

Multiplex profiling of serum proteins in solution using barcoded antibody fragments and next generation sequencing

Mattias Brofelth ^{1,2,6}, Anna Isinger Ekstrand^{1,2,6}, Shashank Gour³, Ronnie Jansson³, My Hedhammar³, Björn Elleby⁴, Anders Kvist ⁵, Christer Wingren^{1,2}, Ulrika Axelsson ^{1,2} & Carl A. K. Borrebaeck ^{1,2}✉

The composition of serum proteins is reflecting the current health status and can, with the right tools, be used to detect early signs of disease, such as an emerging cancer. An earlier diagnosis of cancer would greatly increase the chance of an improved outcome for the patients. However, there is still an unmet need for proficient tools to decipher the information in the blood proteome, which calls for further technological development. Here, we present a proof-of-concept study that demonstrates an alternative approach for multiplexed protein profiling of serum samples in solution, using DNA barcoded scFv antibody fragments and next generation sequencing. The outcome shows high accuracy when discriminating samples derived from pancreatic cancer patients and healthy controls and represents a scalable alternative for serum analysis.

¹Department of Immunotechnology, Lund University, Medicon Village, Lund, Sweden. ²CREATE Health Translational Cancer Center, Lund University, Medicon Village, Lund, Sweden. ³Department of Protein Science, School of Engineering Sciences in Chemistry, Biotechnology and Health, KTH Royal Institute of Technology, Stockholm, Sweden. ⁴Immunovia AB, Medicon Village, Lund, Sweden. ⁵Division of Oncology and Pathology, Department of Clinical Sciences Lund, Lund University Cancer Center, Lund University, Lund, Sweden. ⁶These authors contributed equally: Mattias Brofelth, Anna Isinger Ekstrand. ✉email: carl.borrebaeck@immun.lth.se

Human serum is a complex proteome to analyze, providing major technological challenges. However, mining the serum proteome for differentially expressed molecular biomarkers provides an attractive and minimally invasive way for precision diagnostics¹. Planar antibody microarray is one of the technologies in the forefront^{2–5} and has delivered clinically actionable information for differential and early diagnosis of cancer^{6–8}. Although highly sensitive for multiplexed protein expression profiling, planar antibody arrays strive with inherent limitations such as surface performance, signal-to-noise ratio, limit of detection, dynamic range, and printing logistics.

A solution-based platform could circumvent several limitations but has so far not been developed for the serum proteome to achieve both the necessary sensitivity and scalability. Conventional technologies are limited in target multiplexity, partly by the need of multiple antibodies per target analyte⁹. Alternative approaches have, however, been developed in recent years utilizing antibody–DNA conjugates allowing multiplexed protein analysis of fine needle aspirate using NanoString nCounter¹⁰, high-throughput phenotyping of cells using next-generation sequencing (NGS)^{11–14}, as well as more focused approaches using, e.g., DNA-binding factors¹⁵. Assays can, however, be designed using multi-well plates in automated systems for parallel and consistent serum analysis in solution, which in combination with NGS could reach ultra-high sensitivity.

Here we present a proof-of concept study for profiling serum from pancreatic cancer patients, using ProMIS, *Protein detection using Multiplex Immunoassay in Solution*. ProMIS is a streamlined platform for profiling of serum proteins with a solution-based bead array. The assay utilizes antibody fragments (scFv) that are site specifically conjugated to DNA oligonucleotide barcodes, in a 1:1 manner, using a Sortase A-mediated coupling

strategy. The barcoded scFvs are mixed with biotinylated serum proteins coupled to streptavidin-coated magnetic beads, and bound antibodies are detected, using NGS allowing for both a multiplex and sensitive read-out.

Results

The ProMIS concept. The concept is based on immobilization of serum proteins onto magnetic beads, followed by target binding of DNA barcoded scFv antibodies and a subsequent PCR step prior to detection by NGS (Fig. 1a). Proof-of-concept was consecutively established by selecting 17 scFv antibodies (Table 1), of which 14 have been previously reported to discriminate between serum samples derived from patients with pancreatic ductal adenocarcinoma (PDAC) and healthy controls⁷, while the three additional scFvs provide orthogonal information (unpublished data). All scFv antibodies were redesigned and produced with a C-terminal Sortase A recognition motif (LPETG), resulting in a typical yield of 1–5 mg/L. A protocol was established to conjugate the LPETG-tagged scFv antibodies to tri-glycine-modified barcode oligonucleotides, containing a tag sequence unique to each scFv. A subsequent purification step, using filtration with 30 kDa cutoff, allowed us to include only pure scFv-oligo in the assay (Fig. 1b). The use of Sortase A to conjugate oligonucleotides to scFv antibodies enables the crucial and necessary site-specific barcoding, with a 1:1 ratio. Successful conjugation and purification were confirmed by gel electrophoresis (Supplementary Fig. 1).

