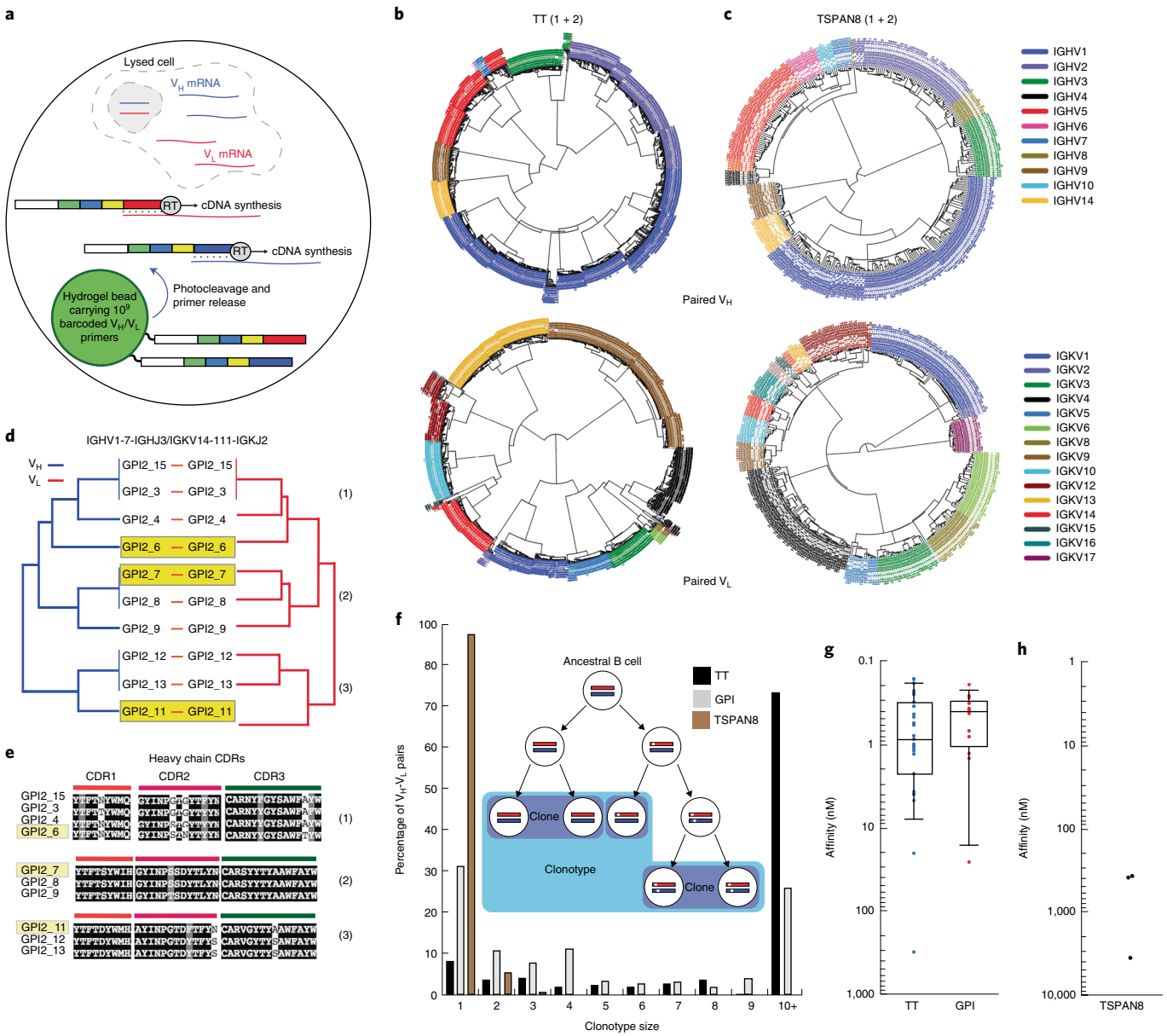


Fig. 3 | Sequencing, expression and characterization of IgGs expressed by sorted cells. a, In-droplet single-cell barcoding of V_H and V_L cDNA. Cells from sorted droplets were recovered and re-compartmentalized individually in new droplets in the presence of reverse transcription reagents and a hydrogel bead bearing ~10^9 copies of V_H-V_L RT primers tagged with a bead-specific DNA barcode, which was released from the hydrogel beads by ultraviolet photocleavage. After breaking the emulsion, the pooled, barcoded V_H-V_L cDNAs were amplified and processed for NGS. **b**,**c**, Diversity of V-gene families use in paired V_H-V_L sequences from the TT-immunized mice (TT1 + TT2) (**b**) and the TSPAN8-immunized mice (TSPAN8-1 + TSPAN8-2) (**c**). Phylogenies were created using the full-length amino acid sequences of the V regions. Sequences are color-coded by the closest germline V-gene family. See also Supplementary Figs. 13–15. **d**, Cross-phylogeny of ten unique paired V_H-V_L sequences from the GPI2 sort derived from the same recombined germline V(D)J genes, illustrating co-evolution of V_H and V_L during affinity maturation. One candidate from each cluster (indicated by (1), (2) and (3) on the right side of tree) was expressed and was tested for binding to GPI (shaded yellow boxes); the three IgGs had similar affinity (K_d = 0.195–0.375 nM). **e**, The sequences of the complementarity determining regions (CDRs) from the paired V_H-V_L sequences presented in panel **d**. One candidate from each cluster was expressed and was tested for binding to GPI (shaded yellow boxes); the three IgGs had similar affinity (K_d = 0.195–0.375 nM). **f**, Histogram of frequency (percentage) of paired V_H-V_L sequences versus clonotype size. The clonotype size is equal to the number of non-redundant V_H-V_L sequences derived from the same ancestral V_H-V_L germline recombination event (that is, derived from the same germline V-J genes, with CDR3s of the same length and differing by at least one amino acid across the variable region). Inset: Schematic of IgG clustering on the basis of clonal expansion and somatic hypermutation of a single ancestral cell. Germline V_H and V_L genes are indicated by red and blue rectangles, and vertical white bars indicate mutations. Dark blue regions define clusters of clones (V_H-V_L pairs with exactly the same amino acid sequence but associated with different barcodes and therefore derived from different cells) (size = 2, 2 and 1). The number of dark blue regions in a light blue region defines the clonotype size (size = 3). See also Supplementary Table 4. **g**, Affinities of purified anti-TT IgGs (EC50 by ELISA; n = 27) and anti-GPI IgGs (K_d by bio-layer interferometry; n = 13) expressed from cloned genes that were identified by sequencing of TT-immunized mice and GPI-immunized mice, respectively, tested for binding to the immunogen. Boxes extend from the 25th to the 75th percentile, whiskers indicate the 10th to the 90th percentile and the line in the middle of the box is the median. **h**, Affinities of purified anti-TSPAN8 IgGs (K_d by flow cytometry; n = 3) expressed from cloned genes that were identified by sequencing of TSPAN8-immunized mice. See also Supplementary Table 3.

bioassay were performed using the same transfected cell line may also be a factor. The lower affinities of the anti-TSPAN8 IgGs compared to anti-TT and anti-GPI IgGs may also be due to the immunization strategy; the frequency of somatic hypermutations was lower than that observed with anti-TT and anti-GPI IgGs, which indicates a lower level of affinity maturation (Supplementary Fig. 20). Nevertheless, this shows that Celli*GO* is capable of discovering low affinity antibodies against challenging membrane protein targets, such as TSPAN8, that elicit weak immune responses.

