nature methods

REVIEW ARTICLE

https://doi.org/10.1038/s41592-020-01010-5

Deciphering molecular interactions by proximity labeling

Wei Qin[1,7], Kelvin F. Cho[1,2,7], Peter E. Cavanagh[3,7] and Alice Y. Ting[1,4,5,6] ✉

Many biological processes are executed and regulated through the molecular interactions of proteins and nucleic acids. Proximity labeling (PL) is a technology for tagging the endogenous interaction partners of specific protein 'baits', via genetic fusion to promiscuous enzymes that catalyze the generation of diffusible reactive species in living cells. Tagged molecules that interact with baits can then be enriched and identified by mass spectrometry or nucleic acid sequencing. Here we review the development of PL technologies and highlight studies that have applied PL to the discovery and analysis of molecular interactions. In particular, we focus on the use of PL for mapping protein–protein, protein–RNA and protein–DNA interactions in living cells and organisms.

Cellular functions are tightly regulated by proteins, nucleic acids and their interactions, including protein–protein interactions (PPIs), protein–RNA interactions and protein–DNA interactions[1,2]. Such molecular interaction networks are central to most biological processes, while their dysregulation has been linked to a variety of human diseases including cancers, immune disorders and neurodegeneration. Methods enabling the large-scale discovery of molecular interactions in living cells have provided insights for biological exploration and therapeutic intervention.

The traditional approaches of affinity purification and yeast two-hybrid have been widely applied to discover potential molecular interactions[3,4]. Antibody-based affinity purification, in combination with mass spectrometry–based proteomics, allows the enrichment and identification of stable interaction partners of specific proteins of interest. Such efforts have expanded our understanding of protein interaction networks in a variety of systems, including yeast, flies and human cells. Affinity purification can also be combined with crosslinking and nucleic acid sequencing to interrogate protein–nucleic acid interactions, such as in chromatin immunoprecipitation sequencing (ChIP-seq) and RNA immunoprecipitation sequencing (RIP-seq)[5,6]. The main limitation of affinity purification, however, is that weak or transient interactions are often lost during cell lysis and the subsequent washing steps. To overcome this, affinity purification can be combined with crosslinking[7]; however, this increases the rate of false positives. Moreover, affinity purification is challenging to apply to insoluble targets or protein baits lacking high-affinity antibodies.

Yeast two-hybrid and other protein complementation assays represent another approach for mapping protein–protein, protein–RNA and protein–DNA interactions in living cells[3]. These approaches are often high throughput, enabling the screening of thousands to millions of potential molecular interactions[8]. However, many protein complementation assays have cell type and organelle type restrictions (for example, yeast two-hybrid does not work on membrane proteins), false positives due to overexpression and tagging of both bait and prey, and false negatives due to steric interference by or geometric constraints of the required tags.

Proximity labeling (PL) was developed to provide a complementary approach to these traditional methods for molecular interaction mapping in living cells. PL uses engineered enzymes, such as peroxidases (engineered ascorbate peroxidase 2 (APEX2)[9], horseradish peroxidase (HRP)[10]) or biotin ligases (BioID[11,12], BioID2[13], BASU[14], TurboID[15], miniTurbo[15]), that are genetically tagged to a protein of interest (Table 1). The PL enzyme converts an inert small-molecule substrate into a short-lived reactive species, such as a radical in the case of APEX[16] or an activated ester in the case of BioID and TurboID[17], that diffuses out from the enzyme active site to covalently tag neighboring endogenous species (Fig. 1a,b). The labeling radius is determined by both the half-life of the reactive species and the concentration of quenchers in the environment, such as glutathione for APEX and amines for BioID and TurboID. The experimentally determined labeling radii for HRP, APEX, BioID and TurboID enzymes fall in the range of 1–10 nm in living cells[16,17]. However, instead of a fixed radius, it is more accurate to think of labeling by PL enzymes as a 'contour map' in which the reactant concentration is highest at the PL enzyme and falls off nanometer by nanometer from the source[17,18]. Peroxidase- and biotin ligase–generated reactive species are also membrane impermeant[19], and thus the contour map ends at membrane boundaries. The substrate molecule typically contains a biotin handle to enable subsequent enrichment of tagged species using streptavidin beads and their identification by mass spectrometry (for proteins) or nucleic acid sequencing (for RNA) (Fig. 1c). Depending on the localization and expression of the PL enzyme, PL can be used to interrogate spatial proteomes on several different length scales—from entire cells[20] to organelles and subcellular compartments[18,19,21–25] to macromolecular complexes[26,27]. In this Review, we will focus on the application of PL to the study of molecular interactions, which has more published examples than organelle or cellular mapping[28].

For this review, we define PL as labeling catalyzed by genetically encoded enzymes (as opposed to chemical catalysts) that generate diffusible reactive species in living systems. Many other conceptually related technologies have been described but fall outside the scope of this Review. These include proximity ligation assays

[1]Department of Genetics, Stanford University, Stanford, CA, USA. [2]Cancer Biology Program, Stanford University, Stanford, CA, USA. [3]Department of Biochemistry, Stanford University, Stanford, CA, USA. [4]Department of Biology, Stanford University, Stanford, CA, USA. [5]Department of Chemistry, Stanford University, Stanford, CA, USA. [6]Chan Zuckerberg Biohub, San Francisco, CA, USA. [7]These authors contributed equally: Wei Qin, Kelvin F. Cho, Peter E. Cavanagh. ✉e-mail: ayting@stanford.edu

Table 1 | Overview of PL enzymes

Enzyme	Type	Size (kDa)	Labeling time	Modification sites	Advantages	Limitations
APEX	Peroxidase	28	1 min	Tyr, Trp, Cys, His	High temporal resolution; versatility for both protein and RNA labeling	Limited application in vivo because of the toxicity of H_2O_2
APEX2	Peroxidase	28	1 min	Tyr, Trp, Cys, His	High temporal resolution; versatility for both protein and RNA labeling	Limited application in vivo because of the toxicity of H_2O_2
HRP	Peroxidase	44	1 min	Tyr, Trp, Cys, His	High temporal resolution; versatility for both protein and RNA labeling	Limited application in vivo because of the toxicity of H_2O_2; limited to secretory pathway and extracellular applications
BioID	Biotin ligase	35	18 h	Lys	Non-toxic for in vivo applications	Poor temporal resolution as a result of low catalytic activity
BioID2	Biotin ligase	27	18 h	Lys	Non-toxic for in vivo applications	Poor temporal resolution as a result of low catalytic activity
BASU	Biotin ligase	29	18 h	Lys	Non-toxic for in vivo applications	Poor temporal resolution as a result of low catalytic activity
TurboID	Biotin ligase	35	10 min	Lys	Highest activity biotin ligase; non-toxic for in vivo applications	Potentially less control of labeling window as a result of high biotin affinity
miniTurbo	Biotin ligase	28	10 min	Lys	High activity; non-toxic for in vivo applications; smaller than TurboID	Lower catalytic activity and stability as compared to TurboID

on fixed cells with antibody[29] or nucleic acid probes[30], photocross-linking with unnatural amino acids[31], promiscuous enzymes that label with non-diffusible substrates (for example, PUP-IT)[32] and light-activated chemical catalysts[33].

Proximity labeling for profiling protein–protein interactions

PL has been applied to a wide range of PPI mapping problems, from signal transduction networks (mitogen-associated protein kinase (MAPK)[34,35], Hippo[36], adrenergic[37], G-protein-coupled receptor (GPCR)[38,39]) to enzyme–substrate interactions (E3 ubiquitin ligases[40,41], kinases[42]). These studies have been conducted in a variety of cell types (for example, 2D or 3D culture[43], endothelial cells[44], neuronal cells[22,45,46]) and organisms (bacteria[15,47], yeast[9,15,48,49], flies[15,41,50,51], worms[15], plants[52,53], mice[37,54–56], primary human tissue[57]). In this section, we highlight some areas where PL has offered advantages over traditional methods in identifying molecular interactions and enabling biological discovery. These include characterizing the architecture of insoluble protein complexes (for example, the nuclear envelope), capturing transient PPIs (for example, enzyme–substrate interactions), dissecting dynamic processes (for example, GPCR signaling), and enabling the specific interrogation of interactomes in live organisms (Table 2).

PL has enabled the study of insoluble baits that are difficult to analyze by affinity purification, such as lamin A/C, a nuclear envelope resident protein critical for maintaining nuclear envelope structure[11]. To map lamin A's interaction partners by PL, Roux et al. used BioID, a promiscuous mutant of the *Escherichia coli* biotin ligase BirA[12], and fused it directly to lamin A in HeLa cells[11]. Residents of the nuclear membrane and other previously unknown interactors, such as the nuclear pore complex, were identified[11]. Subsequent work using PL enzymes have built on this work by interactome mapping of other lamins, nuclear envelope proteins[17,58,59] and nuclear transporters[60].