Serum samples from PDAC patients and controls were biotinylated and allowed to bind to streptavidin-coated magnetic beads. Unbound serum proteins were removed by thorough washing of the beads before a cocktail with excess of each of the 17 barcoded scFv antibodies was added to each

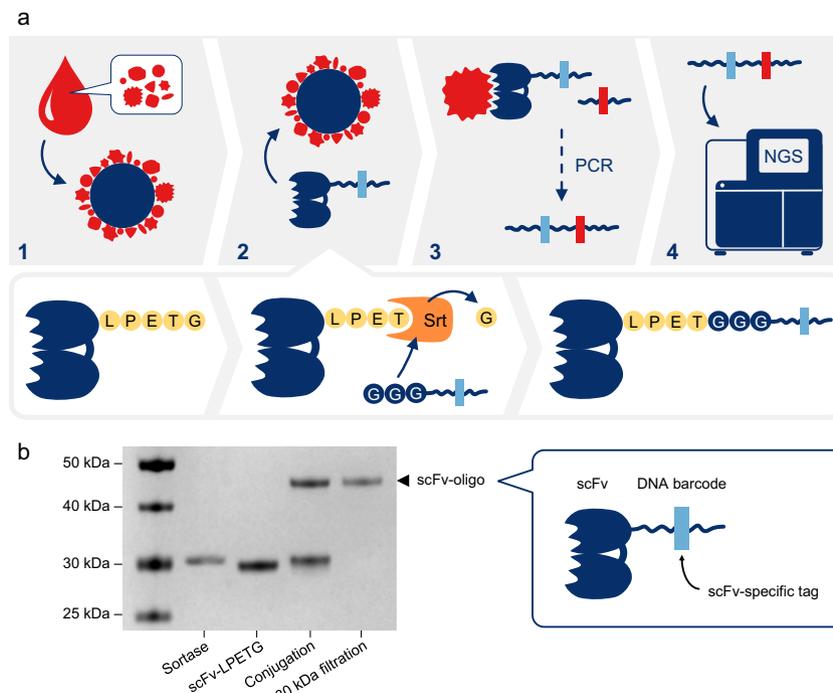


Fig. 1 The concept of the ProMIS assay. **a** Assay principles (1–4) and the conjugation of scFv with oligonucleotides, using Sortase A (Srt). (1) Biotinylated serum proteins are captured and displayed on streptavidin-coated magnetic beads. (2) Recombinant antibody fragments (scFvs) are site-specifically conjugated 1:1 with a unique DNA oligo containing a scFv-specific tag, using a Sortase A-mediated coupling strategy. The scFv-oligos are then mixed with the beads coated with serum proteins. (3) After washing, adapter PCR is performed to equip bound scFv-oligos with a sample-specific DNA tag. (4) PCR products obtained from the combined scFv and sample tags are finally analyzed, using NGS. **b** SDS-PAGE analysis of conjugation of scFvs (containing the Sortase A recognition motif LPETG) to tri-glycine modified oligonucleotides, followed by filtration (30 kDa cutoff) to isolate only the conjugated scFv-oligos.

Table 1 scFv antibody fragments and corresponding target antigens.

#	Antibody	Target antigen	Target antigen (full name)
1	scFv(1)-Srt	IL-4	Interleukin-4
2	scFv(2)-Srt	MCP-1	C-C motif chemokine 2
3	scFv(3)-Srt	Lewis x	Lewis x
4	scFv(4)-Srt	Sialyl Lewis x	Sialyl Lewis x
5	scFv(5)-Srt	C1q	Complement C1q
6	scFv(6)-Srt	C5	Complement C5
7	scFv(7)-Srt	C1 est inh	Plasma protease C1 inhibitor
8	scFv(8)-Srt	Properdin	Properdin
9	scFv(9)-Srt	VEGF	Vascular endothelial growth factor
10	scFv(10)-Srt	IL-4	Interleukin-4
11	scFv(11)-Srt	C1 est inh	Plasma protease C1 inhibitor
12	scFv(12)-Srt	C5	Complement C5
13	scFv(13)-Srt	CDK2	Cyclin-dependent kinase 2
14	scFv(14)-Srt	HADH2	HADH2 protein
15	scFv(15)-Srt	APLF	Aprataxin and PNK-like factor
16	scFv(16)-Srt	MARK1-1	Serine/threonine-protein kinase MARK1
17	scFv(17)-Srt	PRKCZ	Protein kinase C zeta type

sample for multiplex target detection. After incubation and washing, the scFv-oligos bound to the beads were PCR amplified with primers adding Illumina-compatible adapters and sample-specific indexes. The final PCR product thus contained both the scFv-specific barcode for quantification of binding events and an index to identify the sample. This allowed us to pool all the PCR products from all samples in a single NGS run on an Illumina NextSeq 500 system (Fig. 1a). The sequencing data were quality filtered and demultiplexed into counts for each scFv antibody in each sample, providing a direct digital readout. Data analysis was performed on median normalized and log₂ transformed counts for supervised classification, using support vector machine (SVM) and leave-one-out (LOO) cross-validation to generate receiver operating characteristic (ROC) curves and area under the curve (AUC) values. In addition, the data were analyzed with unsupervised principal component analysis (PCA).