The Celli*GO* pipeline provides the capacity to deeply mine the IgG repertoire by high-throughput screening of millions of IgG-secreting cells on the basis of coupled analysis of phenotype (activity of the secreted IgG) and genotype (the sequence of the IgG), on both soluble (antigen or autoantigen) and membrane-bound targets. From 337 to 1,053 non-redundant IgGs were identified from splenocytes of a single TT-immunized mouse (Supplementary Table 1) with >90% of cloned IgGs binding TT. This compares to the identification of only 21 non-redundant IgGs using hybridomas, 42 IgGs by phage-display[30], and 86 IgGs by FACS of TT-stained human peripheral IgG^+ plasmablasts followed by single-cell mRNA capture and NGS[9]. Furthermore, the median affinity of the cloned anti-TT IgGs was 0.8 nM, compared to 3.4 nM, 9.0 nM and 2.15 nM with hybridomas, phage display and FACS plus single-cell mRNA capture and NGS, respectively[9,30].

The system can be simply adapted to analyze circulating human memory B-cell repertoires, by inducing peripheral B cells in vitro to secrete IgG prior to encapsulation in droplets. Human memory B cells enriched from peripheral blood mononuclear cells from two donors were activated with anti-CD40 monoclonal antibody and CpG oligodeoxynucleotides prior to encapsulation. Analysis of three different runs resulted in the identification of ~460–900 paired V_H–V_L sequences per run from ~2,300 IgG-secreting cells in each run (Supplementary Fig. 23). Simply changing the primer regions on the barcoded primers on the hydrogel beads allows antibodies from different species to be sequenced, or indeed, enables transcriptomic profiling of cells (Supplementary Fig. 24). Alternatively, paired V_H–V_L sequencing on sorted cells can be performed by single-cell reverse transcription-PCR in microtiter plate wells prior to NGS (Supplementary Fig. 25), which allows direct cloning of the amplified V_H and V_L genes into expression vectors without gene synthesis.

Celli*GO* is capable of analyzing cells from different species and allows phenotypic screening of IgG repertoires derived from large numbers of plasma cells, plasmablasts and in vitro activated memory B cells from different anatomical compartments at high throughput, and using a wide range of assays, including functional and cell-based assays (Supplementary Fig. 1). It should facilitate the analysis of immune responses and the identification of potent antibodies with unique functional activities for therapeutic purposes, including those against difficult targets, such as insoluble membrane antigens.

Online content

Any methods, additional references, Nature Research reporting summaries, source data, extended data, supplementary information, acknowledgements, peer review information; details of author contributions and competing interests; and statements of data and code availability are available at https://doi.org/10.1038/s41587-020-0466-7.

Received: 22 March 2018; Accepted: 25 February 2020;
Published online: 30 March 2020

References

1. Georgiou, G. et al. The promise and challenge of high-throughput sequencing of the antibody repertoire. *Nat. Biotechnol.* **32**, 158–168 (2014).
2. Eyer, K. et al. Single-cell deep phenotyping of IgG-secreting cells for high-resolution immune monitoring. *Nat. Biotechnol.* **35**, 977–982 (2017).
3. Bannard, O. & Cyster, J. G. Germinal centers: programmed for affinity maturation and antibody diversification. *Curr. Opin. Immunol.* **45**, 21–30 (2017).
4. Gunn, B. M. & Alter, G. Modulating antibody functionality in infectious disease and vaccination. *Trends Mol. Med.* **22**, 969–982 (2016).
5. Nutt, S. L., Hodgkin, P. D., Tarlinton, D. M. & Corcoran, L. M. The generation of antibody-secreting plasma cells. *Nat. Rev. Immunol.* **15**, 160–171 (2015).
6. Meijer, P. J. et al. Isolation of human antibody repertoires with preservation of the natural heavy and light chain pairing. *J. Mol. Biol.* **358**, 764–772 (2006).
7. Di Niro, R. et al. High abundance of plasma cells secreting transglutaminase 2-specific IgA autoantibodies with limited somatic hypermutation in celiac disease intestinal lesions. *Nat. Med.* **18**, 441–445 (2012).
8. Robinson, W. H. Sequencing the functional antibody repertoire–diagnostic and therapeutic discovery. *Nat. Rev. Rheumatol.* **11**, 171–182 (2015).
9. DeKosky, B. J. et al. High-throughput sequencing of the paired human immunoglobulin heavy and light chain repertoire. *Nat. Biotechnol.* **31**, 166–169 (2013).
10. Busse, C. E., Czogiel, I., Braun, P., Arndt, P. F. & Wardemann, H. Single-cell based high-throughput sequencing of full-length immunoglobulin heavy and light chain genes. *Eur. J. Immunol.* **44**, 597–603 (2014).
11. McDaniel, J. R., DeKosky, B. J., Tanno, H., Ellington, A. D. & Georgiou, G. Ultra-high-throughput sequencing of the immune receptor repertoire from millions of lymphocytes. *Nat. Protoc.* **11**, 429–442 (2016).
12. Wang, B. et al. Functional interrogation and mining of natively paired human VH:VL antibody repertoires. *Nat. Biotechnol.* **36**, 152–155 (2018).
13. Adler, A. S. et al. A natively paired antibody library yields drug leads with higher sensitivity and specificity than a randomly paired antibody library. *MAbs* **10**, 431–443 (2018).
14. Czerkinsky, C. C., Nilsson, L. A., Nygren, H., Ouchterlony, O. & Tarkowski, A. A solid-phase enzyme-linked immunospot (ELISPOT) assay for enumeration of specific antibody-secreting cells. *J. Immunol Methods* **65**, 109–121 (1983).
15. Story, C. M. et al. Profiling antibody responses by multiparametric analysis of primary B cells. *Proc. Natl Acad. Sci. USA* **105**, 17902–17907 (2008).
16. El Debs, B., Utharala, R., Balyasnikova, I. V., Griffiths, A. D. & Merten, C. A. Functional single-cell hybridoma screening using droplet-based microfluidics. *Proc. Natl Acad. Sci. USA* **109**, 11570–11575 (2012).
17. Shembekar, N., Hu, H., Eustace, D. & Merten, C. A. Single-cell droplet microfluidic screening for antibodies specifically binding to target cells. *Cell Rep.* **22**, 2206–2215 (2018).
18. Bradbury, A. R., Sidhu, S., Dubel, S. & McCafferty, J. Beyond natural antibodies: the power of in vitro display technologies. *Nat. Biotechnol.* **29**, 245–254 (2011).
19. Fitzgerald, V. & Leonard, P. Single cell screening approaches for antibody discovery. *Methods* **116**, 34–42 (2017).
20. Klein, A. M. et al. Droplet barcoding for single-cell transcriptomics applied to embryonic stem cells. *Cell* **161**, 1187–1201 (2015).
21. Gurney, M. E., Heinrich, S. P., Lee, M. R. & Yin, H. S. Molecular cloning and expression of neuroleukin, a neurotrophic factor for spinal and sensory neurons. *Science* **234**, 566–574 (1986).
22. Liotta, L. A. et al. Tumor cell autocrine motility factor. *Proc. Natl Acad. Sci. USA* **83**, 3302–3306 (1986).
23. Park, C. S. et al. Therapeutic targeting of tetraspanin8 in epithelial ovarian cancer invasion and metastasis. *Oncogene* **35**, 4540–4548 (2016).
24. Anna, S. L., Bontoux, N. & Stone, H. A. Formation of dispersions using "flow focusing" in microchannels. *Appl. Phys. Lett.* **82**, 364–366 (2003).
25. Mazutis, L. et al. Single-cell analysis and sorting using droplet-based microfluidics. *Nat. Protoc.* **8**, 870–891 (2013).
26. Abate, A. R., Chen, C. H., Agresti, J. J. & Weitz, D. A. Beating Poisson encapsulation statistics using close-packed ordering. *Lab Chip* **9**, 2628–2631 (2009).
27. Odegard, V. H. & Schatz, D. G. Targeting of somatic hypermutation. *Nat. Rev. Immunol.* **6**, 573–583 (2006).
28. Roost, H. P. et al. Early high-affinity neutralizing anti-viral IgG responses without further overall improvements of affinity. *Proc. Natl Acad. Sci. USA* **92**, 1257–1261 (1995).
29. Murugan, R et al. Clonal selection drives protective memory B cell responses in controlled human malaria infection.*Sci. Immunol.* **3**, eaap8029 (2018).
30. Sorouri, M., Fitzsimmons, S. P., Aydanian, A. G., Bennett, S. & Shapiro, M. A. Diversity of the antibody response to tetanus toxoid: comparison of hybridoma library to phage display library. *PLoS One* **9**, e106699–106614 (2014).