Protein aggregates are extreme examples of insoluble baits. Chou et al. used BioID to identify interactors of TDP43 aggregates, a common histopathological marker of neurodegenerative disease, including amyotrophic lateral sclerosis and frontotemporal dementia[61]. By fusing BioID to TDP43 to perform PL, the authors identified nucleocytoplasmic transport machinery, and follow-up studies implicated TDP43 aggregates' disruption of nucleoporin and transport factor functions as a mechanism for pathology[61].

PL has proven especially useful in dissecting signaling pathways in which upstream and downstream effectors interact only transiently with the protein of interest. For example, Amber et al. used BioID to probe interactors along the Hippo pathway[36], a highly conserved signaling cascade that controls cell proliferation and apoptosis to dictate organ size. By mapping the interactomes of 19 pathway proteins using BioID, the authors generated protein interaction networks for the Hippo pathway and identified many putative regulators and kinase substrates[36]. PL-based interactome mapping has also been used to map other signaling processes, such as NF-κB[62], Ras[63,64], MAPK[34,35] and Hedgehog[65] pathways. PL-based interactome mapping can also uncover the remodeling of signaling pathways in the context of disease[66] and upon pathway activation to discover critical mediators of signal transduction[37,39]. In addition to intracellular interactome mapping, PL has also been used to identify extracellular ligand–receptor interactions[32,67,68].

PL-based PPI mapping has also been informative for the study of enzyme–substrate interactions, which are intrinsically transient as a result of substrate turnover[40–42]. E3 ubiquitin ligases in particular, which influence many aspects of cellular biology by controlling protein ubiquitination and degradation, each have numerous adaptor proteins and substrates[69]. Etienne et al. used BioID in conjunction with pharmacological proteasome inhibition to probe interactors of SCF E3 ligases β-TrCP1 and β-TrCP2[40]. Using this approach, the authors validated 12 new substrates, including proteins involved in nuclear membrane integrity and translation control. PL has also been used to interrogate substrates of protein kinases. For instance, Cutler et al. fused BioID to p190 and p210 BCR-ABL tyrosine kinases, oncogenic protein fusions that result from chromosomal translocations[42]. Using PL, the authors identified distinct interactomes of each fusion and revealed that the Src family kinase Lyn, critical for transformation and drug resistance, is a preferential substrate of the p190 BCR-ABL fusion.

The short time frame of APEX labeling (<1 min) has been leveraged to capture temporally resolved snapshots of changing interactomes of proteins involved in dynamic cellular processes, such as in Wnt[70] and GPCR signaling[38,39]. Paek et al. applied APEX-based PL to angiotensin II type 1 receptor (AT1R) and β2-adrenergic receptor (β2AR) GPCR signaling in response to agonist activation[38], and proteomic analysis of the changing interactome supported the role of endocytosis in sequestering GPCRs from G proteins and

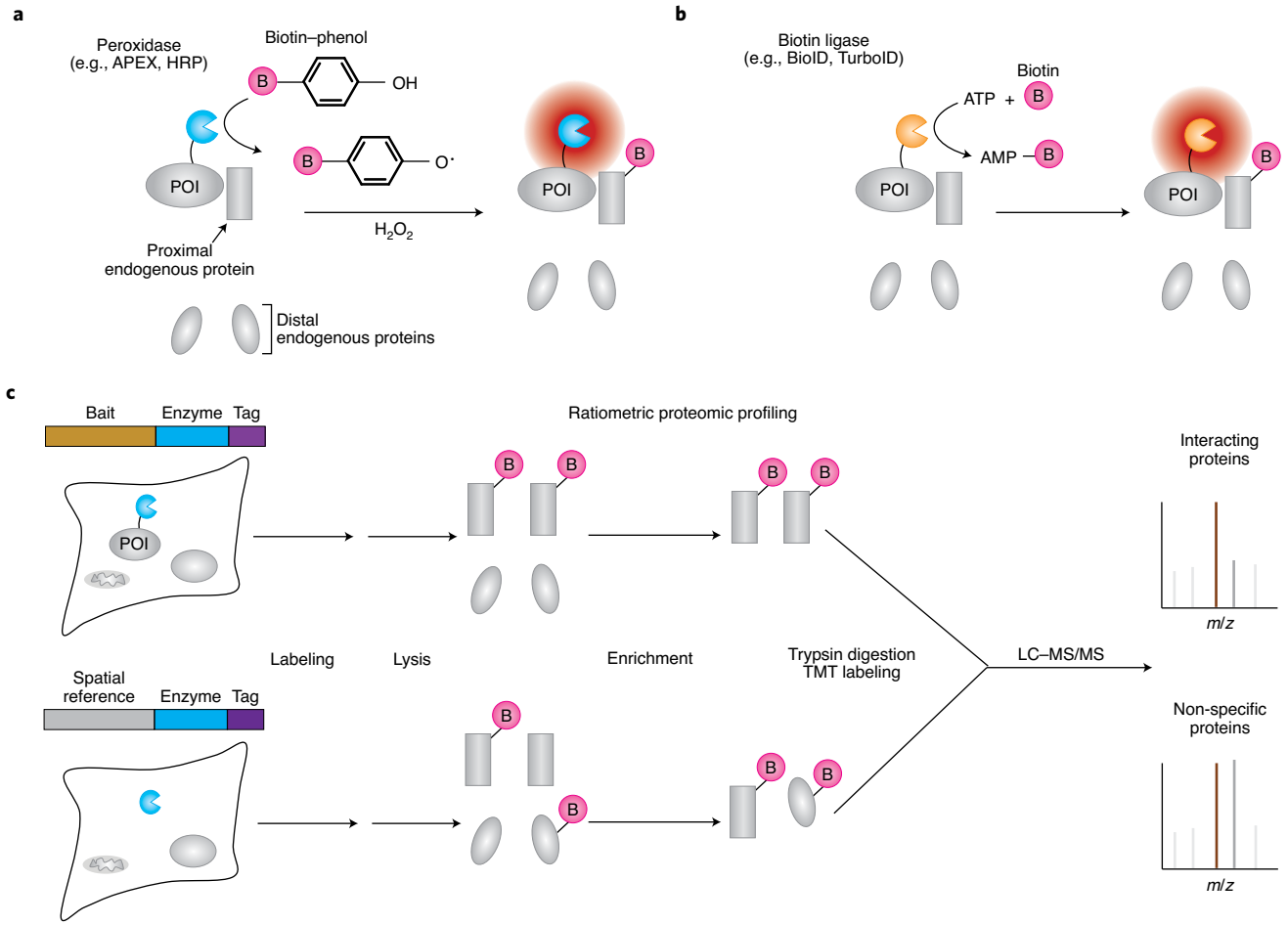

Fig. 1 | Peroxidase- and biotin ligase-based proximity labeling methods for PPI mapping. a, Peroxidase-based approaches, such as APEX or HRP, oxidize biotin-phenol into reactive phenoxyl radicals using hydrogen peroxide, which preferentially labels proximal over distal endogenous proteins. **b**, Biotin ligase-based approaches, such as BioID or TurboID, utilize ATP and biotin to catalyze the formation of reactive biotin-5′-AMP, which diffuses and labels proximal proteins. **c**, Example proteomic workflow for mapping PPIs. PL enzymes fused to the bait of interest and a spatial reference control are expressed in separate samples. Biotinylated proteins from each sample are enriched and analyzed via quantitative mass spectrometry. Proteins that preferentially interact with the bait of interest can be identified by ratiometric analysis. PPI, protein–protein interaction; POI, protein of interest; LC–MS/MS, liquid chromatography and tandem mass spectrometry; TMT, tandem mass tag.

demonstrated differing endocytosis kinetics for different GPCRs. APEX has also been used to capture snapshots of the δ-opioid receptor (DOR) interactome following treatment with agonist[39]. By identifying a time course of protein interactions and using a set of spatial references to increase specificity in the context of receptor internalization and trafficking, Lobingier et al. implicated two ubiquitin-pathway proteins as mediators of DOR endosomal trafficking to the lysosome[39]. APEX has also been used to dissect the specificity of Wnt signaling. After demonstrating that Wnt9a signals by binding the Fzd9b receptor through an unknown factor, Grainger et al. leveraged the rapidity of APEX labeling to map the proteome specifically during receptor activation and identified the epidermal growth factor receptor (EGFR) as a key mediator of Wnt9a–Fzd9b interactions[70]. Overall, these studies and others have capitalized on the rapid in situ labeling of APEX to dissect their respective pathways on a minute time scale, demonstrating the full potential of PL to probe dynamic interactions.

