nature protocols

https://doi.org/10.1038/s41596-024-01030-x

Protocol

Check for updates

Proximity sequencing for the detection of mRNA, extracellular proteins and extracellular protein complexes in single cells

Luke Vistain[1,2,3], Bijentimala Keisham[1,2], Junjie Xia[1,2], Hoang Van Phan[1,2,4] & Savaş Tay[1,2]

Abstract

Complex cellular functions occur via the coordinated formation and dissociation of protein complexes. Functions such as the response to a signaling ligand can incorporate dozens of proteins and hundreds of complexes. Until recently, it has been difficult to measure multiple protein complexes at the single-cell level. Here, we present a step-by-step procedure for proximity sequencing, which enables the simultaneous measurement of proteins, mRNA and hundreds of protein complexes located on the outer membrane of cells. We guide the user through probe creation, sample preparation, staining, sequencing and computational quantification of protein complexes. This protocol empowers researchers to study, for example, the interplay between transcriptional states and cellular functions by coupling measurements of transcription to measurements of linked effector molecules, yet could be generalizable to other paired events. The protocol requires roughly 16 h spread over several days to complete by users with expertise in basic molecular biology and single-cell sequencing.

Key points

- Prox-seq is suitable for the analysis of primary single cells. The approach adopts various existing single-cell sequencing methods to enable characterization of proteins in close proximity (50–70 nm) along with mRNAs.
- The throughput of hundreds of targets simultaneously enables identification of protein targets forming complexes and is complementary to approaches such as FRET that sense macromolecules in contact over shorter distances.

Key references

Vistain, L. et al. *Nat. Methods* **19**, 1578–1589 (2022): https://doi.org/10.1038/s41592-022-01684-z

Xia, J. et al. *PLoS Comput. Biol.* **20**, e1011915 (2024): https://doi.org/10.1371/journal.pcbi.1011915

[1]Pritzker School of Molecular Engineering, The University of Chicago, Chicago, IL, USA. [2]Institute for Genomics and Systems Biology, The University of Chicago, Chicago, IL, USA. [3]Present address: Lymphocyte Biology Section, Laboratory of Immune System Biology, National Institute of Allergy and Infectious Diseases, National Institutes of Health, Bethesda, MD, USA. [4]Present address: Division of Infectious Disease, University of California, San Francisco, San Francisco, CA, USA. e-mail: tays@uchicago.edu

Introduction

The association and dissociation of multiple proteins in complexes is the process that mediates the execution of most cellular functions. When characterizing the conditions that favor a particular cellular function, the network of protein associations serves as the primary information-processing hub that enables the cells to adopt an appropriate biochemical state. Although protein complexes are clearly central to understanding cellular function, their measurement has lagged progress compared to that of nucleic acids and individual proteins, particularly in single-cell technologies. Recently, we developed an approach that adds this capability to the molecular biology toolbox. Proximity sequencing (Prox-seq) combines a proximity ligation assay (PLA) with single-cell mRNA sequencing, thus enabling the measurement of hundreds of extracellular protein complexes while simultaneously measuring mRNA and protein from single cells[1].

When measuring protein complexes, there is a notable difference in capacity between the processing of single-cell and bulk samples. For bulk samples, one can measure thousands of protein complexes by using various methods such as affinity purification mass spectrometry or yeast two-hybrid assays[2–4]. However, within the context of single-cell sampling approaches, the capacity to measure protein complexes typically falls below 10 species per cells (Table 1). Prox-seq offers a throughput of hundreds of analytes and is suitable for the analysis of primary single cells. The method enables the characterization of each individual cell for three categories: proteins, mRNA and protein complexes.

The purpose of this protocol is to provide the reader with all the information necessary to run a Prox-seq experiment successfully. The basic workflow is very similar to a commercial method called 'Total-seq'[5]. Both methods involve staining cells with DNA-barcoded antibodies that are subsequently measured by using single-cell sequencing. We also guide the reader through the selection of the antibodies, the creation of Prox-seq probes, sample preparation, library preparation for DNA sequencing across multiple platforms and data analysis.

Development of the procedure

Proximity sequencing is a variation of the PLA. This assay creates strands of DNA that can be measured only if two strands are close enough to ligate. It is a versatile assay that can detect nucleic acids and antibodies by using various different measurement technologies[6–9]. The PLA is well suited to our goal of making highly multiplexed protein measurements. The requirement of a ligation step allowed us to incorporate proximity information into our measurements. The resulting DNA molecule can be measured by sequencing, which offers high multiplexing potential. Prox-seq was designed to incorporate into existing single-cell sequencing methods. To accomplish this, we designed Prox-seq oligos and library-preparation steps to be compatible with Drop-seq[10], SMART-Seq2[11] and Chromium 10× 3′ sequencing kits. The data that Prox-seq generates is uniquely complex, because each measurement affects the overall dataset. We have separately performed a thorough evaluation of data analysis methods for Prox-seq[12].

Table 1 | Technical features of approaches to measure more than one protein complex in single cells

Method	Interaction range (nm)	Reported multiplexing	Measured output	Throughput (number of cells)	Reference
FRET	5–10	<10 pairs	Fluorescence	1,000s (flow)–100,000s (imaging)	20
Super-resolution microscopy	5–100	12 targets	Fluorescence	10s	24
PLA-rolling circle amplification (RCA) (DuoLink)	40	4 pairs	Fluorescence	1,000s (flow)–100,000s (imaging)	25
Proximity sequencing	50–70	38 targets, 741 pairs	Sequencing	10,000s	1

Overview of the procedure

The overall workflow for Prox-seq has four phases (Fig. 1a). First, antibodies must be selected and used to produce a panel of Prox-seq probes. Second, cells are stained with Prox-seq probes and ligated. If two probes are close enough to be ligated, then they produce a PLA product. The next step, single-cell isolation, can proceed down three different paths depending on the preference of the user. Prox-seq is compatible with Drop-seq, 10× 3′ Chromium and a plate-based methods similar to SMART-Seq2. Next, sequencing libraries are produced by using slightly different procedures depending on the single-cell isolation method. Samples are then sequenced and finally passed through the data-analysis pipeline.

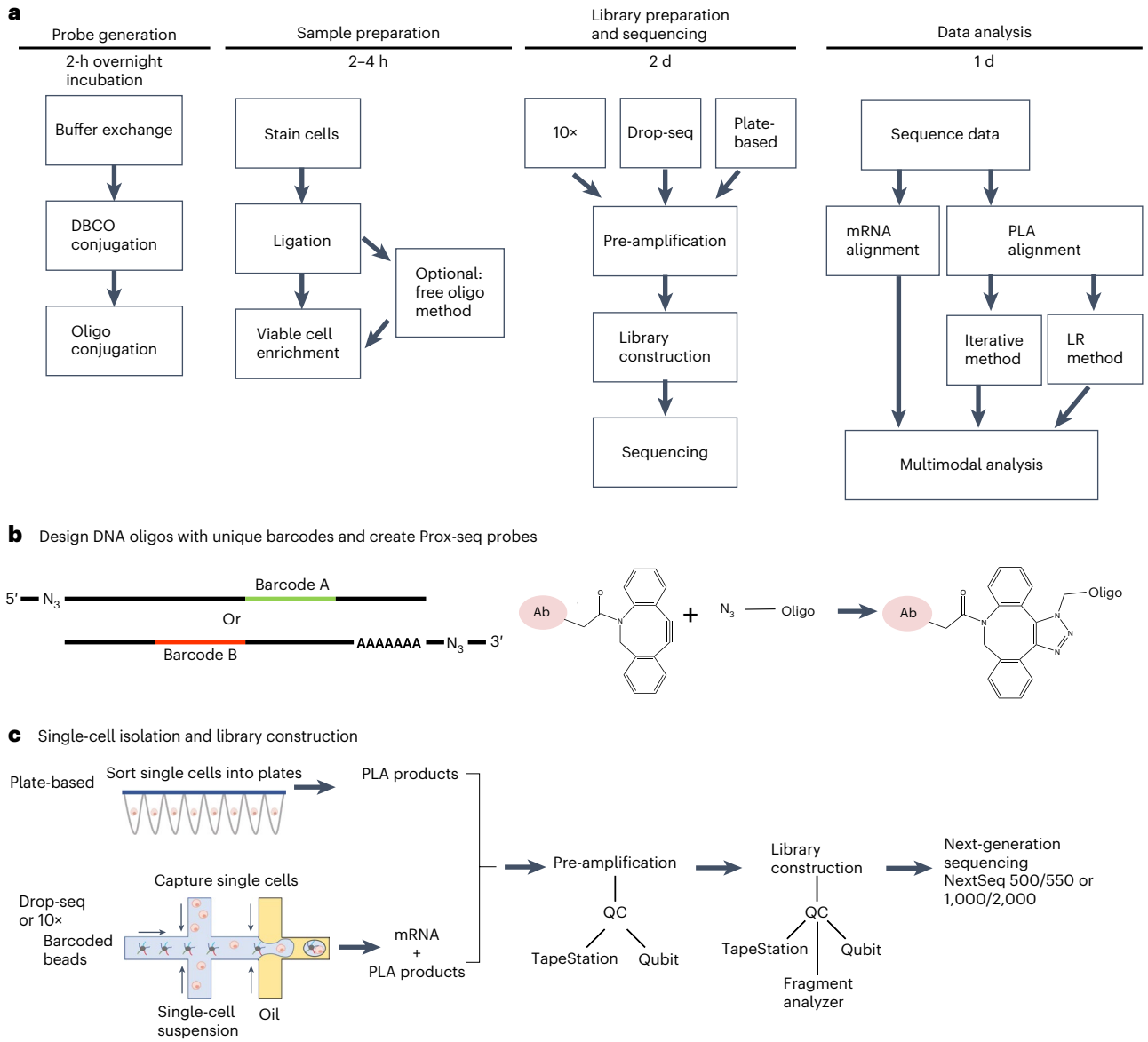


Fig. 1 | Overview of Prox-seq workflow. a, Overall workflow of Prox-seq from probe generation to data analysis. **b**, Creating a PLA probe starts with selecting azide-functionalized DNA oligonucleotides encoded with unique barcodes. Next, antibodies are functionalized with dibenzocyclooctyne (DBCO). The antibody–DBCO conjugates are then combined with the azide-functionalized oligos by using copper-free click chemistry. **c**, Single-cell isolation by the plate-based, Drop-seq or 10× protocol. For the plate-based protocol, single cells are sorted into a 96-well plate containing lysis buffer by using a FACS sorter. For the droplet-based method, Drop-seq or 10×, single cells are captured in droplets along with the barcoded beads and the lysis buffer. The samples are then processed according to their respective protocols consisting of pre-amplification and library construction. Each step includes several QC checks. Finally, the mRNA and PLA libraries are sequenced by using next-generation sequencing.