Methods

Immunization. BALB/c mice (Janvier Labs, age 7–10 weeks at start) were immunized intraperitoneally with 10 µg of formaldehyde-inactivated TT (Statens Serum Institut) or 100 µg of rabbit GPI (89% homologous to mouse GPI; Sigma) emulsified in complete Freund's adjuvant (Sigma, F5881) mixed 1:1 with 0.9% (wt/vol) NaCl (Versylene Fresenius) for primary immunization (day 0), or emulsified in incomplete Freund's adjuvant (Sigma, F5506) mixed 1:1 with 0.9% (wt/vol) NaCl for secondary and tertiary immunization (day 14, 42), or only in 0.9% NaCl for boosting (day 46). BALB/c mice were injected intraperitoneally with 2×10^6 human TSPAN8-transfected m300.19 mouse pre-B cells (originating from the outbred NIH Swiss mouse) in PBS on days 1, 15, 29, 43 and, if titers were estimated to be too low, again on days 57 and 71. Experiments using mice were validated by the CETEA ethics committee number 89 (Institut Pasteur) under agreement no. 2013-0103, and the French Ministry of Research under agreement no. 00513.02, as well as by the French ethics committee number 80 (CellVax) under reference APAFIS no. 6063-2016071213009170 v5.

Extraction of cells from spleen and B-cell enrichment. Cells were extracted from spleens of mice, harvested at day 46 of the immunization schedule for soluble antigens (4 d after boosting), as described[2], or at day 61 for the membrane-bound antigen (4 d after boosting). On average 100–200 million cells were obtained per spleen following red blood cell lysis. Cells from the B-cell lineage (including IgG-secreting cells) were enriched using the Pan B-cell isolation kit II (Miltenyi Biotec) as described[2] and were further enriched by depletion with anti-IgM (Miltenyi Biotec, 130-095-904) and anti-IgD (Miltenyi Biotec, 130-101-899) antibodies. We observed an approximate sixfold enrichment of CD138+ cells by flow cytometry (plasmablast and plasma cells).

Aqueous phase I: preparation of cells for compartmentalization in droplets. Cell suspensions were prepared for compartmentalization in droplets at 4 °C as described[2], and cells were resuspended in DMEM/F12 supplemented with 0.1% Pluronic F68 (LifeTechnologies), 25 mM HEPES pH 7.4 (LifeTechnologies), 5% HyClone super low IgG defined fetal bovine serum (GE Healthcare), and 1% Pen/Strep (ThermoFisher), so as to achieve a λ (mean number of cells per droplet) of ~0.3 for B cells in soluble antigen binding assays and of ~0.6 for B cells in cell-based assays.

Aqueous phase II: preparation of bioassay reagents and reporter cells for compartmentalization in droplets. For the soluble antigens (TT and GPI), paramagnetic nanoparticles (Streptavidin Plus (300 nm), Ademtech) were coated with biotinylated V_HH anti-mouse κ light chain as described[2], and resuspended in working buffer containing rabbit $F(ab')_2$ anti-mouse IgG Fc-specific antibody (Alexa Fluor 647-labeled, Jackson ImmunoResearch) diluted to 75 nM and TT (purified monomer, Alexa Fluor 488-labeled) or GPI (Alexa Fluor 488-labeled) diluted to 50 nM, (resulting in final in-droplet concentrations of 37.5 nM anti-mouse IgG antibody and 25 nM antigen, respectively). The mouse reporter cells M300.19 that stably expressed the transfected membrane-bound antigen TSPAN8 were pre-labeled with calcein green AM and were resuspended as defined above in medium containing rabbit $F(ab')_2$ anti-mouse IgG Fc-specific antibody (Alexa Fluor 647-labeled, Jackson ImmunoResearch) diluted to 75 nM and supplemented with 23.6% (vol/vol)) Nycodenz, to achieve a λ (mean number of cells per droplet) of ~1.5 for reporter cells in cell-based assays.

Microfluidic chips. Separate microfluidic chips were used: device 1 was used to compartmentalize single cells with bioassay reagents in droplets, or to co-compartmentalize IgG-secreting cells with reporter cells expressing a membrane antigen and bioassay reagents in droplets; device 2 was used to sort droplets by fluorescence-activated dielectrophoresis[31]; device 3 produced hydrogel beads and device 4 compartmentalized single sorted cells with single hydrogel beads (Supplementary Fig. 26). All chips were manufactured by soft-lithography in poly-dimethylsiloxane (PDMS)[32] (Sylgard) as described[25]. Masters were made using one or two layers of SU-8 photoresist (MicroChem), depending on the design. The list depth of the photoresist layers for devices 1 and 2 was 40 µm ± 1 µm and for device 3 was 55 µm ± 1 µm. For device 4, the first layer (70–75 µm deep) was for the hydrogel bead inlet and the second layer (130–145 µm deep) was for the cell inlet, for the reverse transcriptase inlet and for the outlet. Electrodes were prepared by melting a 51In 32.5Bi 16.5Sn alloy (Indium Corporation of America) into the electrode channel[33].