Many proteins participate in multiple distinct protein complexes that each carry out different cellular functions, but fusing PL enzymes directly to the bait in these scenarios would result in labeling proximal interactors of each complex, thereby reducing the confidence for those of a certain subpopulation. To overcome this, PL tools have been further adapted using various strategies for mapping interactomes of specific subcellular pools of a particular protein of interest, with the potential to dramatically improve specificity. For example, James et al. developed a strategy to probe only the inner nuclear membrane–localized pool of VAMP (vesicle-associated membrane protein)-associated protein B (VAPB), which localizes to both the endoplasmic reticulum membrane and nuclear membrane. By taking advantage of the chemically inducible dimerization FRB–FKBP system, the authors employed rapamycin-dependent recruitment of nuclear-targeted APEX2–FKBP to inner nuclear membrane–localized FRB-VAPB but not endoplasmic reticulum membrane–localized FRB-VAPB[71]. More generalizable PL approaches for increasing spatial specificity have been developed in the form of split PL enzymes. Split PL enzymes consist of two inactive fragments that can be brought together by PPIs or membrane apposition to reconstitute enzymatic activity[20,72–76]. Split-APEX2[75], split-HRP[20], various versions of split-BioIDs[72–74], and split-TurboID[76] have all been developed, with advantages and disadvantages mirroring those of their full-length counterparts. While not yet widely adapted for PPI mapping, the application of split enzymes for PL could markedly improve spatial specificity for mapping certain PPIs. For example, Schopp et al. successfully used split-BioID to probe interactors of

Table 2 | Examples of proximity labeling for mapping PPIs

PPI category	Notes	Enzyme	Baits	Refs.
Protein aggregates	Insoluble complexes by definition	BioID	TDP43 aggregates	61
Nuclear membrane and nuclear structures	Low-solubility complexes as a result of membrane function and/or complex size	BioID	Lamin A	11
		BioID	Lamin B1	59
		BioID	Various nuclear transport receptors	60
		BioID2	Lamin A, Sun2	58
Enzyme–substrate interactions	Low-affinity or transient interactions as a result of enzyme turnover	BioID	Hippo pathway (including Mst1/Mst2 kinases)	36
		BioID	p190/p210 BCR-ABL kinases	42
		APEX2	p38 MAPK	34
		BioID2	p38 MAPK	35
		BioID	SCF E3 ligases	40
		APEX2	KREP, Kelch E3 ligase adaptors	41
		BioID	ClpP protease	124
Other signaling pathways	Low-affinity or transient interactions	BioID2	TLR9, MYD88 (NF-κB pathway)	62
		BioID2	KRas4B	63
		APEX2	$Ca_v1.2$ GPCR (adrenergic pathway)	37
Intracellular sorting	Transient interactions, low-affinity interactors for trafficking machinery	BioID	Golgin-97, Golgin-245	125
		APEX2	LAMP1	45
		BioID2	Golgi glycosyltransferases	126
Dynamic processes	APEX for minute-scale interactome capture	APEX2	DOR (GPCR)	39
		APEX2	AT1R, β2AR (GPCRs)	38
		APEX2	Fzd9b (GPCR)	70
		APEX2	Gal8, Gal3, Gal9	127
		APEX2	TssA (bacteria)	47
In vivo PL in plants	PL in plant systems	BioID	HopF2	52
		BioID	AvrPto	77
		TurboID	N NLR	78
		TurboID	FAMA	53
In vivo PL in other organisms	Biotin ligase–based in vivo PL	BioID	Sun1 (*Dictyostelium*)	79
		BioID	CDK5RAP2 (*Dictyostelium*)	80
		BioID	ISP3 (*Toxoplasma gondii*)	82
		BioID	Cyst wall proteins (*T. gondii*)	85
		BioID	TbMORN1 (*Trypanosoma brucei*)	81
		BioID	TbPLK (*T. brucei*)	83
		BioID	Parasitophorous vacuole (*Plasmodium falciparum*)	84
		BioID	c-MYC (mouse xenograft)	56
		BioID	Gephyrin (mouse)	55
		TurboID	Rmt3 (*Schizosaccharomyces pombe*)	49
		TurboID	Dcp-1, Drice, Dronc (*Drosophila melanogaster*)	86

the miRNA-induced silencing complex (miRISC)[73]. During complex maturation, the protein subunit Ago2 participates in two distinct subcomplexes containing either Dicer or TNRC6. By fusing fragments of split-BioID to Ago2 and TNRC6, and then to Ago2 and Dicer, the authors were able to differentiate distinct interactomes of each of the respective subcomplexes[73].

The development of biotin ligase–based PL approaches has also enabled PL studies in vivo across organisms. While peroxidase-based approaches have been applied in various ex vivo studies[37,50], the requirement for hydrogen peroxide limits their use in vivo. Furthermore, peroxidase-based PL in plants is problematic because of background activity from endogenous plant peroxidases. BioID has been applied for proteomic mapping in *Arabidopsis thaliana* and *Nicotiana benthamiana*[52,77], two key plant models. The development of more active TurboID and miniTurbo has improved these approaches[53]. Zhang et al. used TurboID to identify interactors of a plant immune receptor called N[78]. By using TurboID to perform biotin labeling in live *N. benthamiana* plants, the authors identified the interactor UBR7, a putative E3 ligase that downregulates N and mediates immunity against plant pathogens.

Studies in many model and non-model organisms have benefited from the simple and non-toxic labeling conditions of biotin ligase–based PL. PL has been carried out in live bacteria[15,47], yeast[15,48,49], slime molds[79,80], various parasites[81–85], worms[15], flies[15,86,87] and mice[55,56]. In the first in vivo mouse PL study, Dingar et al. fused BioID to the oncogene c-Myc, expressed this fusion construct in