Analyzing pancreatic cancer samples. The performance of ProMIS was demonstrated, using samples from cohorts that had previously been discriminated with ROC-AUCs of >0.90 when classifying PDAC stage I–IV versus healthy controls, using IMMray[®] antibody microarrays⁷. A first test was performed using 20 samples (10 PDAC stage IV and 10 healthy controls) and 16 scFv-oligos. All sequences corresponding to each specific scFv-oligo could be detected in each sample, demonstrating the functionality. The samples grouped according to disease status (healthy versus PDAC) in the unsupervised PCA plot (Fig. 2a) and resulted in a ROC–AUC of 0.82 using SVM analysis (Fig. 2b). The experiment was repeated in a second test with 20 independent samples (10 PDAC stage IV and 10 controls) analyzed with 17 scFv-oligos (one more was available at that time). Again, all specific scFv-oligos were detected, and cases and controls could be separated by PCA (Fig. 2c) and resulted in a ROC–AUC of 0.86 (Fig. 2d). The increased ROC–AUC value might be attributed to the supplementary information by the additional scFv. A final proof-of-concept was performed with 40 independent

samples, including 20 PDAC, this time including samples from 10 stage III and 10 stage IV PDAC patients and 20 healthy controls. Again, all 17 specific scFv-oligos could be detected and the cases and controls were separated in PCA (Fig. 2e) and resulted in a ROC–AUC value of 0.90 (Fig. 2f).

To study the intra-assay precision of each scFv antibody, serum samples were reanalyzed using all 17 scFv-oligos with 10 replicates of each sample (Supplementary Fig. 2a). The results showed a high reproducibility with a coefficient of variation (CV) below 1% for the majority of the antibodies (Supplementary Fig. 2b).

Discussion

This proof-of-concept study exemplifies the ability of the ProMIS platform for multiplex analysis of the human serum proteome, which provides an unparalleled approach in precision diagnostics of complex diseases. Key to the concept is the oligonucleotide barcoding of the scFv antibodies that enables the sequencing readout. The enzymatic reaction with Sortase A was found to be an effective and convenient strategy for conjugating scFv with a barcode in a 1:1 ratio under protein-compatible reaction conditions. Site-specific conjugation avoids the risk of blocking the antigen binding site, which becomes a challenge in non-specific conjugation methods. Multi-barcoding would limit the performance, which is also avoided with the 1:1 scFv:oligo conjugation principle of ProMIS.

In terms of multiplexity, the number of recombinant scFv antibodies carrying the Sortase recognition motif can easily be expanded to target virtually any antigen, using phage-display library-derived antibody fragments. The scFv antibodies used here were selected based on their combined power to discriminate PDAC versus healthy samples in antibody microarrays⁷ and demonstrated a similar performance also in ProMIS. In all three assay runs tested, each with independent serum samples, this set of scFv antibodies detected the biological differences separating the two groups with similar accuracy. Together with the low technical variation (CV < 1%), this is an indication of the robustness of the ProMIS assay.

The solution-based assay allows for easy conversion to a multi-well plate format that facilitates automation of the assay steps, which would provide consistent performance and a high sample throughput. In addition, with the PCR amplification step in combination with the NGS detection, it should be possible to tune the assay for excellent sensitivity. However, the sensitivity of the ProMIS assay will be dependent on the context of, e.g., specific biomarker signature and the number of samples that is analyzed in parallel, using a given NGS kit size (i.e. number of reads). In the present context, the sensitivity of ProMIS was shown to be comparable to an antibody microarray, using the same scFvs, demonstrating a LOD in pM–fM range^{16,17}. The rapid development of NGS technology is already resulting in kits with increased coverage and decreased costs, which will catalyze how ProMIS can perform multiplexed analysis without a reduction in dynamic range.

In conclusion, we present a proof-of-concept study of the ProMIS platform, which has the potential to analyze differentially expressed proteins in serum samples with a higher throughput, multiplexity, and sensitivity, thus circumventing some of the inherent limitations with planar microarrays in precision diagnostics of complex diseases.

Methods

Sample preparation and capture to beads. Human serum samples from patients with PDAC (stage III or IV) and negative controls were collected by Oregon Health & Science University (OHSU), USA. All procedures were in accordance with the Institutional Review Board of Oregon Health and Science University approval.