The first section of this protocol covers the creation of Prox-seq probes. These probes are created by using a copper-free 'click' reaction[13]. First, antibodies are conjugated to a dibenzocyclooctyne (DBCO) moiety by using *n*-hydroxysuccinimide chemistry (Fig. 1b). Second, DBCO-conjugated antibodies are allowed to react with azide-bearing DNA oligomers to produce finished Prox-seq probes. This procedure is compatible with many commercially available antibodies. However, there are some limitations. The selected antibodies cannot have any protein in their buffer, including BSA and gelatin. This procedure is compatible with formulations that include azide, Tris and glycerol. We typically use a batch size of 100 µg of antibody. Batches as low as 25 µg have been attempted, but the final yield is inconsistent.

The second section of this protocol covers the treatment of the cell sample to produce the PLA products that will be used to identify protein complexes and their expression levels. This procedure is very similar to typical cell-staining methods used for flow cytometry, with the addition of a ligation step. Briefly, cells are stained with the entire Prox-seq probe panel, washed and then ligated. Prox-seq normally measures only proteins that are close enough to allow their Prox-seq probes to ligate. For panels that target several highly expressed proteins, this results in the vast majority of probes being measured through sequencing. However, if the researcher is interested in also measuring Prox-seq probes that fail to ligate to another probe, then a variation called the 'free oligo method' can be incorporated at this step. Briefly, additional free oligos are added to the ligation step to ensure that all probes are ligated and measurable, with a specific barcode on the free oligo that indicates that the paired probe did not ligate to another probe.

The third section guides the reader through single-cell isolation. The preferred method will depend on the user. We expect Prox-seq to be compatible with any single-cell sequencing method that relies on poly-A capture to obtain mRNA (Fig. 1c). We advise researchers to use the method with which they are already comfortable, if possible. If they have no preferred method, then we suggest 10× Chromium. It offers the best trade-off between data quality, number of cells analyzed and ease of use. The researcher should consider the cell count requirements for the various single-cell isolation methods. At the end of the sample-processing step, Drop-seq generally requires 100,000 viable cells, 10× requires 20,000 viable cells and plate-based methods require a few thousand cells. The cell loss that occurs during the procedure will depend on the sample type that the researcher uses. We strongly advise performing Steps 27–39 (omit Prox-seq probes in Step 32 to reduce waste) on a test sample to determine how much cell loss is normal and selecting an appropriate starting cell count.

Completion of the single-cell isolation protocol will yield cDNAs and PLA products. These must undergo separate library preparation for sequencing. This procedure changes slightly depending on the single-cell isolation method, but each covers the same basic steps. These include PCR amplification of cDNA and PLA products, purification, attachment of adapter sequences via another PCR step and an additional purification step. Quality control (QC) steps are typically taken after cDNA generation and after final library preparation (Fig. 1c).

Finally, samples are sequenced and passed through the data-analysis pipeline. Raw reads are aligned for each cell. cDNA data are aligned to a refence genome that produces a digital gene expression matrix. PLA product data are aligned to the user-supplied list of barcodes used to make the Prox-seq probe panel, which produces a Digital PLA product Expression (DPE) matrix. The user must ensure that they apply the alignment procedure that matches their single-cell isolation method. The DPE matrix is further processed to estimate the levels of protein expression and protein complexes. To calculate protein complex levels, one must account for the distinctive form of noise that affects Prox-seq, wherein two protein molecules are close enough to be ligated by chance, although they do not have any functional interaction. We call this effect 'proximity noise' and have developed methods to account for it. The free oligo method can improve accuracy of estimates for both protein expression levels and protein complex levels, but we have found that for most applications of Prox-seq (when many highly expressed proteins are targeted), the benefits of the free oligo method are minor[14]. The ultimate outcome of the protocol is a set of count matrices that list mRNA levels, protein levels and protein complex levels for each cell. They can be integrated and analyzed by using typical single-cell analysis software.

Applications

We have demonstrated the use of Prox-seq for measuring changes in protein complex abundances during signaling events, identifying cell types by their extracellular protein complexes and discovering novel protein complexes[1]. Prox-seq can also be performed with several variations, including the combination of various single-cell sequencing platforms, by using fixed cells as samples and omitting the mRNA collection to reduce cost. Samples can be hashed by staining multiple samples with the same antibody but different Prox-seq oligos. This allows samples to be combined during single-cell isolation and then deconvoluted during data analysis[1]. We therefore expect Prox-seq to be adapted for a plethora of applications that involve the characterization of single cells.

The ability to measure the formation of protein complexes during a cell signaling event is among the most impactful uses of Prox-seq. Even relatively simple extracellular receptors will often incorporate numerous co-receptors and accessory proteins to facilitate signaling. For example, the detection of lipopolysaccharide through the TLR2 requires the coordinated action of at least four extracellular proteins (TLR2, MD2, LBP and CD14) in addition to potentially many more proteins that direct the organization of lipid rafts[15]. For any set of proteins, the members of that set can be combined in many different ways to form unique complexes. Prox-seq measures complexes as pairs of proteins. The number of protein pairs that can be created scales quadratically with the number of component proteins. Because of this, even these simple systems can potentially display tens to hundreds of complexes. Prox-seq is currently the only single-cell technology that can measure the hundreds of protein complexes that may be required to facilitate a signaling event. Previous work has shown that Prox-seq can measure changes in protein complex formation during a signaling event and that Prox-seq data can be used to identity the ligands that induced the signaling event[1]. Prox-seq could be a powerful tool for understanding why individual cells display such a wide variance in their response to signaling molecules.

Profiling extracellular complexes offers a complementary data mode to identify novel cell types, which is typically done by the analysis of single-cell mRNA and protein expression data. Our previous work has shown that different cell types can display different protein complexes even when the expression levels of the proteins in complex are otherwise the same[1]. Specifically, when measuring surface protein complexes in peripheral blood mononuclear cells, we observed that the CD8 T cell compartment could be divided on the basis of the preference for CD9 to exist in a homodimer form. Because Prox-seq enables the simultaneous measurement of mRNA, we were able to identify CD9-homodimer-high-CD8 T cells as effector T cells, whereas the compartment that disfavored the CD9 homodimer was composed of naïve CD8 T cells. This basic methodology allows the potential to integrate mRNA, protein and protein complex multi-omics data to discover more cell subpopulations that may be enriched for specific protein complexes. This procedure could be particularly powerful for known cell types that do not have a known surface marker, such as regulatory T cells[16].

Although the requirement to use antibodies prevents Prox-seq from measuring the entire cellular proteome, its throughput is nonetheless high enough to identify novel protein complexes and investigate patterns of protein interactions that could provide insights into basic biological mechanisms. Interested users must first of all choose which sets of proteins to target via their Prox-seq probe panel; the approach can then measure any protein complexes that have the potential to form between pairs of targeted proteins. We applied a medium throughput Prox-seq panel, which could measure a maximum of 741 protein complexes, to peripheral blood mononuclear cells and discovered a novel association between CD9 and CD8. This association was specific to naïve CD8 T cells and would have been difficult to identify from a bulk sample. As a further advantage, unlike commonly used proteome-wide screens (affinity purification mass spectrometry or yeast two-hybrid assays), Prox-seq can be applied directly to primary cells for in situ protein complex detection and quantification.

Several variations of Prox-seq can widen the breadth of possible applications. Prox-seq is compatible with fixed samples (e.g., in plate-based formats[1]), making it particularly useful for cell-signaling studies when the researcher wants to capture cells at a precise time point or if they are concerned that the antibodies used in the assay may induce signaling artefacts. It should be

noted that fixing samples will reduce the quality of the mRNA data[17,18]. We have obtained good results by using a 10× workflow on fixed samples, and Prox-seq might be compatible with fixed-sample mRNA recovery using a 10× workflow. Prox-seq is, in theory, also compatible with frozen samples to the same extent that 10× and Drop-seq are. However, frozen tissue is typically not recommended for single-cell sequencing, and we have therefore not tested its use in this setting. A second possible variation of Prox-seq omits the mRNA collection step. This can be very useful for running pilot experiments in a cost-efficient manner by using the plate-based workflow. We anticipate that Prox-seq is compatible with hashing[19]. Hashing would involve labeling different samples with Prox-seq probes that have identical antibodies but that differ in their Prox-seq barcode. Samples could then be pooled, batch-sequenced and demultiplexed by using the barcode identity. This is a valuable tool for minimizing costs and reducing batch effects by ensuring that various samples undergo the same library preparation and sequencing steps.

Comparison with other methods

Currently, there are few options for studying protein complex formation in single cells. Probably the most used method is Förster Resonance Energy Transfer (FRET). In this method, the adjacency of two fluorophores can be measured by energy transfer between them. However, the multiplexing capacity of FRET is limited to usually less than five complexes and requires substantial information about the distance and orientation of targets[20]. It also has a very short interaction range, on the order of 5–10 nm, whereas Prox-seq has a range of 50–70 nm[1,20]. FRET and Prox-seq therefore cater to different targets and are unlikely to serve as good validation methods for each other. Another commonly used method is a PLA, which is the platform on which Prox-seq is built[9]. PLA kits are sold commercially under the name 'DuoLink'. PLA has the advantage of being much easier to implement than Prox-seq, because the kit enables relatively simple fluorescent readouts by using flow cytometry or microscopy. However, the multiplexing capacity is very small (less than five). Furthermore, unlike Prox-seq, there is no clear way to account for proximity noise. This is the noise source that arises from the fact that targeted proteins may be within range to create signal simply due to random chance rather than due to bona fide molecular interaction. Proximity noise has been shown to interfere with the interpretation of DuoLink assays[21]. In conclusion, options are available for researchers interested in a single important interaction, but Prox-seq is currently the only option for researchers interested in many protein complexes in single cells.

Experimental design

The most important consideration for a Prox-seq experiment is the choice of antibodies that will be used for the Prox-seq probes. If the user is interested in targeting a specific protein complex, then selecting antibodies with epitopes known to be permissive for complex formation is advantageous. Conversely, the use of blocking or neutralizing antibodies should be strictly avoided for studying receptor-ligand complexes. Consideration must also be given to the relative effect of choosing monoclonal or polyclonal antibodies. The advantages of monoclonal antibodies include typically having lower dissociation constant values, sometimes having known binding sites and enabling the quantification of homodimers. The biggest disadvantage of monoclonal antibodies is that, because they have a single binding site, they are more vulnerable to having that site occluded by the formation of a complex. The major advantage of polyclonal antibodies is the improved ability to retain target binding in various conditions because of the presence of multiple binding sites. In practice, we generally prefer the use of monoclonal antibodies that have been validated for use in flow cytometry or immunohistochemistry. We have found that if a chosen antibody displays binding to its target via flow cytometry, then it will produce Prox-seq reads that correlate with the expression level of its target[1].