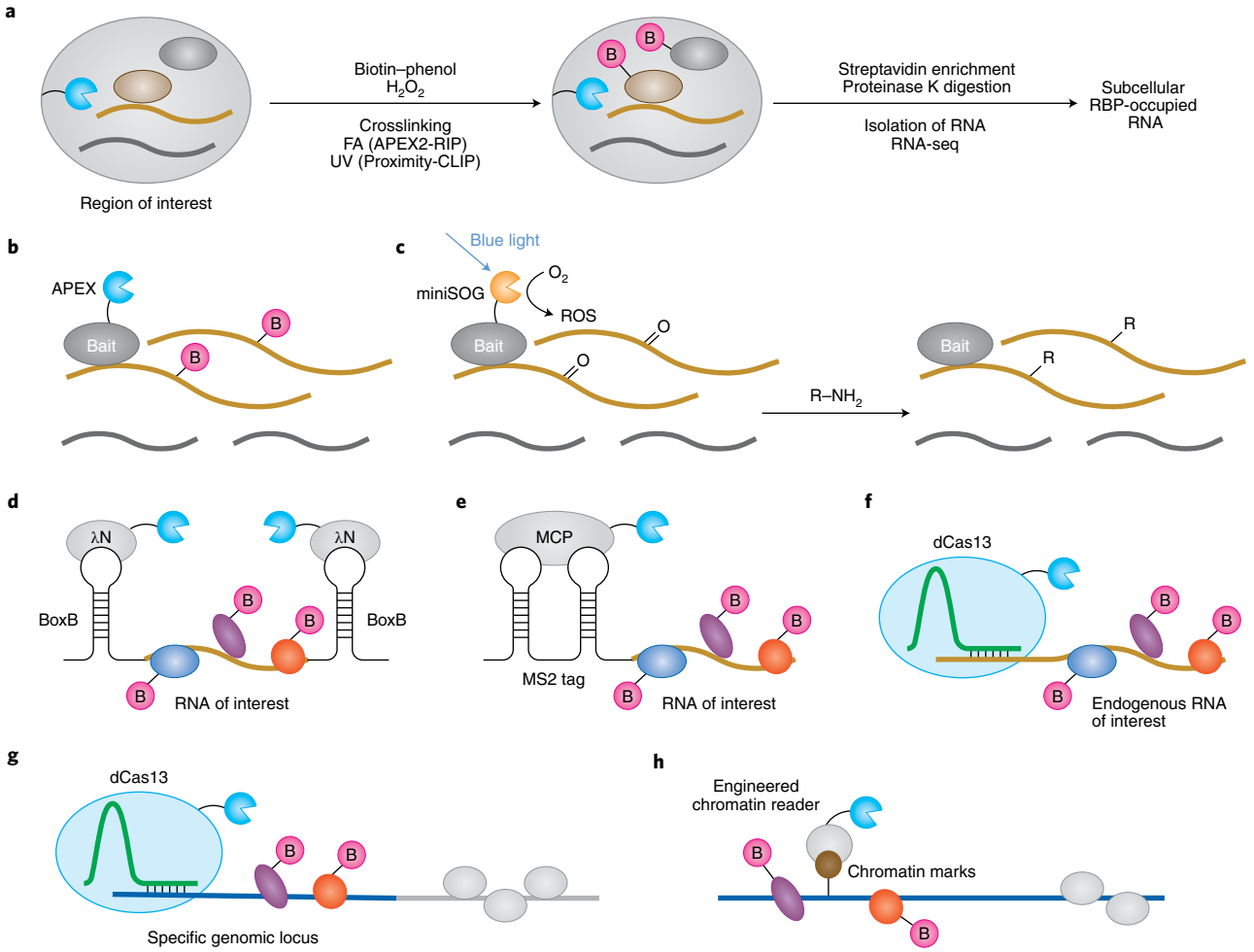

Fig. 2 | PL-based methods to investigate protein–nucleic acid interactions. a, APEX-RIP and Proximity-CLIP. APEX targeted to a specific subcellular location catalyzes the biotinylation (pink B) of proximal proteins, and the protein–RNA interactions are subsequently crosslinked by either UV or formaldehyde (FA). Crosslinked RNAs can be captured by streptavidin-based enrichment. **b**, Schematic of APEX-seq. APEX directly biotinylates RNA proximal (yellow) but not distal (gray) to a protein bait. **c**, Cap-seq. Upon blue light illumination, miniSOG generates reactive oxygen species (ROS) that react with guanine nucleobases in RNA. The photo-oxidation intermediates are intercepted by amine probes ($R\text{-}NH_2$) to form covalent adducts. **d**, RaPID. An RNA of interest is tagged with a BoxB aptamer to recruit a fusion protein of λN and a promiscuous biotin ligase, which can biotinylate associated RBPs. **e**, PL strategies based on MS2 tags and MS2 coat protein (MCP) to capture RBPs associated with an RNA of interest. **f**, dCas13-based PL strategies to biotinylate RBPs associated with an endogenous RNA of interest. **g**, dCas9-based PL strategies to biotinylate DNA-binding proteins at a specific genomic locus. **h**, ChromID. BASU is fused to engineered chromatin readers that can specifically recognize particular chromatin marks, leading to the biotinylation of chromatin-binding proteins.

xenografted cells and performed biotin labeling over the course of 2 days before proteomic analysis[56]. In a subsequent mouse study, Uezu et al. used BioID to map the inhibitory postsynaptic density over a course of 7 days of biotin labeling before proteomic analysis[55]. These long labeling times were likely required for generating sufficient biotinylated material for mass spectrometry because of the low activity of BioID. The application of the more active TurboID or miniTurbo enzymes in future in vivo PL studies should offer increased temporal resolution for mapping dynamic processes in live organisms.

Proximity labeling for profiling protein–RNA interactions

Interactions between proteins and RNA are critical for a wide range of cellular functions, from transcription and translation to innate immunity and stress response[2], and many approaches have been developed to study these interactions[88]. Existing methods can be broadly classified as protein-centric or RNA-centric (Table 3). In protein-centric methods, the RNA interaction partners of a specific protein bait of interest can be identified by RNA sequencing. In RNA-centric methods, the protein partners of a specific RNA bait are identified[88]. There are many more protein-centric methods for mapping protein–RNA interactions owing to the availability of antibodies for protein pulldown and the ease of RNA sequencing.

In protein-centric methods, the addition of a chemical or ultraviolet (UV) crosslinking step before protein bait immunoprecipitation (CLIP) improves the efficiency of RNA capture. CLIP-seq and related methods have been widely applied to the detection of RNAs associated with a particular protein, and generally these methods are highly specific and can be carried out without the exogenous expression of any components[89–94]. However, existing approaches are limited by antibody quality, and UV crosslinking has low efficiency. Furthermore, these methods query protein–RNA interactions across the entire cell, while there may exist compartment-specific variability; for instance, a specific protein bait may localize to both the nucleus and cytosol and interact with different RNA partners in each location[95].

Table 3 | Comparison of protein–nucleic acid approaches

	Method	Description	Pros	Cons	Refs.
RNA	**RNA-centric (identifying the proteins interacting with an RNA of interest)**				
	RAP-MS, PAIR, TRIP, CHART, ChIRP	UV or FA crosslinking followed by RNA pulldown using biotinylated nucleic acid probes.	No genetic engineering or exogenous expression of components. UV crosslinking is highly specific for direct protein–RNA interactions.	Crosslinking—UV in particular—has low efficiency, requiring 10^8–10^9 cells. FA is less specific and results in protein–protein crosslinks that increase background. DNA probes must be optimized and can lead to non-specific capture of RNAs or contribute to lower efficiency.	102–106
	MS2-Biotrap	UV crosslinking followed by RNA pulldown via MS2–coat protein interaction.	Improves the pulldown workflow by avoiding ASO capture.	Crosslinking is low efficiency, requiring 10^8–10^9 cells. Exogenous expression of MS2-tagged RNA may not recapitulate physiological concentrations or conditions	107
	RaPID	MS2-modified endogenous RNA recruits a coat protein–PL enzyme fusion to biotinylate proteins interacting with the RNA of interest.	Avoids crosslinking and associated problems. Enables direct biotinylation of interacting proteins. Can be applied in vivo.	Biotinylation of the general location necessitates spatial references (for example, scrambled RNA control) to eliminate false positives. PL captures indirect interactors. Exogenous expression of MS2-target RNA may not recapitulate physiological concentrations or conditions. Biotinylated proteins may be proximal to the MS2 site and not the RNA in general, making this method better for shorter RNAs.	14,108,109
	CRUIS,CBRPP, CARPID, dCas13d-dsRBD-APEX2	PL enzyme (PafA, BioID, BASU, APEX2) fusion to catalytically inactive dCas13 to biotinylate proteins interacting with an endogenous transcript.	Enables direct biotinylation of proteins interacting with endogenous RNA transcripts. In vivo compatible and can be easily engineered for different targets. Avoids crosslinking.	Incomplete localization of Cas13 can produce high background. May require guide optimization, as well as spatial references (non-targeting guide) to account for non-specific labeling. PL captures indirect interactors. Biotinylated proteins are proximal to the guide RNA site and not to the entire target RNA in general.	109,112
	Protein-centric (identifying the RNAs interacting with a protein of interest)				
	CLIP-seq, eCLIP, iCLIP, irCLIP, PAR-CLIP, fCLIP	Crosslinking immunoprecipitation. There are many variations of the CLIP-seq protocol, but generally, crosslinking of proteins to RNA is carried out by UV (CLIP-seq), by UV using incorporated thiouridine (PAR-CLIP), or using FA (fCLIP). A protein of interest is isolated by antibody pulldown and the covalently bound RNA is sequenced.	UV crosslinking is highly specific. Does not require genetic engineering or exogenous expression of components.	Can be difficult to obtain enough crosslinked RNA because of low efficiency of crosslinking, poor antibody pulldown or low abundance of the RBP–RNA complex. Requires IP-grade antibodies.	89–94
	RIP-seq	Antibody pulldown of a protein of interest under non-denaturing conditions to recover the associated RNAs.	Higher RNA yield than CLIP. Simple protocol without genetic engineering or exogenous expression.	Lower SNR than CLIP; may capture indirect interactors and has a higher chance of false positives.	128
	RNA tagging, TRIBE	RNA tagging uses a poly(U) polymerase fused with the POI to extend poly uracil at the 3′ end of proximal RNAs, which can be subsequently enriched using poly(A) ASO capture. TRIBE uses ADAR fused with the POI and mediates A-to-I editing of interacting RNAs, which can then be identified by sequencing.	Does not require antibody purification. Does not require crosslinking.	Exogenous expression of RBPs can lead to false positives or negatives. RNA tagging may be biased toward 3′ interactors.	129,130
	APEX-RIP, Proximity-Clip	Proteins are biotinylated by APEX2 labeling, and RNA and proteins are crosslinked by UV and 4SU (proximity CLIP) or FA (APEX-RIP). Streptavidin pulldown enables the enrichment of RNA of a specific subcellular location.	Does not rely on antibody purification. Can recover organelle- or location-specific RNAs. UV crosslinking captures direct interactors.	FA crosslinking results in poor specificity, which can be overcome by UV crosslinking at the expense of efficiency. Adapting this method to RBP-specific capture necessitates IP-grade antibodies or genetic tagging of RBP of interest.	96,97
	APEX-seq, CAP-seq	PL of RNAs directly by a PL enzyme enables the enrichment of RNA that interacts with a POI or is located in specific subcellular locations.	Direct labeling of RNA improves workflow, specificity and efficiency. Can be performed in vivo.	PL can capture indirect interactors. Adapting these techniques to studying specific RBPs requires exogenous expression of the RBP–PL fusion protein.	98–101

Continued

Table 3 | Comparison of protein–nucleic acid approaches (continued)