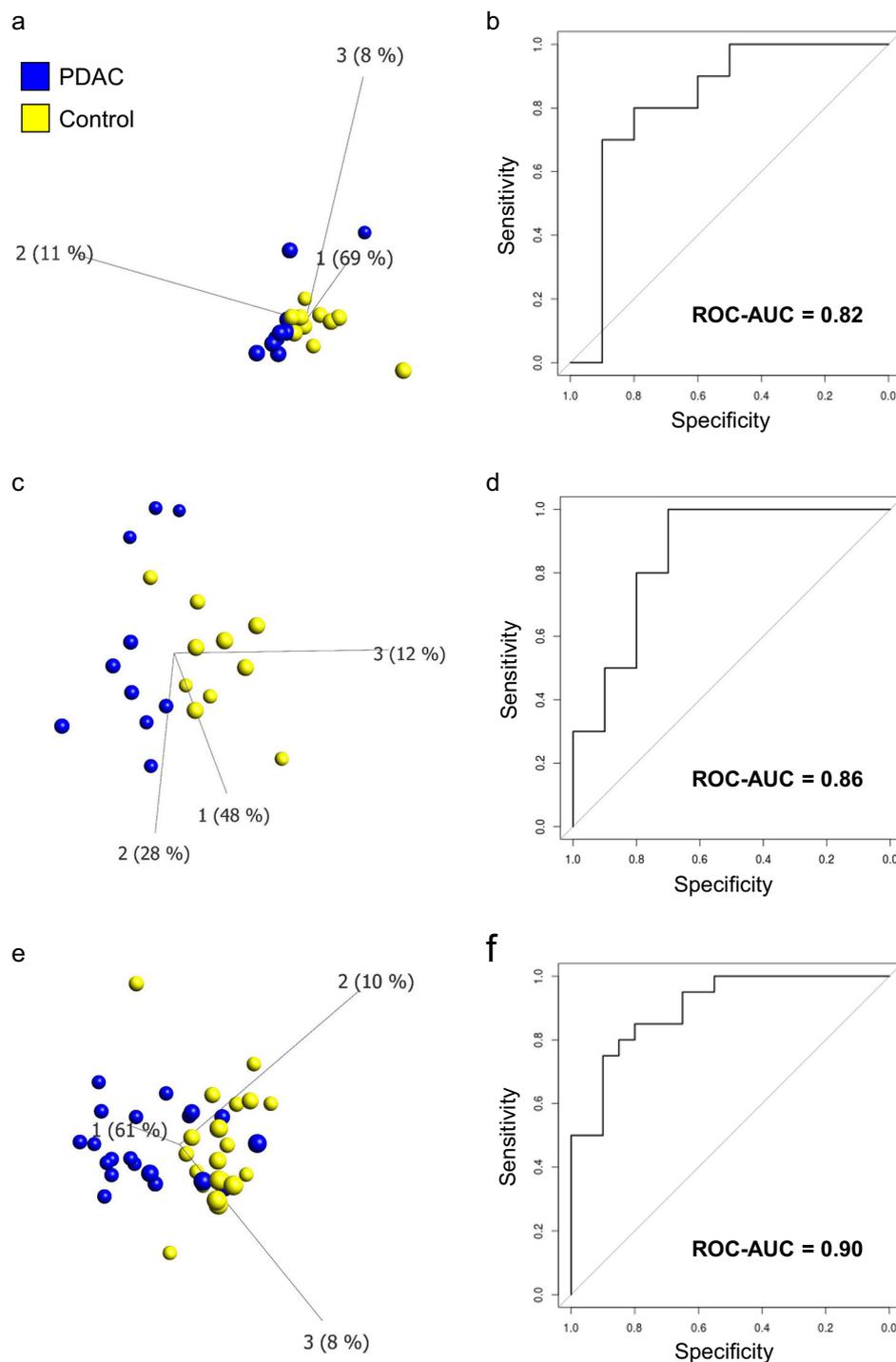


Fig. 2 ProMIS test results from the analysis of serum samples from patients with pancreatic ductal adenocarcinoma (PDAC) compared to healthy controls. Three independent tests were performed on in total 80 distinct serum samples. Test 1: 10 PDAC samples (stage IV) versus 10 healthy controls with 16 scFv-oligos. Test 2: 10 PDAC samples (stage IV) versus 10 healthy controls with 17 scFv-oligos. Test 3: 20 PDAC samples (10 stage III and 10 stage IV) versus 20 healthy controls with 17 scFv-oligos. PCA: unsupervised visualization using principal component analysis (PCA) plotted with individual serum samples. SVM: supervised two-group classification, using support vector machine (SVM) leave-one-out cross-validation. The results are presented with receiver operating characteristic curves (ROC) and their corresponding area under the curve (AUC) values. **a** Test 1 (PCA), **b** Test 1 (SVM), **c** Test 2 (PCA), **d** Test 2 (SVM), **e** Test 3 (PCA), and **f** Test 3 (SVM).