The experimental procedure has several sections that should be considered before beginning an experiment (Fig. 1a). The most important consideration is the choice of single-cell isolation method. One major advantage of plate-based methods is that they can be used to run comparatively inexpensive pilot experiments by limiting the number of analyzed cells and foregoing mRNA recovery. During developmental stages in our laboratory, this was the most

common method for testing new Prox-seq probe panels. Plate-based methods also yield the highest number of PLA products per cell. The second major consideration is whether to include the free oligo component in the workflow. This step will improve the downstream analysis to better quantify protein and protein complex abundance; however, it is not strictly necessary to produce high quality Prox-seq data[14]. The major trade-off of including the free oligo step is the risk of overprocessing cells. If samples have low cell counts or are highly sensitive to processing steps, then the free oligo step can be skipped.

Required expertise

Most of the Prox-seq protocol is simple to implement for a laboratory equipped to perform common cell biology experiments. All the components needed to make probes are commercially available. Only basic expertise in molecular biology is required to concentrate an antibody, mix it with conjugation reagent, purify it by using a concentrator and finally measure the degree of conjugation by using UV–visible absorbance. Cell handling requires expertise in tissue culturing and its associated equipment.

Expertise in FACS is strongly recommended. This is because flow cytometry offers the most effective method for validating probe binding to cells. Furthermore, plate-based methods will require single-cell sorting. It is possible to perform a Prox-seq experiment without these methods, but we expect that most users will prefer to use them.

Expertise in single-cell sequencing is required. Researchers must have access to a method that isolates single cells for Drop-seq, 10× Chromium and plate-based methods. FACS sorting expertise is required for plate-based methods. Furthermore, expertise in operating an Illumina sequencer is essential. This protocol is written for the NextSeq 550, but other sequencing platforms can be used with appropriate modification to library preparation. Users will need expertise in single-cell data analysis to interpret the sequencing data. Finally, the analysis of single-cell data can be greatly accelerated by using supercomputing facilities.

Limitations

The most obvious limitation for Prox-seq is the requirement to use antibodies, which ultimately limits the number of proteins and complexes that can be measured. Antibodies must be selected very carefully because Prox-seq probes must be able to bind their targets both individually and in complex.

The upper limit for Prox-seq multiplexing is estimated to be in the low hundreds of targets. We have previously shown multiplexing potential of up to 38 targets, and we expect that this number can be increased. Our own data[1] along with those of others[22,23] suggest that DNA-conjugated antibodies display minimal scaling noise, and similar technologies have been multiplexed up to around 100 targets[22]. A Prox-seq experiment that targets 100 proteins could measure up to 5,050 protein complexes. Because of these scaling properties, we expect that the multiplexing limit of Prox-seq will be determined by antibody availability and cost. The initial investment cost for a new Prox-seq probe pair is roughly $1,000; however, this cost includes oligomers that can be used for many different Prox-seq experiments.

Another limitation is that it is difficult to validate whether the proteins that form a complex are in direct physical contact because Prox-seq data provide a measure of distance between proteins. The range of Prox-seq is expected to be 50–70 nm. This is a large enough range that makes it difficult to rule out the possibility that other proteins are responsible for bringing the two targeted proteins in proximity, but not bound to each other. Direct contact can then be tested with measures of shorter ranges (such as FRET).

Finally, we anticipate that Prox-seq may provide a more holistic understanding of complex formation when a large number of proteins are probed. Although lowly abundant proteins that are not in complex with another target in the panel can be hard to measure (because they have little chance to form a valid ligation product that can be sequenced afterwards), having a large number of targets may help measure their expression level because the lowly expressed targets will be adjacent to a targeted protein by chance. However, lowly abundant proteins on the cell surface will probably remain difficult to identify.

Materials

Reagents

- 1× PBS (GIBCO, cat. no. 20012027)
- Amicon Ultra-0.5 centrifugal filter unit (50,000-molecular-weight-cutoff concentrator; EMD Millipore, cat. no. UFC5050)
- Dibenzocyclooctyne-PEG4-*N*-hydroxysuccinimidyl ester (Sigma-Aldrich, cat. no. 764019)
- Dimethyl sulfoxide (DMSO; Sigma-Aldrich, cat. no. D8418)
  ▲ **CAUTION** DMSO is flammable. Handle by using appropriate safety equipment.
- Glycerol (Sigma-Aldrich, cat. no. G5516)
- Antibodies selected by the user. A list of antibodies used in this protocol can be found in Supplementary Table 1
  ▲ **CRITICAL** Avoid antibodies formulated with protein, including BSA and gelatin.
- All custom DNA oligomers were purchased from IDT. All Prox-seq oligomers can be found in an article by Vistain et al.[1]. This includes both oligomers used in library preparation and for conjugation to antibodies
- UltraPure DNase/RNase-free distilled water (Invitrogen, cat. no. 10977023)
- Blocker BSA (Thermo Scientific, cat. no. 37520)
- Trypan blue (Gibco, cat. no. 15250061)
- Ultrapure salmon sperm DNA solution (Invitrogen, cat. no. 15632011)
- ATP (10 mM; New England Biolabs, cat. no. P0756L)
- T4 ligase (New England Biolabs, cat. no. M0202L)
- Propidium iodide (PI; Invitrogen, cat. no. P1304MP)
- TruStain FcX (Biolegend, cat. no. 422302)
- Dead cell removal kit (Miltenyi Biotec, cat. no. 130-090-101)
- RNase inhibitor, murine (New England Biolabs, cat. no. M0314L)
- Triton X-100 (98%; Thermo Scientific, cat. no. 327371000)
  ▲ **CAUTION** Triton X-100 is a hazardous material and harmful if swallowed. It can also cause serious eye damage/irritation. Handle with care by using appropriate safety measurements.
- dNTP mix (10 mM each; Thermo Scientific, cat. no. R0193)
- Proteinase K (New England Biolabs, cat. no. P8107S)
- Ethanol, 200 proof (Decon Labs, cat. no. 2705)
  ▲ **CAUTION** Ethanol is flammable. Handle by using appropriate safety equipment.
- Drop-seq beads (ChemGenes, cat. no. Macosko-2011-10(V+))
- QX200 droplet generation oil for EvaGreen (Bio-Rad, cat. no. 1864006)
- Saline-sodium citrate solution (10×; Fisher BioReagents, cat. no. AC611520010)
- 1H,1H,2H,2H-Perfluorooctan-1-ol (Synquest Laboratories, cat. no. 2101-3-20)
  ▲ **CAUTION** 1H,1H,2H,2H-Perfluorooctan-1-ol is flammable. It can also cause skin, eye and respiratory irritation. Avoid inhaling fumes or contact with the skin. Handle with care by using appropriate safety equipment.
- 5× Maxima H minus RT buffer (Thermo Scientific, cat. no. EP0752)
- Ficoll Solution (20%; Sigma-Aldrich, cat. no. F5415-25 ml)
- Exonuclease I and 1× exonuclease I buffer ((New England Biolabs, cat. no. M0293)
- Nextera XT DNA library preparation kit (Illumina, cat. no. FC-131-1024)
- Chromium Next GEM Chip G single cell kit (10× Genomics, cat. no. 1000127)
- Chromium Next GEM single cell 3′ GEM, library and gel bead kit v.3.1 (10× Genomics, cat. no. 1000268)
- Dual index kit TT set A, 96 reactions (10× Genomics, cat. no. 1000215)
- AMPure XP (Beckman Coulter, cat. no. A63881)
- Buffer EB (Qiagen, cat. no. 19086)
- Agilent high-sensitivity D5000 ScreenTape (Agilent, cat. no. 5067-5592)
- Agilent high-sensitivity D5000 reagents (Agilent, cat. no. 5067-5593)
- Agilent HS NGS fragment kit (1–6,000 bp), 500 (Agilent, cat. no. DNF-474-0500)
- Qubit 1× double-stranded DNA high-sensitivity assay kit (Invitrogen, cat. no. Q33231)

- Qubit assay tubes (Invitrogen, cat. no. Q32856)
- KAPA HiFi HotStart ReadyMix (2×; KAPA Biosystems, cat. no. KK2601)
- SPRIselect beads (Beckman Coulter, cat. no. B23317)
- NextSeq PhiX control kit (Illumina, cat. no. FC-110-3002)
- NextSeq 500/550 mid-output kit v2.5 (150 cycles; Illumina, cat. no. 20024904)
- NextSeq 500/550 high-output kit (150 cycles; Illumina, cat. no. 20024907)
- 10% Sarkosyl Solution (1,000 ml; Teknova, cat. no. S3377)
- DL-Dithiothreitol solution (DTT, 1 M; Sigma-Aldrich, cat. no. 43816-10 ml)

Equipment

- Microcentrifuge tube, 1.5 ml
- DNA LoBind tube, 1.5 ml (Eppendorf, cat. no. 022431021)
- Falcon polystyrene conical tube, 15 and 50 ml
- Filter tips (P1000, P200, P100 and P10)
- Rainin tips (P20 and P200)
- 96-well plate (Eppendorf, Bio-Rad)
- Hemocytometer (Hausser Scientific, cat. no. 3520)
- Falcon cell strainers, 40 µm (Fisher Scientific, cat. no. 08-771-1)
- Magnetic disc (V&P Scientific, cat. no. 772DP-N42-5-2)
- Plastic tubes (Scientific Commodities, cat. no. BB31695PE/2)
- Magnetic stirrer (V&P Scientific, cat. no. 710D2)
- DiaMag 0.2-ml magnetic rack (Diagenode, cat. no. B04000001)
- DynaMag-96 side magnet (Invitrogen, cat. no. 12331D)
- Themal cycler (Bio-Rad C1000 Touch, cat. no. 1851196)
- PX1 PCR plate sealer (Bio-Rad, cat. no. 1814000)
- PCR plate heat seal, foil, pierceable (Bio-Rad, cat. no. 1814040)
- Adhesive PCR plate foils (Thermo Scientific, cat. no. AB-0626)
- Nanodrop UV–visible spectrophotometer (Thermo Scientific, cat. no. ND-ONE-W)
- Qubit 4.0 fluorometer (Invitrogen, cat. no. Q33238)
- Agilent 5200 fragment analyzer (Agilent)
- Agilent 4200 TapeStation (Agilent)
- Fortessa 4-15 (BD Biosciences)
- Cell sorter (BD FACSAria II)
- Pump 11 Pico Plus Elite infusion/withdrawal programmable dual syringe pump (Harvard Apparatus, cat. no. 70-4506)
- Chromium Controller (10× Genomics, cat. no. 1000204)
- NextSeq 500/550 (Illumina)