Droplet production, collection and incubation. Aqueous phases I and II were co-flowed and partitioned into droplets using hydrodynamic flow-focusing[24] in dripping mode on a microfluidic chip with a nozzle 15 µm wide, 40 µm deep and 10 µm long (Supplementary Fig. 26a). The continuous phase comprised 2-3% (wt/wt) 008-FluoroSurfactant (RAN Biotechnologies) in Novec HFE7500 fluorinated oil (3M). The flow rates were adjusted to generate monodisperse droplets of 40 pl ± 2 pl volume (for assays with soluble antigens), or 80 pl ± 8 pl (for cell-based assays with membrane-bound antigen). Shortly after generation, the droplets were collected into a 10 ml hemolysis tube filled with 10 m of Novec HFE7500 fluorinated oil containing 0.1% (wt/wt) 008-FluoroSurfactant.

In the case of the soluble antigens (TT and GPI) the tube was inserted into a ring magnet (diameter 25/13 mm, 7 kg, Magnets4you) to force the superparamagnetic nanoparticles to form an elongated aggregate (a beadline) inside each droplet. For TT sorts only, in addition to the screening emulsion containing splenocytes, 2 control emulsions (positive and negative) were produced that contained 10 nM purified anti-TT7 antibody[2] and 10 nM mouse anti-horseradish peroxidase antibody (Jackson Immuno Research, 223-005-024), respectively; these were differentiated using sulforhodamine B (Sigma Aldrich, S1402-5G) and DY405 (Dyomics, 405-00) to optically barcode the droplets in the positive- and negative-control emulsions, respectively (see Supplementary Fig. 3).

In the case of the membrane-bound target TSPAN8, enriched B cells were co-compartmentalized in droplets with $TSPAN8^+$ M300.19 reporter cells. The $TSPAN8^+$ M300.19 reporter cells were pre-labeled with 300 nM calcein green AM (Thermo Fisher Scientific, C34852) and the enriched B-cell population was pre-labeled with 300 nM calcein violet AM (Thermo Fisher Scientific, C34858). In addition to the screening emulsion containing splenocytes, two control emulsions—negative and positive—were produced, that contained reporter cells only, or reporter cells and 50 nM anti-hTSPAN8 51D3 IgG antibody (Pfizer; see Supplementary Fig. 4), respectively. These were differentiated using sulforhodamine B to optically barcode the droplets in the positive- and negative-control emulsions, respectively (see Supplementary Fig. 4).

Microfluidic platform. Droplet fluorescence analysis and sorting was performed on a dedicated droplet microfluidic station, similar to that described in ref. [25], but including a fixed focus laser line (solid state laser of wavelength 405 nm, 488 nm, 561 nm or 635 nm, Omicron) oriented parallel to the beadline for fluorescence analysis using photomultiplier tube bandpass filters of 440/40–25 nm, 525/40–25 nm, 593/46–25 nm and 708/75–25 nm (Hamamatsu).

Gating strategy for droplet sorting. The droplets were first gated to eliminate coalesced droplets and retain only droplets of the desired size (Supplementary Figs. 3a,4a). Using the optical droplet barcoding, the negative-control droplets, positive-control droplets and droplets containing splenocytes were detected (Supplementary Figs. 3b,4b).

For the soluble antigens (TT and GPI), fluorescence relocalization to the beadline in droplets was measured by plotting the maximum peak fluorescence signal (F_p) in a droplet against the integrated fluorescence signal (F_i) from the droplet (Supplementary Fig. 3c–e). Fluorescence relocalization to the beadline results in an increase in the ratio F_p/F_i (Supplementary Figs. 2,3c–h). Uncoalesced droplets containing splenocytes were sorted if they satisfied all of the following criteria: 1) relocation of antigen (green fluorescence) to the beadline, 2) relocation of the $F(ab')_2$ anti-mouse IgG Fc (red fluorescence) to the beadline and 3) colocalization of the green and red fluorescence peak (Supplementary Fig. 3g–i). The colocalization value, c, was thus bounded between 0 and 1, 1 being the perfect colocalization of the two peaks. Droplets were sorted if $c > 0.95$.

A similar gating strategy was employed for the membrane antigen (TSPAN8) (Supplementary Fig. 4d–i), except that the droplets were first selected for the presence in the drop of both a reporter cell and an enriched B cell on the basis of their fluorescent labeling (Supplementary Fig. 4c). Relocalization of red fluorescence to the reporter cell results in an increase in the ratio F_p/F_i (Supplementary Fig. 2). Uncoalesced droplets that contained splenocytes were sorted if they satisfied all of the following criteria: 1) presence of a reporter cell overexpressing TSPAN8 (green fluorescence) in the droplet, 2) presence of an enriched B cell (violet fluorescence) in the droplet, 3) relocation of the $F(ab')_2$ anti-mouse IgG Fc (red fluorescence) on the reporter cell and 4) colocalization of the green and red fluorescence peaks (Supplementary Fig. 4i).

For all antigens the colocalization parameter, c, was calculated from the time interval between the peaks in the red and green fluorescence channels, t_p, and the time interval from the beginning to the end of the droplet, t_d: $c = 1 - (t_p/t_d)$. The value of c was thus bounded between 0 and 1, 1 being the perfect colocalization of the two peaks. Droplets were sorted if $c > 0.95$ for TT and GPI, and $c > 0.92$ for TSPAN8.

Droplet sorting and cell recovery. For the soluble antigens (TT and GPI), droplets were sorted by dielectrophoresis as described[31], while a magnetic field oriented perpendicularly to the droplet flow was applied using permanent magnets (K&J Magnetics). The device (Supplementary Fig. 26b) can sort droplets into up to two bins, by activation of the electrodes above or beneath the channel; however, only one bin was used in this study. The inlet flows, Q_{em} for the emulsion and Q_{oil} for the oil, were adjusted to sort droplets at 600 s^{-1}; typical parameters for sorting were $Q_{em} = 50$ µl h^{-1} (180 mbar), $Q_{oil} = 50$ µl h^{-1} (180 mbar), F (the frequency of the sorting pulse) = 10 kHz, τ_{sort} (the duration of the sorting pulse) = 2,000 µs and U_{sort} (the peak-to-peak voltage applied across the electrodes) = 400 kVp-p. Sorted droplets were collected in a 1.5 ml tube cooled to 4 °C, which contained an emulsion of 40,000 Chinese hamster ovary cells that were used as carrier cells during recovery. Cells were recovered by addition of 100 µl of DMEM/F12 supplemented with 5% low IgG serum, followed by 100 µl of 1*H*,1*H*,2*H*,2*H*-Perfluoro-1-octanol (Sigma, 370533); cells were then mixed gently and centrifuged at 700*g* for 10 min at 4 °C to favor complete phase separation.