The serum proteins were biotinylated, according to a previously described protocol^{18,19}. Biotinylation is widely used for protein applications, such as immobilization and functionalization^{16,18,20,21} and to verify that biotinylation of the serum sample did not considerably result in misrepresentation of the actual protein distribution we performed an analysis, using immunoprecipitation and mass spectrometry. The results showed a similar number of matching peptides for the target antigen both before and after biotinylation. In brief, 5 μ L of serum

samples were diluted 1:45 in phosphate-buffered saline (PBS) and labeled with 0.6 mM EZ-Link Sulfo-NHS-LC-Biotin (Thermo Fisher Scientific). Unbound biotin was removed by dialysis against PBS for 72 h using a 3.5 kDa MWCO Slide-A-Lyzer MINI dialysis unit (Thermo Fisher Scientific), changing buffer every 24 h. The labeled serum samples were aliquoted and stored at -80°C . In the assay, 75 μ L of streptavidin-coated magnetic beads, Dynabeads M-280 (Life Technologies), were used to immobilize and display 1 μ L of biotinylated serum proteins.

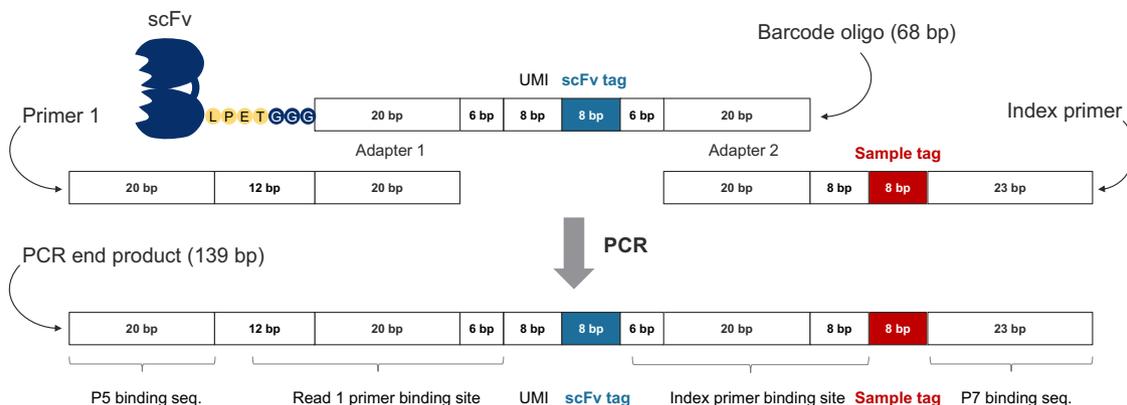


Fig. 3 Barcode oligo design and adapter PCR. The oligonucleotide barcode contains a scFv-specific tag and is conjugated to the scFv using Sortase A. After binding to the target, the barcode is extended in both ends in the adapter PCR step with two primers. Primer 1 contains the P5 sequence needed for binding to the NGS flow cell and the Read 1 sequencing primer-binding site. The Index primer contains the index sequencing primer site, the index (sample tag), and the P7 sequence needed for binding to the NGS flow cell. The sample tag allows pooling of multiple samples and post-NGS demultiplexing of reads.

Generation and production of scFv-LPETG antibody fragments. Seventeen single-chain fragment variable antibodies were selected from in-house designed large phage-display libraries^{22,23}. The specificity of antibodies from the libraries were previously validated with well-characterized serum samples (including spiking, blocking, and depletion of antigens) on antibody microarrays and several orthogonal methods such as mass spectrometry, enzyme-linked immunosorbent assay, and MesoScaleDiscovery cytokine assay, using various samples^{16,18,24–30}.

The scFvs were used as templates in PCR reactions with primers introducing an N-terminal *Nco*I restriction endonuclease site, and a C-terminal (GS)₃-Srt-XhoI (Srt = LPETG, Sortase tag) sequence. The generated PCR products were further used for insertion into a pET-26b(+) vector (Novagen), harboring an N-terminal pelB signal sequence and a C-terminal His₆ tag, generating the 17 scFv gene constructs pelB-scFv-(GS)₃-Srt-His₆. The final gene constructs were verified by DNA sequencing.

All constructs were transformed into *Escherichia coli* BL21(DE3) cells (Merck Biosciences) and produced as previously described¹⁸. In brief, O/N cultures of *E. coli* were grown in 2xYT medium with appropriate antibiotics at 37 °C and induced with 1 mM isopropyl β-D-1-thiogalactopyranoside when OD reached 0.6–0.9. After O/N expression, the antibody fragments were harvested by centrifugation, lysed, and then purified using His MultiTrap 96-well filter plates (GE Healthcare).