Reagent setup

- TE-TW (10 mM Tris pH 8.0, 1 mM EDTA, 0.01% (wt/vol) Tween-20 and nuclease-free water)
- TE-SDS (10 mM Tris pH 8.0, 1 mM EDTA, 0.5% (wt/vol) SDS and nuclease-free water)
- 10× HEPES buffer
- Probe binding buffer (PBS, 0.1% (wt/vol) BSA (Thermo Scientific), 0.1 mg/ml sonicated salmon sperm DNA (Invitrogen), 6.7 nM for each isotype)
  ▲ **CRITICAL** Most solutions are made fresh with each experiment. However, the probe binding buffer can be stored at 4 °C.
- Ligation buffer (50 mM HEPES pH 7.5, 10 mM $MgCl_2$, 1 mM rATP, 9.5 nM connector oligomer (TTTCACGACACGACACGATTTAGGTC), 130 U/ml T4 ligase and nuclease-free water)
- Oligo solution in PBS (80 µM)
- Drop-seq lysis buffer (6% (wt/vol) Ficoll type 400, 200 mM Tris pH 7.5, 20 mM EDTA, 0.2% (wt/vol) Sarkosyl, 50 mM DTT and nuclease-free water)
  ▲ **CRITICAL** A 50-ml stock of Drop-seq lysis buffer (without DTT) can be prepared and stored at 4 °C. Before use, add DTT at 50 mM: 50 µl of 1 M DTT/1 ml of lysis buffer. Filter the lysis buffer through a 40-µm cell strainer.

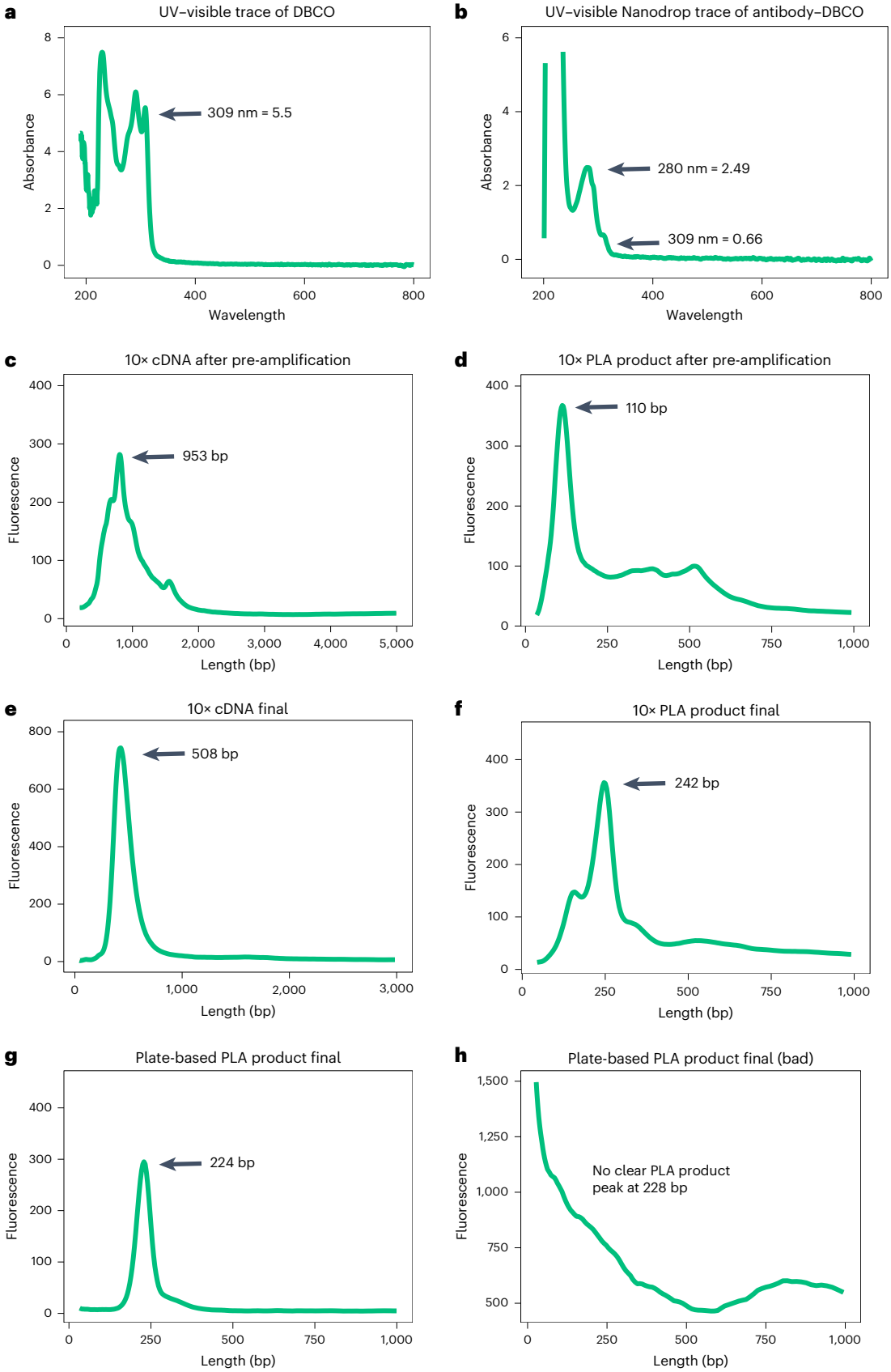

a
UV–visible trace of DBCO
Absorbance
Wavelength
309 nm = 5.5
b
UV–visible Nanodrop trace of antibody–DBCO
Absorbance
Wavelength
280 nm = 2.49
309 nm = 0.66
c
10× cDNA after pre-amplification
Fluorescence
Length (bp)
953 bp
d
10× PLA product after pre-amplification
Fluorescence
Length (bp)
110 bp
e
10× cDNA final
Fluorescence
Length (bp)
508 bp
f
10× PLA product final
Fluorescence
Length (bp)
242 bp
g
Plate-based PLA product final
Fluorescence
Length (bp)
224 bp
h
Plate-based PLA product final (bad)
Fluorescence
Length (bp)
No clear PLA product
peak at 228 bp

Fig. 2 | Overview of the QC results. a, UV–visible Nanodrop trace for DBCO indicating the absorbance at 309 nm. **b**, DBCO-conjugated antibodies showing the absorbances at 280 nm and 309 nm. The peak at 210 nm is truncated for clarity. The complete dataset can be found at https://github.com/tay-lab/NProt_data-src. **c**, The TapeStation traces of cDNA using a 10× pipeline after pre-amplification with a peak around 1,000 bp. **d**, The TapeStation traces of PLA using a 10× pipeline after pre-amplification with a peak around 100 bp. **e**, The TapeStation traces of final cDNA products using a 10× pipeline with a peak around 500 bp. **f**, The TapeStation traces of final PLA products using a 10× pipeline with a peak around 228 bp. **g**, The trace of PLA products generated by using the plate-based protocol measured by a fragment analyzer. **h**, The TapeStation traces of PLA products generated from a suboptimal experiment using the plate-based protocol.

- Plate cell lysis buffer (0.1% (wt/vol) Triton X-100, 1 U/µl RNase inhibitor (murine), 2.5 mM dNTPs, 2.5 µM SmartSeq2_oligodTVN and 2.5 µM SmartSeq2_oligodTGT in nuclease-free water)
  ▲ **CRITICAL** If using fixed cells, include 20 U/ml Proteinase K in the lysis buffer to improve the RNA yield.
- 80% (vol/vol) ethanol
  ▲ **CRITICAL** Always use freshly prepared 80% (vol/vol) ethanol.

Software

- Java, version 8 (www.java.com/en/)
- R 4.0.5 or higher (https://www.r-project.org/) and R Studio 2021.09.0-351 or higher (https://www.rstudio.com/products/rstudio/download/)
- Python 3.11 or higher (https://www.python.org/)
- A custom Java program for PLA sequencing data alignment, available at https://github.com/tay-lab/Prox-seq/blob/master/PLA_alignment.jar
- A custom Python script for PLA data analysis, available at https://github.com/tay-lab/NProt_data-src

Procedure

Preparing Prox-seq probes

● **TIMING ~2 h of work time, overnight incubation**

1. For each antibody, wet a 50,000-molecular-weight-cutoff concentrator membrane with 500 µl of PBS.
2. Spin each concentrator at 14,000*g* at 4 °C for 1 min.
3. Discard the liquid that flowed into the bottom of the concentrator.
4. Dilute antibodies to 500 µl with PBS.
5. Transfer the diluted antibodies to the concentrator.
6. Spin concentrators at 14,000*g* and 4 °C for 5 min.
7. Discard the liquid that flowed into the bottom of the concentrator.
8. Add 450 µl of PBS to the concentrator, rediluting the antibody.
9. Spin concentrators at 14,000*g* and 4 °C for 5 min.
10. Discard the liquid that flowed into the bottom of the concentrator.
11. Add 450 µl of PBS to the concentrator, rediluting the antibody.
12. Spin concentrators at 14,000*g* and 4 °C for 10 min.
13. Dissolve a small amount (the amount that sticks to the tip of a glass pipette when pressed into the waxy pellet) of dibenzocyclooctyne-PEG4-*N*-hydroxysuccinimidyl ester in 10 µl of DMSO. This is the stock DBCO solution.
14. Take a 1-µl sample from the stock DBCO solution and dilute 100-fold in PBS.
15. Measure the absorbance at 309 nm of the diluted DBCO stock with a UV–visible spectrophotometer (Nanodrop) (Fig. 2a).
16. Calculate the concentration of the stock DBCO solution by using the molar extinction coefficient 12,000 M^{-1} cm^{-1}.
17. Prepare a 2 mM DBCO solution in DMSO by using the stock DBCO solution. If the concentration of DBCO is too low, more DBCO can be added to the stock solution by repeating Steps 13–16.