ELISpot. IgG-secreting cells were enumerated after TT sorts and TSPAN8 sorts by ELISpot using the MabTech ELISpotPLUS Kit (Mabtech, 3825-2AW-Plus) on an AID ELISpot Reader. ELISpot 96-well plates were coated with 0.5 mg ml^{-1} anti-IgG mouse monoclonal antibody (Mabtech, 3825-2AW-Plus) or, specifically for cells from TT sorts, with 30 ug ml^{-1} antigen TT (Statens Serum Institut), and IgG secretion was detected using 0.5 mg ml^{-1} biotinylated anti-IgG mouse monoclonal antibody (Mabtech, 3825-2AW-Plus).

Labeling and preparation of sorted cells for single-cell sequencing. Sorted cells and carrier cells were labeled for 20 min at room temperature in the dark with 1 µM Calcein AM (Thermo Fisher Scientific, C3099). The cells were then washed in 400 µl of 0.1% Pluronic F-68 non-ionic surfactant (Thermo Fisher Scientific, 24040032), 25 mM Hepes–KOH, 5% (vol/vol) low IgG serum, in DMEM/F12, centrifuged for 5 min at 400*g* at 4 °C and were then resuspended in 50 µl of 1x PBS containing 21.82% (vol/vol) Optiprep density gradient solution (Sigma) and 0.01 mg ml^{-1} BSA.

Production of barcoded hydrogel beads. Polyacrylamide hydrogel beads of 60-µm diameter were produced by polymerization in droplets made with a microfluidic device (Supplementary Fig. 26c) using a method similar to that described previously[20,34]. However, barcoded primers were then added to the beads by split-and-pool synthesis using ligation rather than primer extension (Supplementary Fig. 7a,b). One million beads—each of which carried ~10^9 copies of a double-stranded DNA oligonucleotide with a 5′-overhang (complementary to the first index 5′-overhang sequence), a photo-cleavable site and the T7-SBS12 sequence—were distributed into 96 wells of a microtiter plate. Each well contained 10 µl of 5 µM double-stranded DNA with a different first index (index A), a complementary 5′-overhang to the first DNA at one end and a different 5′-overhang at the other end (Supplementary Table 5), and these were ligated for 15 minutes at 23 °C using T7 DNA ligase (New England Biolabs) according to the manufacturer's instructions. The hydrogel beads were then pooled, washed as described[34] and re-distributed as above into the wells of a second microtiter plate, each of which contained a double-stranded DNA with a different second index (index B), a complementary 5′-overhang to index A at one end and a different 5′-overhang at the other end, which was ligated to index A. Repetition of this splitting and pooling process 3 times in total (adding 3 indexes) results in 96^3 combinations, which generates ~10^6 different barcodes. After addition of the last index, the beads were pooled, and a mixture of double-stranded DNA molecules with a complementary 5′-overhang to index C, and which contained gene-specific primer regions complementary to the regions encoding the mouse J_H regions and the mouse Cκ domain (Supplementary Fig. 9), were ligated to the barcodes on the beads. The second strand of the primer was then removed by incubating for 2 min at 22 °C with 300 mM NaOH. After completion of the process, each hydrogel bead has a total of ~10^9 primers that carry the same bead-specific barcode. See Supplementary Fig. 7 for further details and quality control.

Single-cell barcoded complementary DNA synthesis. Individual sorted cells (and Chinese hamster ovary carrier cells) were co-compartmentalized in droplets with individual barcoded hydrogel beads and lysis and reverse transcription reagents using a microfluidic device (Supplementary Fig. 26d) as described[34] (see Supplementary Fig. 8). Droplets of ~1 nl volume were formed at 250 s^{-1}. The droplets were collected in a 1.5 ml tube containing the solvent HFE-7500 (Fluorochem) and 0.1% surfactant, and were photo-cleaved by ultraviolet light for 90 seconds (OmniCure, AC475; 365 nm) and then incubated at 50 °C for cell lysis and cDNA synthesis.

Sequencing library preparation. The emulsion containing the barcoded cDNA was broken by adding one volume of 1*H*,1*H*,2*H*,2*H*-Perfluoro-1-octanol. The pooled, barcoded cDNAs were run on a Novex TBE–Urea 6% polyacrylamide gel and cDNAs above 250 bases were excised, purified using a Qiaex II kit (Qiagen, 20051), eluted in 10 mM Tris-HCl pH 8.0, further purified with Agencourt RNA CleanUp beads (Beckman, A63987) at a 1:1 ratio (vol/vol), and eluted in 40 µl DNase- and RNase-free H_2O. The sequencing library was generated by two-step nested PCR using GoTaq Polymerase (Promega), external (PCR1) and internal (PCR2) reverse primers (T7 primer followed by Illumina TruSeq indexed primer) and external (PCR1) and internal (PCR2) forward primers specific to the V_H and V_κ leader and framework sequences, as described[35] with minor modifications (Supplementary Fig. 9). In brief, the PCR2 forward primers described in ref. [35], which have a binding site downstream of the first two amino-acids of the framework sequences, were replaced by the corresponding PCR1 forward primers during the PCR2 reaction.

Sequencing. Final products were sequenced on an Illumina MiSeq (using 2 × 300 bp paired-end reads), which allowed sequencing of the entire V_H and V_L domain as well as the barcode sequence (Supplementary Fig. 9 and Supplementary Table 5).

Bioinformatics data processing. A bioinformatics pipeline was developed to allow sequence read trimming, merging, barcode extraction and clustering, and antibody sequence characterization and filtering. For all barcode clusters that contained productive V_H and V_L made up of at least 40 reads each (a threshold arrived at on the basis of pairing rates from mixtures of mouse hybridomas, see Supplementary Fig. 10a), the most abundant V_H and V_L regions were considered to be paired (Supplementary Fig. 11). V_H–V_L pairs were assumed to derive from the same progenitor rearrangement event if they shared the same V and J gene assignments and had CDR3s of the same length. Somatic mutations were identified by comparing the V_H and V_L sequences to the closest germline V gene. Detailed bioinformatics methods can be found in the Supplementary Methods.

Antibody selection for expression. From the TT sort, 42 V_H–V_L pairs with the highest numbers of reads were selected for expression. recombinations, and that differed from the germline V gene sequence by either 1–4 amino acids, or 4–10 amino acids or >10 amino acid changes. From the GPI sort, 14 V_H–V_L pairs were selected for expression from among several V_H–J_H recombinations that represented various V_H gene families, that associated with (if possible) different V_L–J_L recombinations, and that differed from the germline V gene sequence by either 1–4 amino acids, or 4–10 amino acids or >10 amino acid changes. Four V_H–V_L pairs that presented germline configurations were also selected for antibody expression from the GPI sort. From the 2 TSPAN8 sorts, 20 V_H–V_L pairs were chosen for expression, and candidates were divided into three groups depending on the amino acid differences in comparison to those seen in the closest germline V gene (<4, 4–9, and ≥10 changes in the V-gene region), and in cases of multi-member families choosing only one IgG per family.