	Method	Description	Pros	Cons	Refs.
DNA	**Protein-centric (identifying the DNAs associated with a protein of interest)**				
	ChIP-seq	Chromatin immunoprecipitation. Antibody pulldown of a POI under non-denaturing conditions allows the identification of associated DNA fragments.	Widely adopted and straightforward protocol, relatively unbiased, does not require exogenous expression.	Requires IP-grade antibodies.	6
	ALaP	APEX2 is fused to a protein of interest to detect associated DNA.	Does not require antibody pulldown.	Requires a spatial reference to improve SNR. Exogenous expression of fusion protein may not reflect physiological conditions.	114
	Chromatin-modification-centric (identifying proteins associated with a specific chromatin modification)				
	ChromID	Fusion of BASU promiscuous biotin ligase to 'reader domains' that specifically bind to chromatin modifications (for example, H3K4me3), which enables the identification of proteins associated with specific chromatin modifications.	Direct labeling of proteins associated with a specific chromatin modification.	Overexpression of the reader domains may perturb the normal occupancy of chromatin modifications. PL may require a spatial reference to improve SNR.	121
	DNA-centric (identifying proteins associated with a specific DNA sequence)				
	RIME, ChIP-MS	DNA–protein crosslinking followed by immunoprecipitation.	Enables the assessment of chromatin-bound protein complexes.	Crosslinking has low efficiency and may result in false positives.	119,120
	APEX–DBP fusion	Fusion of a PL enzyme to a DBP enables the labeling of proteins associated with the DNA-binding site of the DBP.	Does not require crosslinking or antibody pulldown. Can be performed in vivo.	May require a spatial reference to improve SNR. Exogenous expression of a DBP can perturb the studied system.	26
	CASPEX, C-BERST	APEX2-dCas9 fusion proteins are expressed in a cell along with targeting guides to enable labeling of proteins associated with a specific DNA sequence.	Easily reprogrammed and simple protocol. Can directly enrich proteins and avoids crosslinking or IP.	May requires a spatial reference (for example, non-targeting guide) to improve SNR. Exogenous expression of Cas9 can perturb the studied system.	116,117

4SU, 4-thiouridine; ASO, antisense oligonucleotide; DBP, DNA-binding protein; FA, formaldehyde; IP, immunoprecipitation; MS, mass spectrometry; PL, proximity labeling; RBP, RNA binding protein; SNR, signal-to-noise ratio; POI, protein of interest.

PL has been combined with protein–RNA crosslinking to discover RNAs proximal to protein baits in specific subcellular locales. APEX-RIP[96] uses formaldehyde, whereas Proximity-CLIP[97] uses UV, to crosslink APEX-biotinylated proteins to RNA just before cell lysis, enabling streptavidin-based enrichment of protein–RNA complexes (Fig. 2a). The methods have been applied to the endoplasmic reticulum membrane[96], nuclear lamina[96] and cell–cell interfaces[97]. Using Proximity-CLIP, Benhalevy et al. observed the enrichment of CUG repeats in the RNA binding protein (RBP)-protected footprints of mRNA 3′ UTRs localized to cell–cell interfaces, among other functional insights into protein–RNA occupancy[97].

In APEX-RIP, the use of formaldehyde adds time and complexity, and degrades spatial specificity. In a more direct approach, APEX-seq bypasses the need for protein–RNA crosslinking altogether, using an APEX fusion protein to directly biotinylate proximal endogenous RNAs (Fig. 2b)[98,99]. After 1 min labeling in live cells, streptavidin is used to enrich tagged RNAs for RNA-seq. An improved variation of APEX-seq uses a more efficient substrate, biotin-aniline, which improves RNA capture efficiency[100]. Fazal et al. used APEX-seq to generate a transcriptome-wide subcellular RNA atlas in human fibroblasts, uncovering functional insights and correlating transcript location with genome architecture and protein localization[98]. By taking advantage of APEX's rapid kinetics, the authors used APEX-seq to quantify RNA dynamics at the outer mitochondrial membrane in response to drug perturbations and identified two distinct pathways for mRNA localization to the outer mitochondrial membrane[98]. APEX-seq has also been used to study stress granules, providing insights into the organization of translation initiation complexes on active mRNAs[99].

An alternative protein-centric PL method, chromophore-assisted proximity labeling and sequencing (CAP-seq), incorporates the light-activated singlet oxygen generator miniSOG for proximity-dependent photo-oxidation of RNA nucleobases, which can be subsequently captured by amine probes and identified by high-throughput sequencing[101] (Fig. 2c). Although the temporal resolution of CAP-seq (~20 min) is lower than that of APEX-seq (1 min), the two approaches offer distinct mechanisms of RNA labeling and may be complementary. As compared to traditional protein-centric sequencing methods, PL-based APEX-seq and CAP-seq do not require antibodies or crosslinking steps and can be easily adapted to identify the interacting or proximal RNAs of specific RBPs.

In contrast to protein-centric methods, RNA-centric methods target an RNA of interest to identify its protein binding partners. Traditionally, these approaches involve crosslinking and RNA capture using biotinylated oligonucleotide probes or MS2 bacteriophage tags[102–107]. However, the development of RNA-centric PL offers an alternative that does not require crosslinking. RaPID (RNA–protein interaction detection) allows the biotinylation of RBPs by tagging an RNA of interest with a BoxB aptamer to recruit a fusion protein of the bacteriophage λ N peptide and the biotin ligase BASU[14] (Fig. 2d). RaPID was used to discover host proteins that interact with Zika virus RNA[14]. In similar approaches, the MS2 coat protein has been fused to BioID[108] and to APEX2[109] to recruit these PL enzymes to MS2-tagged RNAs (Fig. 2e). However, these methods map proteins that interact with exogenously expressed tagged RNA, which may not accurately reflect the interactome of native transcripts.

The development of RNA-directed CRISPR systems offers the opportunity to target endogenous RNAs. For example, Han et al. targeted catalytically inactive RfxCas13d fused with APEX2 and a double-stranded RNA-binding domain (dsRBD) (to enhance its binding affinity) to human telomerase RNA[109] (Fig. 2f). Using this

approach, the authors discovered a previously unknown interaction between human telomerase RNA and the N^6-methyladenosine (m6A) demethylase ALKBH5, and subsequent studies showed that post-transcriptional regulation by ALKBH5 affects both telomerase complex assembly and activity. Alternative methods have been developed that combine inactive dCas13 orthologs with BioID2[110], APEX2[111], PUP-IT[112] or BASU[113] labeling for RBP profiling. These approaches vary in their benefits and drawbacks; for example, different dCas13 orthologs may exhibit differential binding to the accessible regions of the target RNA, and the chosen PL enzyme will have corresponding benefits and limitations, as previously discussed (Table 1). A potential limitation of these approaches is the large size of Cas13, which may sterically interfere with RBP binding; alternative strategies to target PL enzymes to specific RNAs may further improve RNA-centric discovery.

Proximity labeling for profiling protein–DNA interactions

Protein–DNA interactions are vital to the regulation of gene expression, genome integrity and chromatin organization. ChIP-seq is widely used to capture and sequence DNA regions associated with a protein of interest[6]. The PL adaptation of this approach occurs in living cells and uses the peroxidase APEX to biotinylate proteins proximal to a bait, which are in turn crosslinked by formaldehyde to neighboring DNA regions. Subsequently, biotinylated protein–DNA fragment complexes are enriched by streptavidin and analyzed by next-generation sequencing. ALaP (for APEX-mediated chromatin labeling and purification) is conceptually analogous to APEX-RIP for RNA identification[114] and offers improved sensitivity but decreased specificity in comparison to traditional ChIP-seq. ALaP has also been further adapted for mapping the genomic contact sites of promyelocytic leukemia bodies, phase-separated nuclear structures that physically interact with chromatin[115].

For DNA-centric mapping, wherein proteins proximal to a genomic locus or chromatin complex of interest are identified in an unbiased manner, several methods have been developed. Three groups independently combined PL with CRISPR-based genome targeting[116–118]. Fusing APEX2 with catalytically inactive dCas9 to target specific genomic loci (for example, telomeres and centromeres) allowed associated proteins to be biotinylated, enriched, and analyzed by mass spectrometry[116,117] (Fig. 2g). For discovery of proteins associated with specific chromatin complexes, RIME (rapid immunoprecipitation mass spectrometry of endogenous proteins)[119] and ChIP-MS[120] were reported. More recently, ChromID was used to interrogate protein interactomes at specific chromatin marks by fusing BASU to engineered readers specific to chromatin modifications[121] (Fig. 2h). ChromID identified promoter regions modified by trimethylation on histone H3 Lys4 and Lys27[121]. Although the presence of targeting enzymes may affect the interactors that bind to chromatin, these studies provide a tool to investigate the regulatory mechanisms of chromatin functions. Of note, APEX-based PL has also been applied for mapping proteins associated with mitochondrial DNA, uncovering seven previously unknown mitochondrial nucleoid-associated proteins[26].

Limitations of proximity labeling

Molecular interaction mapping with PL-based approaches requires direct fusion of a PL enzyme (27–44 KDa) to the protein of interest, requiring either transfection of the fusion construct or an alternative induction method, such as viral infection. The fusion can potentially affect the function, localization or even interactome of the target. Thus, it is crucial that functional and localization assays are performed to confirm that the PL enzyme fusion construct remains physiologically relevant and behaves similarly to the endogenous protein of interest. Furthermore, the selection of PL enzyme depends highly on the specific application, as each enzyme has its own advantages and disadvantages (Table 1). For example, the requirement for hydrogen peroxide in APEX labeling may compromise redox-sensitive proteins or pathways and hinder in vivo applications, whereas biotin ligases, such as BioID or TurboID, are less toxic and more suitable in these scenarios.