18. Combine 2 mM DBCO in DMSO with the concentrated antibody at a 1:10 volume-by-volume ratio.
    ▲ **CRITICAL STEP** Verify the antibody concentration before DBCO conjugation. The optimal antibody concentration is 0.5–2 mg/ml.
19. Incubate the antibody-DBCO solution on ice for 1 h.
    ▲ **CRITICAL STEP** This reaction time can be increased up to 2 h without negatively affecting the reaction.
20. After incubation, purify the DBCO-conjugated antibodies by repeating Steps 4–12. The same concentrator can be used for both buffer exchange and purification.
21. Using a UV–visible spectrophotometer (Nanodrop), measure the absorbances of the antibody and DBCO at 280 and 309 nm, respectively.
22. Calculate the concentrations of both antibody and DBCO by using their respective molar extinction coefficient (12,000 M^{-1} cm^{-1} for DBCO and 204,000 M^{-1} cm^{-1} for the antibody). The user must account for the absorbance DBCO displays at 280 nm[13] (Fig. 2b).
    ◆ **TROUBLESHOOTING**
23. Divide the molar concentration of DBCO by the molar concentration of antibody to get the antibody-to-DBCO ratio.
    ■ **PAUSE POINT** Divide the DBCO-conjugated antibodies into aliquots and store at −20 °C. Avoid repeated freeze-thaw cycles.
24. Finally, to make the Prox-seq probes, combine 2 µg of DBCO-conjugated antibody with an equal volume of 80 µM azide-functionalized oligonucleotides.
25. Incubate overnight at 4 °C. The reaction time can be extended up to 2 d without any negative effect.
26. To store long term, bring the probe solution to 50% (wt/vol) glycerol in PBS. In our hands, probes stored this way are stable for ≥1 year. Probes stored at 4 °C are typically stable for several months; however, this is dependent on the specific antibody.
    ■ **PAUSE POINT** Store the 50% (wt/vol) glycerol stock solution of Prox-seq probes at −20 °C.

Sample processing (PLA)

● **TIMING ~3 h**

27. Centrifuge the cells. The spin speed and time will depend on the user's sample, but typically 300–500*g* for 5 min at 4 °C is suitable.
28. Discard the supernatant.
29. Resuspend the cells with 1 ml of 1% (wt/vol) BSA/PBS.
30. Spin the cells by using the same centrifuge conditions as in Step 27.
31. Discard the supernatant.
32. Prepare Prox-seq probe solution in probe binding buffer. The final concentration of each probe should be 2.5 nM. Note that because there are two probes for each targeted protein, the total amount of probe for each target is 2.5 nM probe A + 2.5 nM probe B = 5 nM.
33. Resuspend the cell pellet with the probe mix. Use 30 µl of probe mixture for every 100,000 cells, scaling the volume accordingly.
34. Incubate at 37 °C for 30–60 min.
35. Wash three times with 1 ml of 1% (wt/vol) BSA/PBS by centrifugation and resuspension as in Steps 27–29. If overprocessing of cells is a concern, the number of washes can be reduced by at least one.
36. Resuspend the cell pellet in 100 µl of ligation solution.
37. Incubate the sample at 37 °C for 30 min.
38. (Optional) Perform a free oligo experiment. After completion of the 30–60-min incubation in Step 34:
    - Add free oligo A to a final concentration of 50 nM.
    - Add additional connector to a final concentration of 59.5 nM.
    - Incubate at 37 °C for 15 min.
    - Add free oligo B to a final concentration of 60 nM.
    - Add additional connector to a final concentration of 109.5 nM.
    - Incubate at 37 °C for 15 min.

39. Wash twice by repeating Steps 27–29.
40. Proceed to the single-cell isolation step. Start at Step 41 for Drop-seq, Step 87 for 10× Chromium and Step 103 for plate-based methods.
    ◆ **TROUBLESHOOTING**

Drop-seq–based Prox-seq

● **TIMING ~8 h**

41. Remove any dead cells from Step 39 by using Miltenyi Biotec dead cell removal kit and following the standard manufacturer's protocol.
42. Filter the cells with a 40-µm cell strainer to remove any clumps.
43. Process the cells for Drop-seq according to the standard Drop-Seq protocol up to and including the exonuclease digestion step[10].
44. Distribute Drop-seq beads into a 96-well plate at ~5,000 beads per well.
45. Set up a 50-µl PCR reaction for each well, with 1× KAPA HiFi HotStart Readymix, 0.8 µM TSO_PCR primer and 0.8 µM U_fwd primer.
46. Heat-seal the plate. Adhesive seals are more prone to sample evaporation.
47. Incubate in a thermal cycler and perform cDNA amplification according to the following parameters:

Cycle number	Denature	Anneal	Extend	Hold
1	95 °C, 3 min	–	–	–
2–5	98 °C, 20 s	65 °C, 45 s	72 °C, 3 min	–
6–15	98 °C, 20 s	67 °C, 20 s	72 °C, 3 min	–
16	–	–	72 °C, 5 min	–
17	–	–	–	4 °C

48. Add 10 µl of nuclease-free water to each well.
49. Centrifuge at 1,000*g* for 1 min.
50. Transfer 50 µl of the supernatant from each well to a clean 96-well plate (plate 1).
51. Add 30 µl of AMPure XP beads to each well to obtain a 0.6× AMPure bead concentration. Pipette-mix 15 times.
52. Incubate the plate at room temperature for 5 min.
53. Transfer onto a magnetic rack for 5 min.
54. Transfer the 80-µl supernatant into a new 96-well plate (plate 2). This will be used to recover PLA products.
55. Wash each well in plate 1 four times with 200 µl of freshly prepared 80% (vol/vol) ethanol.
56. Air-dry for 2 min.
57. Remove from the magnet and add 11.5 µl of nuclease-free water to each well to elute the cDNA products.
58. Incubate the plate at room temperature for 5 min.
59. Transfer onto a magnetic rack for 3 min.
60. Collect 10 µl of the supernatant from each well.
61. Pool them into a 1.5-ml microcentrifuge tube.
62. Add 60 µl of AMPure beads to each well (1.8×) of plate 2 from Step 54 containing the supernatant with the PLA products. Pipette-mix 15 times.
63. To obtain purified PLA products, repeat Steps 52–60. Skip Step 54; discard the supernatant instead.
64. Quantify the cDNA with an Agilent high-sensitivity DNA TapeStation. The distribution of the cDNA is quite broad, ranging between 400 and >3,000 base pairs, with a peak at ~1,000 bp.
    ■ **PAUSE POINT** The libraries can be stored long term at −20 °C.
    ◆ **TROUBLESHOOTING**
65. On ice, combine 450 pg of cDNA products with 10 µl of Tagment DNA buffer and 5 µl of Amplicon Tagment (enzyme) mix.
    ▲ **CRITICAL STEP** The maximum input volume for the library is 5 µl. Thus, take 450 pg of cDNA and bring up to 5 µl with nuclease-free water.

66. Incubate at 55 °C for 5 min. Place on ice immediately after the incubation.
67. Add 5 µl of neutralization buffer to the mixture.
68. Incubate again at room temperature for 5 min.
69. On ice, add the following components to the mixture for a total volume of 50 µl: 8 µl of nuclease-free water, 1 µl of 10 µM P5_TSO primer, 1 µl of 10 µM P7_N70X_Custom2 primer and 15 µl of Nextera PCR Master Mix.
70. Incubate in a thermal cycler according to the following parameters.

Cycle number	Denature	Anneal	Extend	Hold
1	95 °C, 30 s	–	–	–
2–13	95 °C, 10 s	55 °C, 30 s	72 °C, 30 s	–
14	–	–	72 °C, 5 min	–
15	–	–	–	4 °C

71. Add 30 µl of AMPure beads (0.6×).
72. Incubate the plate at room temperature for 5 min.
73. Transfer onto a magnetic rack for 5 min.
74. Discard the supernatant.
75. Wash each well four times with 200 µl of freshly prepared 80% (vol/vol) ethanol.
76. Air-dry for 2 min.
77. Remove from the magnet.
78. Add 11.5 µl of nuclease-free water to each well.
79. Incubate the plate at room temperature for 5 min.
80. Transfer onto a magnetic rack for 3 min.
81. Collect 10 µl of the supernatant from each well.
82. Pool them into a 1.5-ml microcentrifuge tube.
83. To prepare the PLA product library, use 1 µl of the pooled PLA products to set up a 20-µl PCR reaction (1× KAPA HiFi HotStart Readymix, 0.2 µM P5_TSO primer, 0.2 µM P7_N7XX_Custom2 primer and nuclease-free water).
84. Incubate in a thermal cycler according to the following parameters.

Cycle number	Denature	Anneal	Extend	Hold
1	95 °C, 3 min	–	–	–
2–13	98 °C, 20 s	67 °C, 15 s	72 °C, 20 s	–
14	–	–	72 °C, 5 min	–
15	–	–	–	4 °C

85. Clean up the PCR products with 20 µl of AMPure beads (1.0×) as described in Steps 72–81.
86. Quantify both the final products by using a Qubit fluorometer and an Agilent high-sensitivity DNA TapeStation. Measure the concentration by using Qubit and the library size distribution with an Agilent high-sensitivity chip according to the manufacturers' instructions.
    ■ **PAUSE POINT** The final libraries (cDNA and PLA product) can be stored long term at −20 °C until ready for sequencing.
    ◆ **TROUBLESHOOTING**

10×-based Prox-seq

● **TIMING ~8 h**

87. Remove any dead cells from Step 39 by using a Miltenyi Biotec dead cell removal kit and following the standard manufacturer's protocol.
88. Filter the cells with a 40-µm cell strainer to remove any clumps.
89. Perform 10× 3′ single-cell isolation by following the manufacturer's protocol until just before the start of cDNA amplification.
90. To amplify, add 1 µl of 2 µM U_fwd primer to the amplification mix, along with the cDNA primers included in the 10× kit. Set up the reaction by following the instructions by 10× Genomics with 12 PCR cycles.

91. Separate the PLA products from cDNA by using the SPRI-based size selection step by incubating the amplified cDNA with 60 µl of SPRIselect beads to a final volume ratio of 0.6× for 5 min at room temperature.
92. Save 75 µl of the supernatant for PLA library construction and discard the rest (10× Genomics, CG000317 RevB).
▲ **CRITICAL STEP** saving the supernatant is essential. In a typical 10× protocol, the supernatant is discarded; however, in Prox-seq the supernatant contains the PLA products.
93. The cDNA fragments remain on the SPRIselect beads. Wash the SPRIselect beads with 80% (vol/vol) ethanol twice and elute the cDNA into 35 µl of buffer EB (Qiagen).
94. Quantify the cDNA fragments by using Qubit and an Agilent high-sensitivity DNA TapeStation. Measure the concentration by using Qubit and the library size distribution with an Agilent high-sensitivity chip, according to the manufacturers' instructions. cDNA libraries typically have an average length of 1,000 bp (Fig. 2c).
◆ **TROUBLESHOOTING**
95. Using 10 µl of the purified cDNA product, construct the gene expression library according to the standard 10× Genomics single-cell 3′ V3.1 protocol. The number of PCR cycles is determined by the amount of cDNA input.
96. To construct the PLA product library, clean up the supernatant saved in Step 92 by incubating with 70 µl of SPRIselect beads to a final volume ratio of 2.1×.
97. Wash the SPRIselect beads with 80% (vol/vol) ethanol twice and elute into 40 µl of buffer EB (Qiagen).
98. Verify the presence and concentration of PLA products by using an Agilent high-sensitivity DNA TapeStation (Fig. 2d).
99. Using 10 µl of PLA product, prepare a 50-µl PCR reaction volume with 1× KAPA HiFi HotStart Readymix, 0.2 µM 10×_SI_PCR primer, 0.2 µM P7_N70X_Custom2 primer and nuclease-free water.
100. Incubate in a thermal cycler by using the following program.