Amicon Ultra 10K 0.5 mL centrifugal filters (Merck Millipore) were used both to change the buffer to 450 μL of Sortase ligation buffer (50 mM Tris, 150 mM NaCl, 10 mM CaCl₂, pH 7.5)³¹ and to concentrate the purified scFvs. Purity and concentration were evaluated using 10% SDS-PAGE (Invitrogen) and a Nanodrop-1000 spectrophotometer at 280 nm (Thermo Scientific).

Design of oligonucleotide sequences. The oligonucleotide sequences (68 bp) were designed to include an 8 bp scFv-specific barcode sequence (position 35–42) used to count binding events between scFv-oligo and the target protein (Fig. 3). The sequences of all oligonucleotide barcodes are presented in Supplementary Table 1. The oligonucleotides were designed to carry a tri-glycine (G–G–G) modification in the 5'-end for the Sortase A-mediated conjugation and were purchased from Biomers AG (Ulm, Germany).

Sortase-mediated conjugation of scFv-Srt-His6 antibodies and oligonucleotides. The oligonucleotides, carrying a tri-glycine (G–G–G) modification in the 5'-end, were used for site-specific, enzyme-dependent conjugation to scFv-Srt-His₆. 0.2 nmol (2 μM) of scFv-Srt-His₆ antibodies were mixed with 2 nmol (20 μM) oligonucleotides and 0.1 nmol (1 μM) high-activity mutant Sortase A in sortase ligation buffer (100 μL total reaction volume). The conjugation mixtures were incubated for 2 h at 4 °C. To purify the conjugated scFv-oligos, the conjugation mixtures were added to Amicon Ultra 30 K 0.5 mL centrifugal filters (Merck Millipore) and washed five times with 400 μL PBS. Purity and concentration was evaluated using 10% SDS-PAGE (Invitrogen) and a Nanodrop-1000 spectrophotometer at 280 nm (Thermo Scientific). A cocktail was then created by mixing 85 μL from each of the 17 purified scFv-oligos.

ProMIS assay using barcoded scFvs. Proof-of-concept for the ProMIS assay was demonstrated, using 17 Sortase A-conjugated scFv-oligos in three experiments, two with 20 serum samples and one with 40 serum samples.

In a first step of the assay, 5 μL of biotinylated serum sample was diluted in 20 μL PBS. Five microliters of the diluted serum sample was then mixed with 75 μL of streptavidin-coated magnetic beads in 1.5 mL tubes (1 tube/sample) and incubated for 30 min in room temperature using gently agitation, according to the manufacturer's recommendation. To remove any unbound proteins, the bead/samples were washed

four times with 100 μL of washing buffer (PBS + 0.05% (v/v) Tween-20) by placing the tubes in a magnetic holder for 2 min per washing round.

Next, 32 μL of the cocktail containing a mix of all scFv-oligos was added to each tube with bead/sample and the scFv-oligos were allowed to bind their targets during an incubation at 4 °C for 2 h with gentle agitation. After three rounds of washing with washing buffer, the beads were resuspended in 50 μL of nuclease-free water and used for PCR and NGS.

Library preparation and NGS. For adapter PCR, 8 μL of each sample was mixed with 1× Phusion Master Mix (Thermo Scientific #F-531), 0.5 μM Illumina adapter, index primer (corresponding to each sample), and nuclease-free water in a total volume of 20 μL. PCR program: 98 °C 2 min; 18 repeats of: 98 °C 20 s, 65 °C 30 s, 72 °C 30 s; 72 °C 5 min; 10 °C. PCR product purification was performed using Agencourt® AMPure XP beads according to the manufacturer's recommendation (1.8 ratio). Positive controls contained pure barcode oligos (no scFv) and negative control (water). Quality control of purified PCR products was done using Bioanalyzer and Agilent High Sensitivity DNA kit. Five microliters of each sample was pooled, diluted, and prepared for sequencing on a NextSeq 500/550 High Output v2.5 kit (Illumina) on a NextSeq 500 sequencer (Illumina).

NGS data analysis. NGS raw data (BCL files) generated from the NextSeq 500 were demultiplexed by the sample index reads using bcl2fastq2 Conversion Software v2.20 (Illumina) and ran through an in-house pipeline written in Java programming language to count the total number of scFv-specific tags for each sample. For our analysis we used only reads that passed the sequencer chastity filter and had base call quality for each base over Q30 (Phred Quality score)^{32,33}.

Next, the counts were median normalized and log₂ transformed before two-group classification using SVM LOO cross-validation to generate ROC curves and AUC values. The SVM analysis is a supervised machine learning algorithm and was performed with the R package “e1071” and a linear kernel. No prior data filtration was done before the SVM, i.e., all scFv antibodies used in the assay were also included in the analysis. The SVM finds an optimal hyperplane that separates the two groups and the classification performance is measured by the ROC–AUC value, where the value 1 would mean a perfect classifier and 0.5 a random classifier.