Because some published PL datasets do not derive from quantitative approaches to data collection and analyses, these datasets may be considered candidate lists that may contain considerable false positives. However, PL experiments can produce highly specific datasets if quantitative mass spectrometry is used while including proper controls for ratiometric or statistical analysis[18,122]. For example, we have previously used APEX2 to generate a highly specific proteome of the outer mitochondrial membrane (OMM) by comparing the extent of biotinylation of proteins by APEX2 targeted to the OMM versus APEX2 expressed in the cytosol[23]. However, PL-based technology may have decreased sensitivity for various reasons. For instance, in the example described above, the ratiometric analysis filters out dual-localized proteins—proteins that reside both in the cytosol and on the OMM. Furthermore, proteins that lack surface-exposed tyrosines (in the case of APEX) or lysines (in the case of BioID and TurboID) may not be detected, and different PL enzymes may exhibit biases toward labeling certain protein substrates[123]. More details regarding setting up, optimizing, analyzing and troubleshooting PL experiments may be found in two protocol publications from our laboratory[18,122].

Conclusions and outlook

Technological advances in molecular interaction mapping using PL have enabled biological investigations previously difficult to access. However, further tool development and engineering may allow more comprehensive interactome maps and improve spatiotemporal specificity in a greater diversity of model systems. While biotin ligases such as BioID and now TurboID have been successfully used in many organisms for in vivo proteomic mapping, further optimization such as the use of non-biotin probes to avoid background from endogenously biotinylated proteins may improve compatibility for PL in vivo. For protein–nucleic acid mapping, improving the efficiency of RNA or DNA labeling by PL enzymes will boost sensitivity and analysis of transcriptomes and genomes in distinct cell populations. Furthermore, improvements in CRISPR-based nucleic acid targeting and binding stability should improve PL approaches that use this mechanism and enable application to endogenous transcripts expressed at low levels. Multiplexing PL enzymes and enrichment strategies could allow simultaneous molecular interactome mapping for multiple complexes at a time. While PL has enabled molecular interaction mapping in many previously intractable biological systems (for example, transient interactions, insoluble baits, in vivo interactions, etc.), continuing development of increasingly sophisticated PL technology may vastly expand the range of PL-based discoveries and address more challenging questions, such as determining the affinity, stoichiometry and contact sites of molecular interactions.

Received: 23 June 2020; Accepted: 30 October 2020;
Published online: 11 January 2021

References

1. Keskin, O., Tuncbag, N. & Gursoy, A. Predicting protein-protein interactions from the molecular to the proteome level. *Chem. Rev.* **116**, 4884–4909 (2016).
2. Hentze, M. W., Castello, A., Schwarzl, T. & Preiss, T. A brave new world of RNA-binding proteins. *Nat. Rev. Mol. Cell Biol.* **19**, 327–341 (2018).
3. Brückner, A., Polge, C., Lentze, N., Auerbach, D. & Schlattner, U. Yeast two-hybrid, a powerful tool for systems biology. *Int. J. Mol. Sci.* **10**, 2763–2788 (2009).
4. Dunham, W. H., Mullin, M. & Gingras, A. C. Affinity-purification coupled to mass spectrometry: basic principles and strategies. *Proteomics* **12**, 1576–1590 (2012).