Gene synthesis and cloning of antibody V genes. V_H and V_L gene sequences were codon optimized for expression in human cells using the GeneArt Strings DNA optimization software, and optimized DNA fragments were synthesized by GeneArt (Thermo Fisher Scientific). The synthesized V_H and V_L DNA fragments contained a first adapter sequence (5′-ACAGCTACAGGCGCGCACTC-3′, corresponding to the V-gene proximal region of the mouse leader sequence) and a second adapter sequence (5′-GCGTCGACCAAGGGCCCATCG-3′ or 5′-CGTACGGTGGCTGCACCATCT-3′, overlapping with the V-gene proximal region of the mouse CH1 and Cκ regions, respectively) to allow cloning of the V genes. Synthesized DNA fragments were inserted into corresponding pHFB-IgG1 and pHFB-κ vectors (HiFiBiO heavy and light chain expression plasmids) in frame with the human IgG1 or Cκ regions by In-Fusion DNA assembly (Takara Bio) according to the manufacturer's instructions, and plasmids were transformed into *Escherichia coli* Stellar chemically competent cells (Clontech). Plasmids were purified by the mini-prep technique and were analyzed by Sanger sequencing (Wyzer Biosciences).

Antibody expression. Correctly cloned V_H–V_L pairs were cotransfected into serum-free Expi293F suspension cell cultures (Life Technologies) using ExpiFectamine reagents (Life Technologies) according to the manufacturer's instructions. Transfection reactions were incubated for 4 d, at which point supernatants were collected, clarified by centrifugation at 1400 rpm for 15 min at room temperature, and stored at 4 °C. For affinity measurements, IgGs were purified by affinity chromatography using an AKTA pure fast protein liquid chromatography (FPLC) instrument (GE Healthcare) and HiTrap Protein G Column (GE Healthcare). After purification, IgGs were desalted with a HiTrap Desalting Column (GE Healthcare).

ELISA assays. Maxisorp plates were coated with 100 µl of 1 µg ml^{-1} TT or 5 µg ml^{-1} GPI in PBS overnight at 4°. Plates were blocked with PBS containing 1% BSA. Naive, immune mouse sera, IgGs from culture supernatants or control IgGs at at 1 µg ml^{-1} were diluted in series. Anti-human Fc-HRP (horseradish peroxidase) antibody at 0.05 µg ml^{-1} (Bethyl Laboratories) was used for detection. The assay was revealed with ortho-phenylene diamine and was stopped with 2 M H_2SO_4, after which the absorbance at 492 nm was measured.

Affinities of the recombinant anti-TT antibodies identified by CelliGO. EC50s for anti-TT IgGs were measured using the same protocol as described above.

Bio-layer interferometry and real-time surface plasmon resonance. Affinities of the recombinant anti-GPI antibodies identified by Celli*GO*. Bio-layer interferometry with an anti-human IgG sensor on Octet systems (Pall ForteBio) was used to quantify antibody concentrations in supernatants of transfected Expi293F cells, and to measure affinities of the recombinant mouse–human chimeric anti-GPI antibodies. For the latter, biosensors were loaded with 1 ug ml^{-1} antibody, and then association and dissociation signals (10 min per step) were measured against 5 different concentrations of GPI (1.85 nM, 5.5 nM, 16.6 nM, 50nM and 150 nM), in duplicate. Curves were retrieved as raw data from the Octet Data Analysis software and were processed in Scrubber2 prior to analysis with BIAevaluation software (Biacore). IgG-free buffer and an irrelevant IgG antibody (here, an anti-TT IgG) were used as reference and for background subtraction, respectively. Curves were fitted and K_d values were obtained using the 1:1 Langmuir model.

Affinity of serum from mice immunized against TT or immunized against GPI. Real-time surface plasmon resonance on a Biacore T200 instrument (GE Healthcare)

was used to measure, at 25 °C, the mean affinity of mouse serum against TT or GPI. IgG from mouse serum (diluted 1/500) was captured onto a streptavidin sensor chip (GE Healthcare) that was pre-loaded with biotin-conjugated V_HH anti-IgG-Fc multi-species antibodies (Thermo Fisher Scientific). Purified TT or GPI antigens were injected at concentrations ranging from 7.8nM to 1000 nM for 10 min at a rate of 20 µl min^{-1}, and dissociation followed for 10 min. For negative controls, TT antigen was injected onto immobilized IgG from anti-GPI sera and vice-versa, and was used as a reference surface to calculate specific signals. The association and dissociation profiles were analyzed using Biacore T200 evaluation software (v3.0), assuming a 1:1 simple Langmuir interaction mechanism, to provide an average equilibrium dissociation constant (K_d) for the anti-TT and anti-GPI IgGs contained in each serum.

Characterization of anti-TSPAN8 IgG antibodies. *Specificity.* Parental M300.19 cells and TSPAN8$^+$ M300.19 cells were incubated with 10 µg ml^{-1} recombinant IgG identified from TSPAN8 sorts or with 5 µg ml^{-1} anti-TSPAN8 monoclonal antibody clone 43A6 (positive control) in ice-cold FACS buffer (Ca^{2+}-, Mg^{2+}-free PBS supplemented with 3% FBS + 0.1% (wt/vol) NaN_3) for 30 min on ice. Cells were washed twice, incubated with goat anti-human IgG-PE (Southern Biotech, 1030-09) for 30 min on ice, resuspended in BD Cytofix (BD Biosciences 554655) for 30 min on ice and then analyzed on a FACSCalibur flow cytometer (BD Biosciences). Affinity was determined as follows: anti-TSPAN8 IgG antibodies that demonstrated specific binding to TSPAN8$^+$ M300.19 cells were fluorescently labeled with Alexa Fluor 488 and incubated in serial dilutions with TSPAN8$^+$ M300.19 cells in ice-cold FACS buffer (Ca^{2+}, Mg^{2+}-free PBS supplemented with 10% FBS) for 30 min on ice, and cell-bound fluorescence was quantified using a MACSQuant flow cytometer (Miltenyi Biotech). The K_d values were calculated using BIAevaluation software by fitting the data on the antibody-concentration dependence of the fluorescent signal to equation (1):

$$R = \frac{R_{max} \times C}{K_d + C} \quad (1)$$

where the response, R, is the mean fluorescent signal of the tested IgG minus the mean fluorescent signal of a negative-control antibody (rituximab, an anti-human CD20 IgG), C is the antibody concentration and R_{max} is the fluorescent signal at infinite concentration. K_d values were calculated by fitting to the pooled data from triplicate experiments.