Cycle number	Denature	Anneal	Extend	Hold
1	98 °C, 2 min	–	–	–
2–13	98 °C, 20 s	60 °C, 30 s	72 °C, 20 s	–
14	–	–	72 °C, 5 min	–
15	–	–	–	4 °C

101. Clean up the PCR products by using 60 µl of SPRIselect beads at a final volume ratio of 1.2× and elute into 40 µl of buffer EB (Qiagen).
102. Quantify the purified samples (cDNA and PLA product) with a Qubit fluorometer and an Agilent TapeStation (Agilent) to determine the library size (Fig. 2e,f, respectively).
■ **PAUSE POINT** The final libraries (cDNA and PLA product) can be stored long term at −20 °C until ready for sequencing.
◆ **TROUBLESHOOTING**

Plate-seq–based Prox-seq

● **TIMING ~6 h or more, depending on the number of plates**

103. Prepare 96-well plates by adding 6 µl of lysis buffer per well (0.1% (wt/vol) Triton X-100, 1 U/µl RNase inhibitor (murine), 2.5 mM dNTPs, 2.5 µM SmartSeq2_oligodTVN and 2.5 µM SmartSeq2_oligodTGT in nuclease free water). For PFA-fixed cells, prepare the 96-well plates by using 6 µl of modified lysis buffer per well (0.1% (wt/vol) Triton X-100, 1,000 U/ml RNase inhibitor (), 20 U/ml Proteinase K, 2.5 mM dNTPs, 2.5 µM SmartSeq2_oligodTVN and 2.5 µM SmartSeq2_oligodTGT in TE buffer).
104. Resuspend the cells from Step 40 in 1 ml of 1% (wt/vol) BSA/PBS + 1/500 PI.
105. Perform single-cell (PI-negative) sorting via FACS into the plates prepared in Step 103.
106. Heat-seal the plate.

107. Centrifuge the plates at 700*g* for ≥10 s. For PFA-fixed cells, incubate in a thermal cycler set to 56 °C for 1 h, 95 °C for 10 min and 4 °C for 5 min.
■ **PAUSE POINT** Plates can be stored long term at −80 °C.
108. Thaw the plates on ice.
109. Incubate them at 72 °C for 3 min.
110. Centrifuge at 700*g* for ≥10 s.
111. Take 2–4 µl of the cell lysate as input for pre-amplification and put it in a separate tube.
▲ **CRITICAL STEP** The lysate must be divided in two if the researcher plans to measure both mRNA and PLA products. We expect that mRNA can be recovered by using the procedure outlined by Picelli et al.[11]. Ensure that less than half is taken for each analyte to account for pipetting error. If the researcher is measuring only PLA products, then all the lysate can be used.
112. To the cell lysate, add the pre-amplification buffer, which contains 1× KAPA HiFi HotStart Readymix, 0.1 µM Oligo_dTGT primer, 0.1 µM U_fwd primer and nuclease-free water, for a total volume of 25 µl.
113. Incubate in a thermal cycler by using the following program.

Cycle number	Denature	Anneal	Extend	Hold
1	98 °C, 3 min	–	–	–
2–5	98 °C, 20 s	55 °C, 15 s	72 °C, 1 min	–
6–22	98 °C, 20 s	67 °C, 15 s	72 °C, 1 min	–
23	–	–	72 °C, 5 min	–
24	–	–	–	4 °C

114. Clean up the PCR products by using 45 µl of AMPure beads (yielding a 1.8× volume ratio). Pipette-mix 15 times.
115. Incubate the plate at room temperature for 5 min.
116. Transfer onto a DynaMag-96 side magnet rack for 5 min.
117. Discard the supernatant.
118. Wash twice with 200 µl of freshly preparedof 80% (vol/vol) ethanol.
119. Air-dry for 2 min.
120. Remove from the magnet.
121. Add 17.5 µl of nuclease-free water to each well.
122. Pipette-mix 10 times.
123. Incubate the plate at room temperature for 5 min.
124. Transfer onto a magnetic rack for 2 min.
125. Collect 15 µl of the supernatant from each well into a new 96-well plate.
126. Take 2 µl of the purified PLA products from each well as input to a 25-µl PCR reaction, which contains 1× KAPA HiFi HotStart Readymix, 1 µM SmartPLA_P5_S5XX primer, 1 µM SmartPLA_P7_N7XX primer and nuclease-free water.
127. Incubate the reaction mixture in a thermal cycler by using the following program.

Cycle number	Denature	Anneal	Extend	Hold
1	98 °C, 3 min	–	–	–
2–13	98 °C, 20 s	67 °C, 15 s	72 °C, 1 min	–
14	–	–	72 °C, 5 min	–
15	–	–	–	4 °C

128. Clean up the barcoded PLA with 45 µl of AMPure beads (yielding a 1.8× volume ratio). Follow the purification process described in Steps 115–125.
129. Quantify the purified products with a Qubit fluorometer.
130. Pool each library in equimolar amounts such that each cell contributes ≥4 ng of DNA.
131. Measure the pooled sample by using a high-sensitivity DNA TapeStation and a high-sensitivity fragment analyzer before proceeding with sequencing (Fig. 2g).

▲ **CRITICAL STEP** Note that the library preparation procedure can be performed either manually or automatically with an automatic liquid handling system (PerkinElmer Janus G3 and Tecan Freedom Evo 200).

■ **PAUSE POINT** The final libraries (cDNA and PLA product) can be stored long term at −20 °C until ready for sequencing.

Next-generation sequencing

● **TIMING ~1–2 d**

132. *Drop-seq–based Prox-seq*. mRNA and PLA libraries can be sequenced separately or combined by using the NextSeq 550 platform, depending on the number of reads that the researcher wants to allocate to each.
    - If mRNA and PLA are being sequenced together, pool the libraries in the ratio of 4:1, such that mRNA receives four times the sequencing depth.
    - Spike in PhiX at a 40% vol/vol concentration according to Illumina's instruction.
    - Prepare custom read 1 sequencing primer (Read1CustomSeqB), custom read 2 primer (DropPLA_Read2) and custom i7 index read primer (DropPLA_i7Read) according to Illumina's instructions.
    - Read distribution was 20 bases for read 1, 85 bases for read 2 and 8 bases for i7 index read.

    ▲ **CRITICAL STEP** A high percentage of PhiX is required to obtain quality data from PLA products.

    ◆ **TROUBLESHOOTING**
133. *10×-based Prox-seq*. Sequence the cDNA and PLA libraries separately on the NextSeq 550 platform.
    - Sequence the cDNA library by using a high-output 150-cycle kit (read distribution: read 1, 28 cycles; i7 index, 10 cycles; i5 index, 10 cycles; read 2, 90 cycles) with 1% vol/vol PhiX.
    - Sequence the PLA library by using a mid-output 150-cycle kit (read distribution: read 1, 28 cycles; i7 index, 8 cycles; read 2, 75 cycles) with 40% vol/vol PhiX. In addition, spike in custom read 2 primer (DropPLA_Read2) to Illumina read 2 primer and prepare custom i7 read primer (DropPLA_i7Read) according to Illumina's instructions.

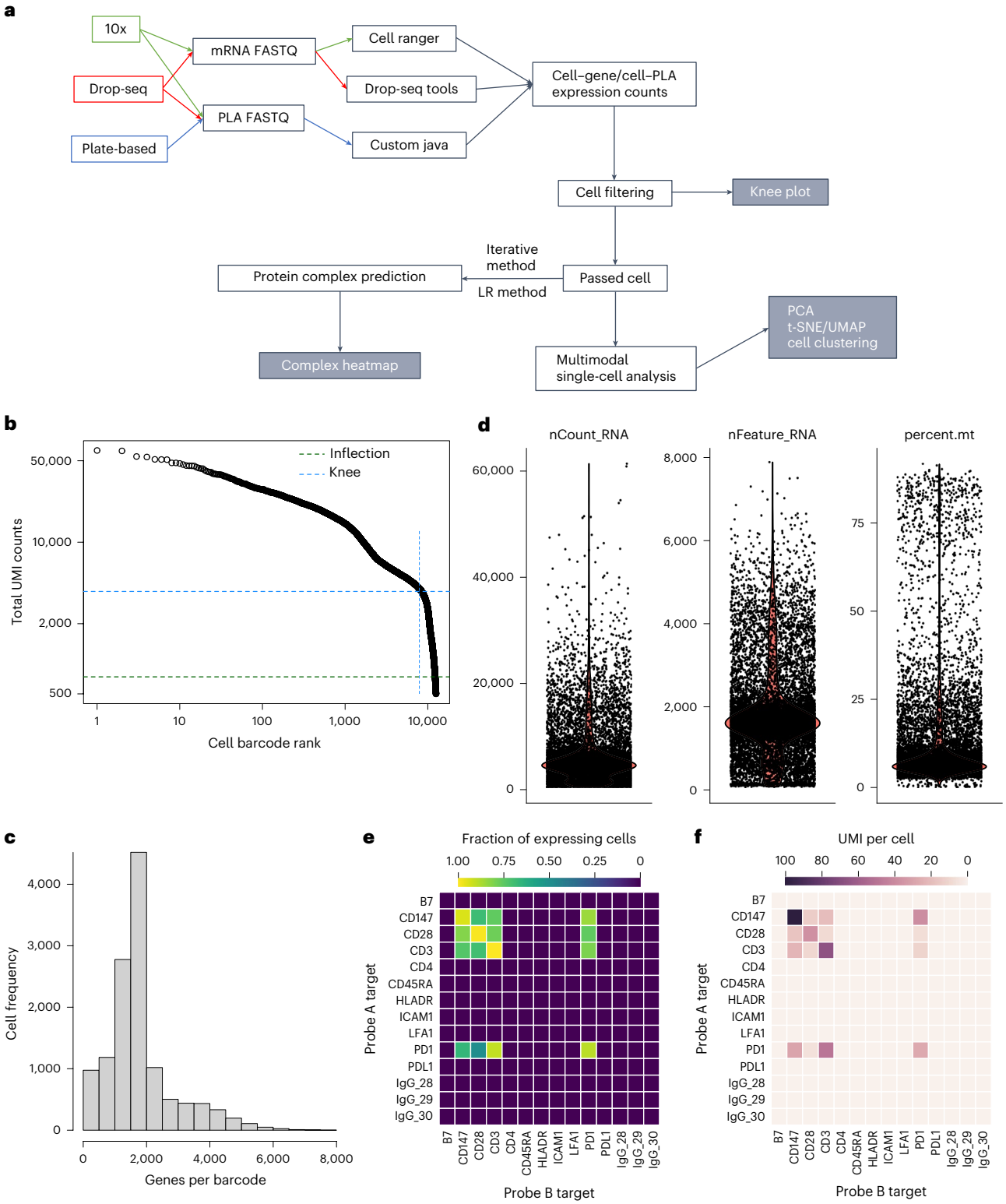
    ◆ **TROUBLESHOOTING**
134. *Plate-based Prox-seq*. Sequence the pooled PLA libraries with a mid-output NextSeq kit.
    - Spike in PhiX at a 40% concentration according to Illumina's instruction.
    - Prepare custom read 1 sequencing primer (SmartPLA_Read1), custom i5 index read primer (SmartPLA_i5Read) and custom i7 index read primer (SmartPLA_i7Read) according to Illumina's instructions.
    - Read distribution was 75 bases for read 1, 8 bases for i5 index read and 8 bases for i7 index read.