Data were also analyzed using PCA in Qlucore Omics Explorer 3.5 (Qlucore AB, Lund, Sweden). PCA was used as an unsupervised method to reduce the dimensionality and allow visual interpretation of the data in a 3D-plot.

Statistics and reproducibility. When analyzing the serum samples from PDAC patients and healthy controls, no replicates were used in order to maximize the number of parallel samples. Instead, a dedicated intra-assay precision study was performed using four individual biological serum samples (two PDAC and two healthy controls) that each was divided into 10 technical replicates, where each technical replicate was analyzed with an equally sized part from a single cocktail of all 17 scFvs. Each replicate was handled in parallel in separate tubes throughout the assay and not mixed until the final pooling for analysis with a single NGS kit. The variability for each scFv is presented as Box plots in Supplementary Fig. 2a, where median value, quartiles, and range for the 10 replicates are shown for each sample. In Supplementary Fig. 2b, the same data are presented as CV values, calculated as the standard deviation divided by the mean of the 10 replicates.

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

Data availability

NGS data (FASTQ files) that support the findings of this study have been deposited in Figshare at <https://doi.org/10.6084/m9.figshare.12370106>. Source data (demultiplexed reads extracted from the NGS data with the Java script) are available in Supplementary Data 1 and the processed data (median normalized and log2 transformed) are available in Supplementary Data 2.

Code availability

The in-house pipeline written in Java for NGS data analysis is available at GitHub: <https://github.com/sunnyveerla/ProMIS/blob/master/ProMIS.java>.

Received: 15 October 2019; Accepted: 11 June 2020;