Statistical methods developed for V-gene usage comparisons. Pairwise comparisons of matrices of the same dimensions (corresponding to usage of V_H genes, V_L genes or V_H–V_L combinations), as well as metric multidimensional scaling (MDS), were performed using comat, an R tool available at https://gitlab.pasteur.fr/gmillot/comat. Comat was run as follows: we first used a normalization method called 'proportional matrix equalization by adjustment to minimal frequency'. Matrix case frequencies were adjusted so that the total absolute frequencies of the two compared matrices were equal. When $n_1 > n_2$, the equation used was:

$$n_{ij1} = n_{ij1} \times \frac{n_2}{n_1}$$

When $n_2 > n_1$, the equation used was:

$$n_{ij2} = n_{ij2} \times \frac{n_1}{n_2}$$

where n_{ij1} is the absolute frequency of the matrix 1 case at the row i and column j, n_{ij2} is the absolute frequency of the matrix 2 case at row i and column j, n_1 is the total absolute frequency of matrix 1 and n_2 is the total absolute frequency of matrix 2. Such normalization was performed to focus the pairwise comparison on the matrix distributions (that is, matrix proportions), without influence related to n_1 and n_2 differences, which depends on uncontrolled experimental fluctuations. This normalization method loses information, but it provides two matrices of similar experimental sensitivity. Then, the observed distance, D_{obs}, between the matrices was computed. The distance, D, between the two matrices was calculated using equation (2):

$$D = \left(\sum_{i=1}^{k} \sum_{j=1}^{c} \left| \frac{n_{ij1}}{n_1} - \frac{n_{ij2}}{n_2} \right| \right) / 2 \quad (2)$$

where k is the number of rows and c is the number of columns. Then, an approximate permutation test was developed. To approximate the distribution of the D statistic, the n_1 frequencies were randomly spread in matrix 1, and the same was done for the n_2 frequencies in matrix 2 (here, n_1 and n_2 are equal due to normalization). Then, a theoretical statistic was computed using equation (2). This was repeated $r = 100{,}000$ times, which provided a probability distribution of D, and thus, the probability to obtain the observed D_{obs} statistic. In a two-sided test, the P value was calculated as:

$$P = \Pr(D \geq D_{obs}) \quad (3)$$

If $P = 0$, then the P value was estimated as $P < 1/r$. The same procedure was performed to compare the V-gene frequencies (sum of the rows of two matrices of the same dimensions). Statistical significance was set to $P \leq 0.05$. Type I error was controlled by correcting the P values according to the Benjamini–Hochberg method ('BH' option in the p.adjust() function of R)[36]. MDS results were obtained using the R package ExPosition[37]. To generate the 95% confidence intervals shown for the MDS value, bootstrap of the initial combined V_H–V_L matrix usage (size n sampled in the non-empty cases of the matrix) was performed 100 times, the MDS values for the 100 new distance matrices were added to the original MDS value, and quantiles 0.025 and 0.975 of the 100 new positions of each initial dot were computed for each dimension.

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

Data availability

Processed sequencing files (quality trimmed, merged and barcode assigned) for all data generated in this study were deposited at Sequence Read Archive (SRA) under accession PRJNA529803.

References

31. Baret, J. C. et al. Fluorescence-activated droplet sorting (FADS): efficient microfluidic cell sorting based on enzymatic activity. *Lab-on-a-Chip* **9**, 1850–1858 (2009).
32. Duffy, D. C., McDonald, J. C., Schueller, O. J. & Whitesides, G. M. Rapid prototyping of microfluidic systems in poly(dimethylsiloxane). *Anal. Chem.* **70**, 4974–4984 (1998).
33. Siegel, A., Bruzewicz, D., Weibel, D. & Whitesides, G. Microsolidics: fabrication of three-dimensional metallic microstructures in poly(dimethylsiloxane). *Adv. Mater.* **19**, 727–733 (2007).
34. Zilionis, R. et al. Single-cell barcoding and sequencing using droplet microfluidics. *Nat. Protoc.* **12**, 44–73 (2017).
35. Rohatgi, S., Ganju, P. & Sehgal, D. Systematic design and testing of nested (RT-)PCR primers for specific amplification of mouse rearranged/expressed immunoglobulin variable region genes from small number of B cells. *J. Immunol. Methods* **339**, 205–219 (2008).
36. R Core Team (R Foundation for Statistical Computing, Vienna, Austria, 2017).
37. Beaton, D., Chin Fatt, C. R. & Abdi, H. An ExPosition of multivariate analysis with the singular value decomposition in R. *Computational Stat. Data Analysis* **72**, 176–189 (2014).

Acknowledgements

We would like to thank M. Holsti and W. Somers (Pfizer, BioMedicine Design) for advice on the technology development and critical reading of the manuscript; R. Nicol (Whitehead Institute, MIT Center for Genome Research) for advice on barcoded sequencing; at the Institut Pasteur, H. Mouquet for advice on antibody sequence selection, and F. Jönsson for advice on flow cytometry analysis; T. Kirk and S. Foulon (ESPCI Paris) for their help developing the barcoded hydrogel beads; S. N. Stewart (HiFiBio Therapeutics Inc.) for her help producing and analysing recombinant antibodies; the Institut Pierre-Gilles de Gennes (IPGG) for use of clean room facilities and the laser engraver (CII08, Axyslaser). This work was supported by the French Agence Nationale de la Recherche (ANR-14-CE16-0011 project DROPmAbs), by the Centre d'Innovation et Recherche Technologique (Citech) through the Institut Carnot Pasteur Microbes & Santé (ANR 16 CARN 0023-01), by BPI France (OSIRIS and CELLIGO projects) and by the French 'Investissements d'Avenir' program via the CELLIGO project and grant agreements ANR-10-NANO-02, ANR-10-IDEX-0001-02 PSL, ANR-10-LABX-31 and ANR-10-EQPX-34. NGS was performed by the ICGex NGS platform of the Institut Curie and the Institut de Biologie Intégrative de la Cellule (I2BC) platform (Gif-sur-Yvette) supported by the grants ANR-10-EQPX-03 and ANR-10-INBS-09-08 (France Génomique Consortium), by the Canceropole Ile-de-France and by the SiRIC-Curie program (SiRIC Grant INCa-DGOS- 4654). C.C. acknowledges financial support from CONCYTEC, Peru. P.C.H. is a scholar in the Pasteur - Paris University (PPU) International PhD program and was also supported by the Fondation pour la Recherche Médicale (FRM; FDT201904008240). K.E. acknowledges financial support from the Swiss National Science Foundation and The Branco Weiss Fellowship - Society in Science.

Author contributions

C.B., P.B., A. Gérard, A.J. and A.D.G. designed and supervised the study; C.B., P.B. and A.D.G. secured funding; L.B.-R., P.C.-H., C.C., M.D., R.D., K.E., S. Elllouze., S. Essono., A. Godina, K.G., B.I., B.J., R.K., V.M., P.M., S.M., G.M., C.O., Y.P., A.P., M.R., O.R.-L., G.R., A.S.-C., B. Saudemont., B. Shen., S.N.S. and A.W. performed the experiments; J.B., C.B., P.B., P.E., A. Gérard, A.D.G., A.J., G.A.M., G.M., C.N. and A.W. analyzed the data, and P.B., A. Gérard, A.D.G. and A.W. wrote the paper.

Competing interests

A.D.G. and C.B. are co-founders of Seven Pines Holding BV, and HiFiBiO Therapeutics SAS is a subsidiary of Seven Pines Holding BV. Patents have been filed on some aspects of this work and future development of the platform, and the inventors may receive payments related to exploitation of these under their employer's rewards to inventors scheme.

Additional information

Supplementary information is available for this paper at https://doi.org/10.1038/s41587-020-0466-7.