    ◆ **TROUBLESHOOTING**

Alignment of cDNA sequencing reads

● **TIMING ~4–6 h (depending on the computer's number of CPU cores and memory)**

135. First, convert raw sequence BCL files to FASTQ format with BCL Convert from Illumina. This conversion is performed automatically if the user has opted to save the sequencing data to Illumina BaseSpace while setting up the next-generation sequencing run.
136. If Drop-seq was used, remove the mosaic sequence from the Nextera transposase by trimming the first 19 bases of read 2 by using the FASTX-toolkit.
137. Download the reference genome and annotation files for the organism of interest from a public database such as the National Center for Biotechnology Information or Ensembl.
138. Align the mRNA sequencing data by using the appropriate software.
    (A) **Drop-seq (For detailed instructions from the McCarroll Lab[10], see** https://github.com/broadinstitute/Drop-seq/blob/master/doc/Drop-seq_Alignment_Cookbook.pdf.)
        (i) Create a STAR genome index.
        (ii) Align by using Drop-seq tools v2.3.0. The first step is to preprocess the FASTQ file by using the `PreprocessReads` command.



(iii) Generate an index for the reference by using the commands `CreateSequenceDictionary` and `CreateIntervals`.

(iv) Align the preprocessed BAM file from the first step to the reference genome by using the `AlignReads` command. This command generates a BAM file with aligned reads.

Fig. 3 | Overview of Prox-seq computational workflow. a, Computational pipeline of processing single-cell mRNA and Prox-seq data, from alignment to protein complex prediction and multimodal single-cell downstream analysis. **b**, Cell filtering based on UMIs to identify cell barcodes that probably correspond to valid single cells[1]. The inflection represents the point when UMI counts start decreasing rapidly, and the knee represents the cutoff of UMI counts to differentiate cells valid for analysis. **c**, Cell filtering based on numbers of detected genes per cell. Cells containing >500 genes are discarded[1]. **d**, Cell filtering based on the percentage of UMIs per barcode associated with mitochondrial genes. Cells containing >15% mitochondrial genes are discarded[1]. **e**, Heatmap showing the fraction of Jurkat cells containing at least one protein complex. **f**, Heatmap showing the mean abundance of protein complexes in Jurkat cells. PCA, principal component analysis; t-SNE, t-distributed stochastic neighbor embedding; UMAP, uniform manifold approximation and projection.

(v) Merge the BAM tags in the first step with aligned reads with command `MergeBamAlignment`.

(vi) Filter the tagged BAM file and check for bead synthesis error. This can be done with command `TagBamWithReadSequenceExtended`.

(vii) Generate a digital gene expression matrix by using the `DigitalExpression` command for downstream analysis.

(B) **10× (For detailed instructions from 10× Genomics, see** https://www.10xgenomics.com/support/software/cell-ranger/latest/tutorials**).**

(i) Align by using Cell Ranger v6.1.1 or higher. First, set up the environment by running the command `cellranger mkref` to build a reference genome index or download premade ones from the 10× website.

(ii) Use the command `cellranger count` and perform alignment with the correct locations of input files and output directory. The output includes gene-barcode, feature-barcode and cell-barcode matrices, which can be loaded in R with the Read10× package for downstream analysis.

Alignment and processing of PLA sequencing reads (detailed guidance is available at https://github.com/tay-lab/Prox-seq/blob/master/alignment_manual.pdf)

● **TIMING ~1–2 h**

139. First, perform alignment by extracting the cell barcode, the unique molecular identifier (UMI) and the Prox-seq probe identities from the FASTQ files. Use command `ReadAlignment10x`, `ReadAlignmentDropseq` or `ReadAlignmentSmartseq` depending on the workflow choice.
◆ **TROUBLESHOOTING**
140. Produce the cell reference list either by Drop-seq tools v2.3.0 function `BAMTagHistogram` or by using function `BulidCellBarcodes` embedded in the custom program or by Cell Ranger function `cellranger count`. This step is not required for plate-based data.
141. Next, perform a cell barcode correction with the `CellBarcodeCorrection` command for Drop-seq or 10× libraries based on the reference list above. This step is not required for plate-based data. For Drop-seq, it is optional to use command `ReadcountHistogram` to export the read count of each single-cell barcode and plot the knee plot (Fig. 3b) to identify good single-cell barcodes.
◆ **TROUBLESHOOTING**
142. Execute the `UMIMerging` command, which merges PLA products from the same single cell with similar UMI sequences together.
143. Export the UMI merged reads into a DPE matrix as a text file by performing the `DigitalCount` command.

Calculation of protein abundance

● **TIMING ~30 min**

144. To determine protein expression level, import PLA product count data to create a PLA object.
145. Perform the 'calculateProteinCount' method, which gives a dataframe attribute called 'protein_count'.
◆ **TROUBLESHOOTING**

Protein complex quantification

● **TIMING ~30 min**

146. To quantify protein complexes, first import PLA product count data to create a PLA object.
147. Call the method 'predictComplex' to generate a dataframe attribute named 'complex_coun't. The default argument for 'method' is 'iterative'. The alternative method is 'lr', which constructs a weighted least squares regression to estimate counts of protein complexes. This method is compatible only with the 'free oligo' variant of Prox-seq (see Step 38).
◆ **TROUBLESHOOTING**

Troubleshooting

Troubleshooting advice can be found in Table 2.

Table 2 | Troubleshooting table

Step	Problem	Possible reason	Solution
22	Low yield of antibody–DBCO conjugate	Low input mass	The antibody–DBCO reaction is more reliable at higher scale (>50 µg). These reactions are most reliable at 100 µg
		Excessive DBCO/antibody ratio	Very high conjugation levels (>10 DBCO/antibody) decrease yield. Either reduce the concentration of DBCO or increase the concentration of antibody
40	Few cells after sample preparation	Loss of cell viability	Reduce the number of wash steps before enrichment of live cells (Step 35)
		Poor cell pelleting	Increase the centrifugation speed. This problem can occur with fixed samples because they often lose membrane integrity, causing a decrease in density
		Cells stuck to surfaces	If cells are transferred into a plastic tube via PBS alone, they will often stick to the tube. Ensure that all solutions include protein (such as BSA) or that all tubes have their surfaces passivated with a protein solution
	Low cell viability	Cells are overprocessed during Prox-seq	Reduce the number of wash steps before enrichment of live cells (Step 35)
64, 86, 94, 102	Low concentration of the libraries	Low elution efficiency	Prevent any loss of the magnetic beads during the cleanup process. Avoid over-drying the magnetic beads. Always check the supernatants for carryover beads before discarding
64, 86, 94, 102	Incorrect size distribution of the library fragment	Inefficient library cleanup	Peaks <200 bp indicate the presence of adapter dimers. Minor peaks do not interfere in sequencing and can be ignored. Repeat the cleanup process if the peaks form a major part of the library. However, the additional cleanup step will reduce the library yield
132, 133, 134	Low Phred quality score when sequencing	Insufficient PhiX spike in	PhiX is critical for obtaining good sequencing data from PLA products because of the low library complexity. Increasing the amount of spike may alleviate poor-quality scores
132, 133, 134	Cluster density outside the optimal range after sequencing	Incorrect loading library concentration	Remeasure library concentration via Qubit or fragment analyzer
139	Error encountered while aligning PLA data without any output	Wrong input or output directory or wrong format of input files (e.g., AB_BC_LIST)	Check the directory and name of every input and output file. Make sure that there is no gap and empty line in the AB_BC_LIST input files. Check the format of the batch file
141	Cell barcode correction error	Wrong suffix added or wrong header parameter set	Set SUFFIX to '–1' when aligning 10× data and to 'None' for other modes. Check if the referred cell barcode list has a header and determine the HEADER parameter correspondingly
145	Failure to identify positive-control protein complexes or identification of implausible protein complexes	Application of each protein complex quantification method requires a homogeneous population of single cells	Use single-cell mRNA data to cluster data into cell types and then apply the quantification method to individual cell types
147	Encounter a run error while applying the algorithm for protein abundance or protein complex abundance calculations	The DPE matrix is not prefiltered to discard 'all-zero' rows or columns, which results in a zero denominator error	Typically for PLA product data, we will do double filtering, which includes filtering out invalid cell barcodes according to mRNA-based barcode correction and filtering out cells with too few or too many PLA product UMIs according to the distribution of the UMI count per cell

Timing

Steps 1–26, preparing Prox-seq probes: 2 h of work time, overnight incubation
Steps 27–40, sample processing (PLA): ~3 h
Steps 41–86, Drop-seq-based Prox-seq: ~8 h
Steps 87–102, 10×-based Prox-seq: ~8 h
Steps 103–131, plate-seq-based Prox-seq: ~6 h or more, depending on the number of plates
Steps 132–134, next-generation sequencing: ~1–2 d
Steps 135–138, alignment of cDNA sequencing reads: ~4–6 h (depending on the computer's number of CPU cores and memory)
Steps 139–143, alignment and processing of PLA sequencing reads: ~1–2 h
Steps 144–145, calculation of protein abundance: ~30 min
Steps 146–147, protein complex quantification: ~30 min

Anticipated results

This Prox-seq protocol includes several QC measures that can identify successful steps. We have highlighted these throughout the text. Overall, cDNA peaks should be in the 1,000-bp range after whole transcriptome amplification and in the 500-bp range after tagmentation (Fig. 2c,d). These ranges are typically for single-cell sequencing. After generation of the cell-by-gene matrix of gene expression counts, QC of single-cell RNA-seq data in Prox-seq follows current mainstream filtering methods. The common metrics and methods include filtering cells by RNA UMI counts, number of RNA features and percentage of mitochondria reads (Fig. 3c,d). PLA products analyzed by TapeStation should be in the range of 100 bp after initial PLA product amplification. After indexing, PLA products should be ~228 bp (Fig. 2d,f,g). The user should be careful to use the correct volume of AMPure beads during purification, because inefficient library cleanup can lead to poor QC results (Fig. 2h). Some knowledge of the protein complex milieu in the researcher's sample can aid in interpretation. Homodimers (e.g., CD147 homodimers detected on Jurkat T cells) are particularly useful for this, because they are common and can serve as positive controls that should be called as protein complexes during data analysis (Fig. 3e,f).