Published online: 03 July 2020

References

- Borrebaeck, C. A. Precision diagnostics: moving towards protein biomarker signatures of clinical utility in cancer. *Nat. Rev. Cancer* **17**, 199–204 (2017).
- Haab, B. B., Dunham, M. J. & Brown, P. O. Protein microarrays for highly parallel detection and quantitation of specific proteins and antibodies in complex solutions. *Genome Biol.* **2**, RESEARCH0004 (2001).
- Schroder, C. et al. Plasma protein analysis of patients with different B-cell lymphomas using high-content antibody microarrays. *Proteomics Clin. Appl.* **7**, 802–812 (2013).
- Sjoberg, R. et al. Exploration of high-density protein microarrays for antibody validation and autoimmunity profiling. *N. Biotechnol.* **33**, 582–592 (2016).
- Wingren, C. & Borrebaeck, C. A. Antibody-based microarrays. *Methods Mol. Biol.* **509**, 57–84 (2009).
- Carlsson, A. et al. Serum proteome profiling of metastatic breast cancer using recombinant antibody microarrays. *Eur. J. Cancer* **44**, 472–480 (2008).
- Mellby, L. D. et al. Serum biomarker signature-based liquid biopsy for diagnosis of early-stage pancreatic cancer. *J. Clin. Oncol.* **36**, 2887–2894 (2018).
- Wingren, C. et al. Identification of serum biomarker signatures associated with pancreatic cancer. *Cancer Res.* **72**, 2481–2490 (2012).
- Lundberg, M., Eriksson, A., Tran, B., Assarsson, E. & Fredriksson, S. Homogeneous antibody-based proximity extension assays provide sensitive and specific detection of low-abundant proteins in human blood. *Nucleic Acids Res.* **39**, e102 (2011).
- Ullal, A. V. et al. Cancer cell profiling by barcoding allows multiplexed protein analysis in fine-needle aspirates. *Sci. Transl. Med.* **6**, 219ra219 (2014).
- Peterson, V. M. et al. Multiplexed quantification of proteins and transcripts in single cells. *Nat. Biotechnol.* **35**, 936 (2017).
- Shahi, P., Kim, S. C., Haliburton, J. R., Gartner, Z. J. & Abate, A. R. Aseq: ultrahigh-throughput single cell protein profiling with droplet microfluidic barcoding. *Sci Rep-Uk* **7**, 44447 (2017).
- Stoeckius, M. et al. Simultaneous epitope and transcriptome measurement in single cells. *Nat. Methods* **14**, 865 (2017).
- van Buggenum, J. A. G. et al. Immuno-detection by sequencing enables large-scale high-dimensional phenotyping in cells. *Nat. Commun.* **9**, 2384 (2018).
- Bilgin, B., Liu, L., Chan, C. & Walton, S. P. Quantitative, solution-phase profiling of multiple transcription factors in parallel. *Anal. Bioanal. Chem.* **405**, 2461–2468 (2013).
- Wingren, C., Ingvarsson, J., Dexlin, L., Szul, D. & Borrebaeck, C. A. Design of recombinant antibody microarrays for complex proteome analysis: choice of sample labeling-tag and solid support. *Proteomics* **7**, 3055–3065 (2007).
- Wingren, C. et al. Microarrays based on affinity-tagged single-chain Fv antibodies: sensitive detection of analyte in complex proteomes. *Proteomics* **5**, 1281–1291 (2005).
- Ingvarsson, J. et al. Design of recombinant antibody microarrays for serum protein profiling: targeting of complement proteins. *J. Proteome Res.* **6**, 3527–3536 (2007).
- Gerdtsson, A. S. et al. Evaluation of solid supports for slide- and well-based recombinant antibody microarrays. *Microarrays (Basel)* **5**, 16 (2016).
- Wingren, C. & Borrebaeck, C. A. Antibody microarray analysis of directly labelled complex proteomes. *Curr. Opin. Biotechnol.* **19**, 55–61 (2008).
- Wilchek, M. & Bayer, E. A. The avidin-biotin complex in bioanalytical applications. *Anal. Biochem.* **171**, 1–32 (1988).
- Sall, A. et al. Generation and analyses of human synthetic antibody libraries and their application for protein microarrays. *Protein Eng. Des. Sel.* **29**, 427–437 (2016).
- Soderlind, E. et al. Recombining germline-derived CDR sequences for creating diverse single-framework antibody libraries. *Nat. Biotechnol.* **18**, 852–856 (2000).
- Pauly, F. et al. Protein expression profiling of formalin-fixed paraffin-embedded tissue using recombinant antibody microarrays. *J. Proteome Res.* **12**, 5943–5953 (2013).
- Kristensson, M. et al. Design of recombinant antibody microarrays for urinary proteomics. *Proteom. Clin. Appl.* **6**, 291–296 (2012).
- Carlsson, A. et al. Serum protein profiling of systemic lupus erythematosus and systemic sclerosis using recombinant antibody microarrays. *Mol. Cell Proteomics* **10**, M110 005033 (2011).
- Dexlin-Mellby, L. et al. Tissue proteome profiling of preeclamptic placenta using recombinant antibody microarrays. *Proteom. Clin. Appl.* **4**, 794–807 (2010).
- Ingvarsson, J., Lindstedt, M., Borrebaeck, C. A. & Wingren, C. One-step fractionation of complex proteomes enables detection of low abundant analytes using antibody-based microarrays. *J. Proteome Res.* **5**, 170–176 (2006).
- Wingren, C., Ingvarsson, J., Lindstedt, M. & Borrebaeck, C. A. Recombinant antibody microarrays—a viable option? *Nat. Biotechnol.* **21**, 223 (2003).
- Borrebaeck, C. A. et al. Protein chips based on recombinant antibody fragments: a highly sensitive approach as detected by mass spectrometry. *Biotechniques* **30**, 1126–1130, 1132 (2001).
- Westerlund, K., Honarvar, H., Tolmachev, V. & Eriksson Karlstrom, A. Design, preparation, and characterization of PNA-based hybridization probes for antibody-molecule-mediated pretargeting. *Bioconjug. Chem.* **26**, 1724–1736 (2015).
- Ewing, B. & Green, P. Base-calling of automated sequencer traces using phred. II. Error probabilities. *Genome Res.* **8**, 186–194 (1998).
- Ewing, B., Hillier, L., Wendl, M. C. & Green, P. Base-calling of automated sequencer traces using phred. I. Accuracy assessment. *Genome Res.* **8**, 175–185 (1998).

Acknowledgements

This study was supported by grants from the Swedish Research Council, Lars Hiertas Memorial Foundation, Royal Physiographic Society, CREATE Health Translational Cancer Center (www.createhealth.lth.se), and Immunovia AB. We thank Therese Törngren (Division of Oncology and Pathology, Lund University) for assistance and sample analysis of the initial NGS experiments, Center for Translational Genomics (CTG), Lund University and Clinical Genomics Lund, SciLifeLab, for providing sequencing service, and May Hassan (Department of Immunotechnology, Lund University) for laboratory assistance. Open access funding provided by Lund University.

Author contributions

M.B.: experimental design and laboratory work, and data analysis; A.I.E.: experimental design and data analysis; R.J.: designed the cloning strategy, supervised the cloning work, analyzed data from cloning, sequencing, and sortase conjugation; S.G.: performed cloning and sequencing; M.H.: experimental design of cloning and purification; B.E.: experimental design, data interpretation, and access to clinical samples; A.K.: designed the NGS experiments; C.W.: designed part of the experimental design; U.A.: technical support and designed part of the experimental design; C.B.: defined the concept and provided financial means. M.B., C.B., and U.A. wrote the manuscript with input from all authors.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information is available for this paper at <https://doi.org/10.1038/s42003-020-1068-0>.

Correspondence and requests for materials should be addressed to C.A.K.B.

Reprints and permission information is available at <http://www.nature.com/reprints>

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

 **Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2020