5. Visa, N. & Jordán-Pla, A. ChIP and ChIP-related techniques: expanding the fields of application and improving ChIP performance. *Methods Mol. Biol.* **1689**, 1–7 (2018).
6. Park, P. J. ChIP-seq: advantages and challenges of a maturing technology. *Nat. Rev. Genet.* **10**, 669–680 (2009).
7. Liu, F., Rijkers, D. T. S., Post, H. & Heck, A. J. R. Proteome-wide profiling of protein assemblies by cross-linking mass spectrometry. *Nat. Methods* **12**, 1179–1184 (2015).
8. Trigg, S. A. et al. CrY2H-seq: a massively multiplexed assay for deep-coverage interactome mapping. *Nat. Methods* **14**, 819–825 (2017).
9. Lam, S. S. et al. Directed evolution of APEX2 for electron microscopy and proximity labeling. *Nat. Methods* **12**, 51–54 (2015).
This study used directed evolution to develop APEX2.
10. Kotani, N. et al. Biochemical visualization of cell surface molecular clustering in living cells. *Proc. Natl Acad. Sci. USA* **105**, 7405–7409 (2008).
11. Roux, K. J., Kim, D. I., Raida, M. & Burke, B. A promiscuous biotin ligase fusion protein identifies proximal and interacting proteins in mammalian cells. *J. Cell Biol.* **196**, 801–810 (2012).
This study introduced a variant of BioID for promiscuous PL.
12. Choi-Rhee, E., Schulman, H. & Cronan, J. E. Promiscuous protein biotinylation by Escherichia coli biotin protein ligase. *Protein Sci.* **13**, 3043–3050 (2004).
13. Kim, D. I. et al. An improved smaller biotin ligase for BioID proximity labeling. *Mol. Biol. Cell* **27**, 1188–1196 (2016).
14. Ramanathan, M. et al. RNA-protein interaction detection in living cells. *Nat. Methods* **15**, 207–212 (2018).
This PL-based study reported a means to identify RBPs associated with an RNA of interest.
15. Branon, T. C. et al. Efficient proximity labeling in living cells and organisms with TurboID. *Nat. Biotechnol.* **36**, 880–887 (2018).
This study used directed evolution to engineer TurboID from BioID.
16. Martell, J. D. et al. Engineered ascorbate peroxidase as a genetically encoded reporter for electron microscopy. *Nat. Biotechnol.* **30**, 1143–1148 (2012).
17. Kim, D. I. et al. Probing nuclear pore complex architecture with proximity-dependent biotinylation. *Proc. Natl Acad. Sci. USA* **111**, E2453–E2461 (2014).
18. Hung, V. et al. Spatially resolved proteomic mapping in living cells with the engineered peroxidase APEX2. *Nat. Protoc.* **11**, 456–475 (2016).
19. Rhee, H. W. et al. Proteomic mapping of mitochondria in living cells via spatially restricted enzymatic tagging. *Science* **339**, 1328–1331 (2013).
20. Martell, J. D. et al. A split horseradish peroxidase for the detection of intercellular protein-protein interactions and sensitive visualization of synapses. *Nat. Biotechnol.* **34**, 774–780 (2016).
21. Markmiller, S. et al. Context-dependent and disease-specific diversity in protein interactions within stress granules. *Cell* **172**, 590–604.e13 (2018).
22. Loh, K. H. et al. Proteomic analysis of unbounded cellular compartments: synaptic clefts. *Cell* **166**, 1295–1307.e21 (2016).
23. Hung, V. et al. Proteomic mapping of cytosol-facing outer mitochondrial and ER membranes in living human cells by proximity biotinylation. *eLife* **6**, e24463 (2017).
24. Youn, J. Y. et al. High-density proximity mapping reveals the subcellular organization of mRNA-associated granules and bodies. *Mol. Cell* **69**, 517–532.e11 (2018).
25. Kehrer, J., Frischknecht, F. & Mair, G. R. Proteomic analysis of the *Plasmodium berghei* gametocyte egressome and vesicular bioid of osmiophilic body proteins identifies merozoite trap-like protein (MTRAP) as an essential factor for parasite transmission. *Mol. Cell. Proteomics* **15**, 2852–2862 (2016).
26. Han, S. et al. Proximity biotinylation as a method for mapping proteins associated with mtDNA in living cells. Cell. *Chem. Biol.* **24**, 404–414 (2017).
27. Miyagawa-Yamaguchi, A., Kotani, N. & Honke, K. Expressed glycosylphosphatidylinositol-anchored horseradish peroxidase identifies co-clustering molecules in individual lipid raft domains. *PLoS One* **9**, e93054 (2014).
28. Gingras, A. C., Abe, K. T. & Raught, B. Getting to know the neighborhood: using proximity-dependent biotinylation to characterize protein complexes and map organelles. *Curr. Opin. Chem. Biol.* **48**, 44–54 (2019).
29. Gullberg, M. et al. Cytokine detection by antibody-based proximity ligation. *Proc. Natl Acad. Sci. USA* **101**, 8420–8424 (2004).
30. Fredriksson, S. et al. Protein detection using proximity-dependent DNA ligation assays. *Nat. Biotechnol.* **20**, 473–477 (2002).
31. Yang, Y. et al. Genetically encoded protein photocrosslinker with a transferable mass spectrometry-identifiable label. *Nat. Commun.* **7**, 12299 (2016).
32. Liu, Q. et al. A proximity-tagging system to identify membrane protein-protein interactions. *Nat. Methods* **15**, 715–722 (2018).
33. Geri, J. B. et al. Microenvironment mapping via Dexter energy transfer on immune cells. *Science* **367**, 1091–1097 (2020).
34. Dumont, A.A., Dumont, L., Berthiaume, J. & Auger-Messier, M. p38α MAPK proximity assay reveals a regulatory mechanism of alternative splicing in cardiomyocytes. *Biochim. Biophys. Acta Mol. Cell Res.* **1866**, 118557 (2019).
35. Prikas, E., Poljak, A. & Ittner, A. Mapping p38α mitogen-activated protein kinase signaling by proximity-dependent labeling. *Protein Sci.* **29**, 1196–1210 (2020).
36. Couzens, A. L. et al. Protein interaction network of the mammalian Hippo pathway reveals mechanisms of kinase-phosphatase interactions. *Sci. Signal.* **6**, rs15 (2013).
This study fused BioID to 19 members of the Hippo pathway for interactome mapping.
37. Liu, G. et al. Mechanism of adrenergic $Ca_V1.2$ stimulation revealed by proximity proteomics. *Nature* **577**, 695–700 (2020).
38. Paek, J. et al. Multidimensional tracking of GPCR signaling via peroxidase-catalyzed proximity labeling. *Cell* **169**, 338–349.e11 (2017).
This study utilized APEX to map the dynamic interactomes of the GPCRs AT1R and β2AR in response to agonist activation.
39. Lobingier, B. T. et al. An approach to spatiotemporally resolve protein interaction networks in living cells. *Cell* **169**, 350–360.e12 (2017).
This study utilized APEX to map the dynamic interactomes of the δ-opioid receptor, a GPCR, in response to agonist activation.
40. Coyaud, E. et al. BioID-based identification of skp cullin F-box (SCF) β-TrCP1/2 E3 ligase substrates. *Mol. Cell. Proteomics* **14**, 1781–1795 (2015).
41. Mannix, K. M., Starble, R. M., Kaufman, R. S. & Cooley, L. Proximity labeling reveals novel interactomes in live *Drosophila* tissue. *Development* **146**, dev176644 (2019).
42. Cutler, J. A. et al. Differential signaling through p190 and p210 BCR-ABL fusion proteins revealed by interactome and phosphoproteome analysis. *Leukemia* **31**, 1513–1524 (2017).
43. Rodríguez-Fraticelli, A. E. et al. Developmental regulation of apical endocytosis controls epithelial patterning in vertebrate tubular organs. *Nat. Cell Biol.* **17**, 241–250 (2015).
44. Holthenrich, A., Drexler, H. C. A., Chehab, T., Naß, J. & Gerke, V. Proximity proteomics of endothelial Weibel-Palade bodies identifies novel regulator of von Willebrand factor secretion. *Blood* **134**, 979–982 (2019).
45. Liao, Y. C. et al. RNA granules hitchhike on lysosomes for long-distance transport, using annexin a11 as a molecular tether. *Cell* **179**, 147–164.e20 (2019).
46. Chung, C. Y. et al. In situ peroxidase labeling and mass-spectrometry connects alpha-synuclein directly to endocytic trafficking and mRNA metabolism in neurons. *Cell Syst.* **4**, 242–250.e4 (2017).
47. Santin, Y. G. et al. In vivo TssA proximity labelling during type VI secretion biogenesis reveals TagA as a protein that stops and holds the sheath. *Nat. Microbiol.* **3**, 1304–1313 (2018).
48. Opitz, N. et al. Capturing the Asc1p/Receptor for Activated C Kinase 1 (RACK1) microenvironment at the head region of the 40s ribosome with quantitative BioID in yeast. *Mol. Cell. Proteomics* **16**, 2199–2218 (2017).
49. Larochelle, M., Bergeron, D., Arcand, B. & Bachand, F. Proximity-dependent biotinylation mediated by TurboID to identify protein-protein interaction networks in yeast. *J. Cell Sci.* **132**, jcs232249 (2019).
50. Li, J. et al. Cell-surface proteomic profiling in the fly brain uncovers wiring regulators. *Cell* **180**, 373–386.e15 (2020).
51. Domsch, K. et al. The Hox transcription factor Ubx stabilizes lineage commitment by suppressing cellular plasticity in *Drosophila*. *eLife* **8**, e42675 (2019).
52. Khan, M., Youn, J. Y., Gingras, A. C., Subramaniam, R. & Desveaux, D. In planta proximity dependent biotin identification (BioID). *Sci. Rep.* **8**, 9212 (2018).
53. Mair, A., Xu, S. L., Branon, T. C., Ting, A. Y. & Bergmann, D. C. Proximity labeling of protein complexes and cell-type-specific organellar proteomes in *Arabidopsis* enabled by TurboID. *eLife* **8**, e47864 (2019).
54. Feng, W. et al. Identifying the cardiac dyad proteome in vivo by a BioID2 knock-in strategy. *Circulation* **141**, 940–942 (2020).
55. Uezu, A. et al. Identification of an elaborate complex mediating postsynaptic inhibition. *Science* **353**, 1123–1129 (2016).
56. Dingar, D. et al. BioID identifies novel c-MYC interacting partners in cultured cells and xenograft tumors. *J. Proteomics* **118**, 95–111 (2015).
This study reported the use of BioID in vivo in mice.
57. Bar, D. Z. et al. Biotinylation by antibody recognition—a method for proximity labeling. *Nat. Methods* **15**, 127–133 (2018).
58. Birendra, Kc. et al. VRK2A is an A-type lamin-dependent nuclear envelope kinase that phosphorylates BAF. *Mol. Biol. Cell* **28**, 2241–2250 (2017).
59. Fu, Y. et al. MacroH2A1 associates with nuclear lamina and maintains chromatin architecture in mouse liver cells. *Sci. Rep.* **5**, 17186 (2015).
60. Mackmull, M. T. et al. Landscape of nuclear transport receptor cargo specificity. *Mol. Syst. Biol.* **13**, 962 (2017).
61. Chou, C. C. et al. TDP-43 pathology disrupts nuclear pore complexes and nucleocytoplasmic transport in ALS/FTD. *Nat. Neurosci.* **21**, 228–239 (2018).
62. Phelan, J. D. et al. A multiprotein supercomplex controlling oncogenic signalling in lymphoma. *Nature* **560**, 387–391 (2018).