Correspondence and requests for materials should be addressed to A.D.G., P.B. or C.B.

nature research

Corresponding author(s): Andrew D. Griffiths, Pierre Bruhns, Colin Brenan

Last updated by author(s): Nov 20, 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	Droplet data acquisition and sorting: - LABVIEW custom software coded by Andrew Griffiths laboratory - available on request for academic collaborations OR - HiFiBio proprietary software: The software uDROP used for droplet analyses and sorting was designed by HiFiBio, or was a custom-written Matlab script. The HiFiBiO B cell screening process is called CelliGO, which is not a commercialized product – and some aspects of it are proprietary. People can use the provided microfluidics design and purchase components separately as described in the manuscript (microscopes, compressors, pumps). Flow cytometry: BD Cell quest
Data analysis	HiFiBio proprietary software: The repertoire analyses were performed using the HiFiBio proprietary software "Absolution". The "Comat" tool is freely available under the terms of the GNU General Public License as published by the Free Software Foundation. Flow cytometry analysis: FlowJo Version 9 Graphical representation: GraphPad Prism Ver 7.0.

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

All data is provided in the manuscript supplemental information and files. All the other data that support the findings of this study are available from the corresponding authors upon 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	The sample size in all experiments ensure a high statistical power to support the observed effect (>95 %) based on null hypothesis testing.
Data exclusions	No data was excluded.
Replication	All experiments were replicated across the two or three replicates. The findings are consistent across the replicates.
Randomization	No method of randomization was chosen.
Blinding	Data collection and analysis were not performed in a blinded manner.

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

anti-IgM (Miltenyi, Cat#130-095-904)

anti-IgD (Miltenyi, Cat#130-101-899)

biotinylated VHH anti-mouse kappa light chain (captureselect, Thermo Fisher #7103152100)

rabbit F(ab')2 anti-mouse IgG Fc-specific (AlexaFluor647-labeled, Jackson ImmunoResearch)

anti-Tetanus Toxoid antibody clone TT7 (DeKosky, B. J., et al. (2013). "High-throughput sequencing of the paired human immunoglobulin heavy and light chain repertoire." Nat Biotechnol 31(2): 166-169.)

mouse anti-horseradish peroxidase antibody (Jackson Immuno Research, Cat#223-005-024)

anti-hTSPAN8 51D3 IgG (Pfizer Inc)

anti-hTSPAN8 43A6 IgG (Pfizer Inc)

biotinylated anti-IgG mouse mAb (Mabtech #3825-2AW-Plus)

Anti-human Fc-HRP (Bethyl)

biotin-conjugated VHH anti-IgG-Fc multi species antibodies (ThermoFisher)

goat anti-human IgG-PE (Southern Biotech #1030-09)

Validation

anti-IgM (Miltenyi, Cat#130-095-904): validated by the manufacturer

anti-IgD (Miltenyi, Cat#130-101-899): validated by the manufacturer

biotinylated VHH anti-mouse kappa light chain (captureselect, Thermo Fisher #7103152100): validated by the manufacturer and in this reference: Eyer, K., et al. (2017). "Single-cell deep phenotyping of IgG-secreting cells for high-resolution immune monitoring." Nat Biotechnol 35(10): 977-982.

rabbit F(ab')2 anti-mouse IgG Fc-specific (AlexaFluor647-labeled, Jackson ImmunoResearch): validated by the manufacturer and in this reference: Eyer, K., et al. (2017). "Single-cell deep phenotyping of IgG-secreting cells for high-resolution immune monitoring." Nat Biotechnol 35(10): 977-982.

anti-Tetanus Toxoid antibody clone TT7: validated in this reference: DeKosky, B. J., et al. (2013). "High-throughput sequencing of the paired human immunoglobulin heavy and light chain repertoire." Nat Biotechnol 31(2): 166-169.

mouse anti-horseradish peroxidase antibody (Jackson Immuno Research, Cat#223-005-024): validated by the manufacturer

anti-hTSPAN8 51D3 IgG (Pfizer Inc): validated by Pfizer Inc, and used as a positive control in this manuscript

anti-hTSPAN8 43A6 IgG (Pfizer Inc): validated by Pfizer Inc, and used as a positive control in this manuscript

biotinylated anti-IgG mouse mAb (Mabtech #3825-2AW-Plus): validated by the manufacturer

Anti-human Fc-HRP (Bethyl): validated by the manufacturer

biotin-conjugated VHH anti-IgG-Fc multi species antibodies (ThermoFisher): validated by the manufacturer

goat anti-human IgG-PE (Southern Biotech #1030-09): validated by the manufacturer

Eukaryotic cell lines

Policy information about cell lines

Cell line source(s)

- m300.19 mouse pre-B cells: LONZA (https://knowledge.lonza.com/cell?id=138)

- Chinese hamster ovary (CHO) cells: ATCC (https://www.lgcstandards-atcc.org/products/all/CCL-61.aspx?geo_country=fr)

- Expi293F suspension cells: Life Technologies (https://www.thermofisher.com/order/catalog/product/A14527?SID=srch-srp-A14527#/A14527?SID=srch-srp-A14527)

- Jurkat human T cell line: ATCC (https://www.lgcstandards-atcc.org/products/all/TIB-152.aspx)

- Jurkat reporter cells stably expressing GFP under control of the NFkB responsive promoter: SYSTEMBIO (https://systembio.com/products/imaging-and-reporter-vectors/signaling-pathway-reporters/nf-kb-jurkat-gfp-transcriptional-reporter-cell-line/)

- Ramos human B cell line: ATCC (https://www.lgcstandards-atcc.org/Products/All/CRL-1596.aspx?geo_country=fr)

Authentication

Cell lines were not independently authenticated.

Mycoplasma contamination

All cell lines tested negative for mycoplasma contamination.

Commonly misidentified lines
(See ICLAC register)

No commonly misidentified cell lines were used.

Animals and other organisms

Policy information about studies involving animals; ARRIVE guidelines recommended for reporting animal research

Laboratory animals

BALB/c mice purchased from Janvier Labs.

Wild animals

N/A

Field-collected samples	N/A
Ethics oversight	Experiments using mice were validated by i) the CETEA ethics committee number 89 (Institut Pasteur, Paris, France) under #2013-0103, and by the French Ministry of Research under agreement #00513.02, and ii) the French ethics committee number 80 (CellVax) under reference APAFIS#6063-2016071213009170 v5.

Note that full information on the approval of the study protocol must also be provided in the manuscript.

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	Parental M300.19 cells and TSPAN8+ M300.19 cells were incubated with 10 µg/mL recombinant IgG identified from TSPAN8 sorts or with 5 µg/mL anti-TSPAN8 mAb clone 43A6 (positive control) in ice-cold FACS Buffer (Ca2+, Mg2+-free PBS supplemented with 3% FBS + 0.1% w/v NaN3) for 30 min on ice. Cells were washed twice, incubated with goat anti-human IgG-PE (Southern Biotech #1030-09) for 30 min on ice, and resuspended in BD Cytofix (BD Biosciences #554655) for 30 min on ice, and analyzed on a FACSCalibur flow cytometer (BD)
Instrument	FACSCalibur flow cytometer (Becton Dickinson)
Software	For data acquisition BD CellQuest; For data analysis FlowJo version 9
Cell population abundance	100% (cell lines)
Gating strategy	N/A

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