Reporting summary

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

Data availability

Source data for the figures are available at https://github.com/tay-lab/NProt_data-src.

Code availability

Custom code used to reproduce Figs. 2 and 3 and Python scripts for PLA data analysis are available at https://github.com/tay-lab/NProt_data-src.

Received: 17 May 2023; Accepted: 24 May 2024;
Published online: 15 August 2024

References

1. Vistain, L. et al. Quantification of extracellular proteins, protein complexes and mRNAs in single cells by proximity sequencing. *Nat. Methods* **19**, 1578–1589 (2022).
2. Sharifi Tabar, M. et al. Illuminating the dark protein-protein interactome. *Cell Rep. Methods* **2**, 100275 (2022).
3. Luck, K. et al. A reference map of the human binary protein interactome. *Nature* **580**, 402–408 (2020).
4. Rhee, H.-W. et al. Proteomic mapping of mitochondria in living cells via spatially restricted enzymatic tagging. *Science* **339**, 1328–1331 (2013).

5. TotalSeqTM-B Human Universal Cocktail, V1.0. Available at https://www.biolegend.com/nl-nl/products/totalseq-b-human-universal-cocktail-v1dot0-20960?GroupID=GROUP28
6. Zhang, W., Xie, M., Shu, M.-D., Steitz, J. A. & DiMaio, D. A proximity-dependent assay for specific RNA–protein interactions in intact cells. *RNA* **22**, 1785–1792 (2016).
7. Darmanis, S. et al. ProteinSeq: high-performance proteomic analyses by proximity ligation and next generation sequencing. *PLoS ONE* **6**, e25583 (2011).
8. Darmanis, S. et al. Sensitive plasma protein analysis by microparticle-based proximity ligation assays. *Mol. Cell. Proteom.* **9**, 327–335 (2010).
9. Söderberg, O. et al. Direct observation of individual endogenous protein complexes in situ by proximity ligation. *Nat. Methods* **3**, 995–1000 (2006).
10. Macosko, E. Z. et al. Highly parallel genome-wide expression profiling of individual cells using nanoliter droplets. *Cell* **161**, 1202–1214 (2015).
11. Picelli, S. et al. Full-length RNA-seq from single cells using Smart-seq2. *Nat. Protoc.* **9**, 171–181 (2014).
12. Xia, J. et al. Computational prediction of protein interactions in single cells by proximity sequencing. *PLoS Comput. Biol.* **20**, e1011915 (2024).
13. Gong, H. et al. Simple method to prepare oligonucleotide-conjugated antibodies and its application in multiplex protein detection in single cells. *Bioconjug. Chem.* **27**, 217–225 (2016).
14. Xia, J. et al. Computational prediction of protein interactions on single cells by proximity sequencing. *PLoS Comput. Biol.* **20**, e1011915 (2024).
15. van Bergenhenegouwen, J. et al. TLR2 & Co: a critical analysis of the complex interactions between TLR2 and coreceptors. *J. Leukoc. Biol.* **94**, 885–902 (2013).
16. Shevyrev, D. & Tereshchenko, V. Treg heterogeneity, function, and homeostasis. *Front. Immunol.* **10**, 3100 (2020).
17. Wang, X., Yu, L. & Wu, A. R. The effect of methanol fixation on single-cell RNA sequencing data. *BMC Genomics* **22**, 420 (2021).
18. Phan, H. V. et al. High-throughput RNA sequencing of paraformaldehyde-fixed single cells. *Nat. Commun.* **12**, 5636 (2021).
19. Stoeckius, M. et al. Cell Hashing with barcoded antibodies enables multiplexing and doublet detection for single cell genomics. *Genome Biol.* **19**, 224 (2018).
20. Kaur, A. & Dhakal, S. Recent applications of FRET-based multiplexed techniques. *Trends Anal. Chem.* **123**, 115777 (2020).
21. Alsemarz, A., Lasko, P. & Fagotto, F. Limited significance of the in situ proximity ligation assay. Preprint at *bioRxiv* https://doi.org/10.1101/411355 (2018).
22. Peterson, V. M. et al. Multiplexed quantification of proteins and transcripts in single cells. *Nat. Biotechnol.* **35**, 936–939 (2017).
23. Stoeckius, M. et al. Simultaneous epitope and transcriptome measurement in single cells. *Nat. Methods* **14**, 865–868 (2017).
24. Schueder, F. et al. Unraveling cellular complexity with unlimited multiplexed super-resolution imaging. Preprint at *bioRxiv* https://doi.org/10.1101/2023.05.17.541061 (2023).
25. Löf, L. et al. Detecting individual extracellular vesicles using a multicolor in situ proximity ligation assay with flow cytometric readout. *Sci. Rep.* **6**, 34358 (2016).

Author contributions

L.V., H.V.P. and B.K. all contributed to developing the protocol. L.V., B.K., J.X. and S.T. all contributed to writing this manuscript.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41596-024-01030-x.

Correspondence and requests for materials should be addressed to Savaş Tay.

Peer review information *Nature Protocols* thanks the anonymous reviewers for their contribution to the peer review of this work.

Reprints and permissions information is available at www.nature.com/reprints.

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

Springer Nature or its licensor (e.g. a society or other partner) holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.

Related links

Key references using this protocol

Vistain, L. et al. *Nat. Methods* **19**, 1578–1589 (2022): https://doi.org/10.1038/s41592-022-01684-z
Xia, J. et al. *PLoS Comput. Biol.* **20**, e1011915 (2024): https://doi.org/10.1371/journal.pcbi.1011915

© Springer Nature Limited 2024

nature portfolio

Corresponding author(s): Savaş Tay

Last updated by author(s): Apr 1, 2024

Reporting Summary

Nature Portfolio 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 Portfolio policies, see our Editorial Policies 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	Illumina Nextseq 500/550 system was used to get the sequencing data.
Data analysis	PLA sequencing data was aligned using custom java program. Digital count matrix was analyzed in Anaconda3/Python v3.8.8. The code and data used to reproduce figures are deposited at https://github.com/tay-lab/NProt_data-src.

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 and reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Portfolio 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 description of any restrictions on data availability
- For clinical datasets or third party data, please ensure that the statement adheres to our policy

The code and data used to reproduce figures are deposited at https://github.com/tay-lab/NProt_data-src.
The raw sequencing data and processed countdata are deposited in NCBI Gene Expression Omnibus (accession numbers GSE149574 and GSE196130).

Human research participants

Policy information about studies involving human research participants and Sex and Gender in Research.

Reporting on sex and gender	*Use the terms sex (biological attribute) and gender (shaped by social and cultural circumstances) carefully in order to avoid confusing both terms. Indicate if findings apply to only one sex or gender; describe whether sex and gender were considered in study design whether sex and/or gender was determined based on self-reporting or assigned and methods used. Provide in the source data disaggregated sex and gender data where this information has been collected, and consent has been obtained for sharing of individual-level data; provide overall numbers in this Reporting Summary. Please state if this information has not been collected. Report sex- and gender-based analyses where performed, justify reasons for lack of sex- and gender-based analysis.*
Population characteristics	*Describe the covariate-relevant population characteristics of the human research participants (e.g. age, genotypic information, past and current diagnosis and treatment categories). If you filled out the behavioural & social sciences study design questions and have nothing to add here, write "See above."*
Recruitment	*Describe how participants were recruited. Outline any potential self-selection bias or other biases that may be present and how these are likely to impact results.*
Ethics oversight	*Identify the organization(s) that approved the study protocol.*

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

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	Sample size in our case is number of cells, which was determined by the pipelines where we can recover single cells. Typically, we can get thousands of cells from either 10X or drop-seq pipelines, and hundreds of cells from plate-seq pipeline.
Data exclusions	Invalid cells were filtered based on total UMI counts, feature numbers, and percent of mitochondria genes.
Replication	Since we got thousands of cells per experiment, there are multiple replicates to give us statistical confidence. Besides, same experiments (e.g. same probe panels, same cell lines) were performed several times to confirm the discovery. Data from different technology (e.g. Prox-seq, flow cytometry) and different pipelines (e.g. 10X, plate-seq) were compared.
Randomization	Cells were identified and analyzed by algorithm to reduce the bias.
Blinding	Researchers were not aware of each cell's identity in Drop-seq and 10X pipeline. Cells were clustered in an unsupervised manner by their RNA or PLA features.

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 and archaeology
☒	☐ Animals and other organisms
☒	☐ Clinical data
☒	☐ Dual use research of concern

Methods

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

Antibodies

Antibodies used: The full list of antibodies used in this study is displayed in Supplementary Table S1.

Validation: The antibodies were validated by the manufacturer via flow cytometry. For this study, the antibodies were supplied by Biolegend, R&D Systems and Millipore. Validation data can be found on the manufacturer's website using the antibody catalog number supplied in the Supplementary Table.

Eukaryotic cell lines

Policy information about cell lines and Sex and Gender in Research

Cell line source(s): Jurkat and Raji cell lines were a gift from Prof. Jun Huang. Originally, the cells were purchased from ATCC.

Authentication: The expression levels of a few key proteins in our cell lines were used to draw the conclusions we make in the manuscript. Flow cytometry was used to measure each of these proteins directly.

Mycoplasma contamination: Not tested

Commonly misidentified lines (See ICLAC register): None

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: Samples were taken directly from cell culture without any pre-processing steps.

Instrument: Fortessa 4-15 (BD Biosciences)

Software: Data were collected with FACS Diva and analyzed using FlowJo V_10.8.1

Cell population abundance: Cell population study was not relevant to this study

Gating strategy: Gating was only applied to identify single cells. Cells were identified using forward and side scatter, and single cells were isolated using forward height and width scatter.

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