63. Cui, Y. et al. The NF2 tumor suppressor merlin interacts with Ras and RasGAP, which may modulate Ras signaling. *Oncogene* **38**, 6370–6381 (2019).
64. Che, Y. et al. KRAS regulation by small non-coding RNAs and SNARE proteins. *Nat. Commun.* **10**, 5118 (2019).
65. Mirza, A. N. et al. LAP2 proteins chaperone GLI1 movement between the lamina and chromatin to regulate transcription. *Cell* **176**, 198–212.e15 (2019).
66. Park, S. S. et al. Kir2. 1 interactome mapping uncovers PKP4 as a modulator of the Kir2.1-regulated inward rectifier potassium currents. *Mol. Cell. Proteomics* **19**, 1436–1449 (2020).
67. Chang, L. et al. Identification of siglec ligands using a proximity labeling method. *J. Proteome Res.* **16**, 3929–3941 (2017).
68. Wu, G., Nagala, M. & Crocker, P. R. Identification of lectin counter-receptors on cell membranes by proximity labeling. *Glycobiology* **27**, 800–805 (2017).
69. Zheng, N. & Shabek, N. Ubiquitin ligases: structure, function, and regulation. *Annu. Rev. Biochem.* **86**, 129–157 (2017).
70. Grainger, S. et al. EGFR is required for Wnt9a–Fzd9b signalling specificity in haematopoietic stem cells. *Nat. Cell Biol.* **21**, 721–730 (2019).
71. James, C. et al. Proteomic mapping by rapamycin-dependent targeting of APEX2 identifies binding partners of VAPB at the inner nuclear membrane. *J. Biol. Chem.* **294**, 16241–16254 (2019).
72. De Munter, S. et al. Split-BioID: a proximity biotinylation assay for dimerization-dependent protein interactions. *FEBS Lett.* **591**, 415–424 (2017).
73. Schopp, I. M. et al. Split-BioID a conditional proteomics approach to monitor the composition of spatiotemporally defined protein complexes. *Nat. Commun.* **8**, 15690 (2017).
74. Kwak, C. et al. Contact-ID, a new tool for profiling organelle contact site, reveals proteins of mitochondrial-associated membrane formation. *Proc. Natl Acad. Sci. USA* https://doi.org/10.1073/pnas.1916584117 (2020).
75. Han, Y. et al. Directed evolution of split APEX2 peroxidase. *ACS Chem. Biol.* **14**, 619–635 (2019).
76. Cho, K. F. et al. Split-TurboID enables contact-dependent proximity labeling in cells. *Proc. Natl. Acad. Sci. USA* **117**, 12143–12154 (2020).
77. Conlan, B., Stoll, T., Gorman, J. J., Saur, I. & Rathjen, J. P. Development of a rapid in planta bioid system as a probe for plasma membrane-associated immunity proteins. *Front. Plant Sci.* **9**, 1882 (2018).
78. Zhang, Y. et al. TurboID-based proximity labeling reveals that UBR7 is a regulator of N NLR immune receptor-mediated immunity. *Nat. Commun.* **10**, 3252 (2019).
79. Batsios, P., Meyer, I. & Gräf, R. Proximity-Dependent Biotin Identification (BioID) in *Dictyostelium amoebae*. *Methods Enzymol.* **569**, 23–42 (2016).
80. Pitzen, V., Askarzada, S., Gräf, R. & Meyer, I. CDK5RAP2 is an essential scaffolding protein of the corona of the *Dictyostelium* centrosome. *Cells* **7**, 32 (2018).
81. Morriswood, B. et al. Novel bilobe components in *Trypanosoma brucei* identified using proximity-dependent biotinylation. *Eukaryot. Cell* **12**, 356–367 (2013).
82. Chen, A. L. et al. Novel components of the Toxoplasma inner membrane complex revealed by BioID. *MBio* **6**, e02357–e14 (2015).
83. McAllaster, M. R. et al. Proteomic identification of novel cytoskeletal proteins associated with TbPLK, an essential regulator of cell morphogenesis in *Trypanosoma brucei*. *Mol. Biol. Cell* **26**, 3013–3029 (2015).
84. Khosh-Naucke, M. et al. Identification of novel parasitophorous vacuole proteins in *P. falciparum* parasites using BioID. *Int. J. Med. Microbiol.* **308**, 13–24 (2018).
85. Tu, V. et al. The *Toxoplasma gondii* cyst wall interactome. *MBio* **11**, e02699–19 (2020).
86. Shinoda, N., Hanawa, N., Chihara, T., Koto, A. & Miura, M. Dronc-independent basal executioner caspase activity sustains *Drosophila* imaginal tissue growth. *Proc. Natl Acad. Sci. USA* **116**, 20539–20544 (2019).
87. Carnesecchi, J. et al. Multi-level and lineage-specific interactomes of the Hox transcription factor Ubx contribute to its functional specificity. *Nat. Commun.* **11**, 1388 (2020).
88. Ramanathan, M., Porter, D. F. & Khavari, P. A. Methods to study RNA-protein interactions. *Nat. Methods* **16**, 225–234 (2019).
89. Licatalosi, D. D. et al. HITS-CLIP yields genome-wide insights into brain alternative RNA processing. *Nature* **456**, 464–469 (2008).
90. Hafner, M. et al. Transcriptome-wide identification of RNA-binding protein and microRNA target sites by PAR-CLIP. *Cell* **141**, 129–141 (2010).
91. Kim, B. & Kim, V. N. fCLIP-seq for transcriptomic footprinting of dsRNA-binding proteins: Lessons from DROSHA. *Methods* **152**, 3–11 (2019).
92. Van Nostrand, E. L. et al. Robust transcriptome-wide discovery of RNA-binding protein binding sites with enhanced CLIP (eCLIP). *Nat. Methods* **13**, 508–514 (2016).
93. Konig, J. et al. iCLIP – transcriptome-wide mapping of protein-RNA interactions with individual nucleotide resolution. *J. Vis. Exp*. https://doi.org/10.3791/2638 (2011).
94. Zarnegar, B. J. et al. irCLIP platform for efficient characterization of protein-RNA interactions. *Nat. Methods* **13**, 489–492 (2016).
95. Trendel, J. et al. The human RNA-binding proteome and its dynamics during translational arrest. *Cell* **176**, 391–403.e19 (2019).
96. Kaewsapsak, P., Shechner, D. M., Mallard, W., Rinn, J. L. & Ting, A. Y. Live-cell mapping of organelle-associated RNAs via proximity biotinylation combined with protein-RNA crosslinking. *eLife* **6**, e29224 (2017).
97. Benhalevy, D., Anastasakis, D. G. & Hafner, M. Proximity-CLIP provides a snapshot of protein-occupied RNA elements in subcellular compartments. *Nat. Methods* **15**, 1074–1082 (2018).
98. Fazal, F. M. et al. Atlas of subcellular RNA localization revealed by APEX-Seq. *Cell* **178**, 473–490.e26 (2019).
This study used APEX to systematically investigate subcellular localization of RNA.
99. Padrón, A., Iwasaki, S. & Ingolia, N. T. Proximity RNA labeling by APEX-Seq reveals the organization of translation initiation complexes and repressive RNA granules. *Mol. Cell* **75**, 875–887.e5 (2019).
100. Zhou, Y. et al. Expanding APEX2 substrates for proximity-dependent labeling of nucleic acids and proteins in living cells. *Angew. Chem. Int. Ed. Engl.* **58**, 11763–11767 (2019).
101. Wang, P. et al. Mapping spatial transcriptome with light-activated proximity-dependent RNA labeling. *Nat. Chem. Biol.* **15**, 1110–1119 (2019).
102. McHugh, C. A. & Guttman, M. RAP-MS: a method to identify proteins that interact directly with a specific RNA molecule in cells. *Methods Mol. Biol.* **1649**, 473–488 (2018).
103. Zeng, F. et al. A protocol for PAIR: PNA-assisted identification of RNA binding proteins in living cells. *Nat. Protoc.* **1**, 920–927 (2006).
104. Matia-González, A. M., Iadevaia, V. & Gerber, A. P. A versatile tandem RNA isolation procedure to capture in vivo formed mRNA-protein complexes. *Methods* **118-119**, 93–100 (2017).
105. Simon, M. D. et al. The genomic binding sites of a noncoding RNA. *Proc. Natl Acad. Sci. USA* **108**, 20497–20502 (2011).
106. Chu, C., Qu, K., Zhong, F. L., Artandi, S. E. & Chang, H. Y. Genomic maps of long noncoding RNA occupancy reveal principles of RNA-chromatin interactions. *Mol. Cell* **44**, 667–678 (2011).
107. Tsai, B. P., Wang, X., Huang, L. & Waterman, M. L. Quantitative profiling of in vivo-assembled RNA-protein complexes using a novel integrated proteomic approach. *Mol. Cell. Proteomics* **10**, 007385 (2011).
108. Mukherjee, J. et al. β-Actin mRNA interactome mapping by proximity biotinylation. *Proc. Natl Acad. Sci. USA* **116**, 12863–12872 (2019).
109. Han, S. et al. RNA-protein interaction mapping via MS2 or Cas13-based APEX targeting. *Proc. Natl Acad. Sci. USA* **117**, 22068–22079 (2020).
110. Li, Y. et al. CBRPP a new RNA-centric method to study RNA-protein interactions. Preprint at *bioRxiv* https://doi.org/10.1101/2020.04.09.033290 (2020).
111. Lin, X., Fonesca, M. A. S., Corona, R. I. & Lawrenson, K. In vivo discovery of RNA proximal proteins in human cells via proximity-dependent biotinylation. Preprint at *bioRxiv* https://doi.org/10.1101/2020.02.28.970442 (2020).
112. Zhang, Z. et al. Capturing RNA-protein interaction via CRUIS. *Nucleic Acids Res.* **48**, e52 (2020).
113. Yi, W. et al. CRISPR-assisted detection of RNA-protein interactions in living cells. *Nat. Methods* **17**, 685–688 (2020).
114. Kurihara, M. et al. Genomic profiling by ALaP-Seq reveals transcriptional regulation by PML bodies through DNMT3A exclusion. *Mol. Cell* **78**, 493–505.e8 (2020).
115. Lallemand-Breitenbach, V. & de Thé, H. PML nuclear bodies. *Cold Spring Harb. Perspect. Biol* **2**, a000661 (2010).
116. Myers, S. A. et al. Discovery of proteins associated with a predefined genomic locus via dCas9-APEX-mediated proximity labeling. *Nat. Methods* **15**, 437–439 (2018).
117. Gao, X. D. et al. C-BERST: defining subnuclear proteomic landscapes at genomic elements with dCas9-APEX2. *Nat. Methods* **15**, 433–436 (2018).
118. Qiu, W. et al. Determination of local chromatin interactions using a combined CRISPR and peroxidase APEX2 system. *Nucleic Acids Res.* **47**, e52 (2019).
119. Mohammed, H. et al. Rapid immunoprecipitation mass spectrometry of endogenous proteins (RIME) for analysis of chromatin complexes. *Nat. Protoc.* **11**, 316–326 (2016).
120. Wang, C. I. et al. Chromatin proteins captured by ChIP-mass spectrometry are linked to dosage compensation in *Drosophila*. *Nat. Struct. Mol. Biol.* **20**, 202–209 (2013).
121. Villaseñor, R. et al. ChromID identifies the protein interactome at chromatin marks. *Nat. Biotechnol*. https://doi.org/10.1038/s41587-020-0434-2 (2020).
This study used biotin ligase to identify proteins associated with specific chromatin marks.
122. Cho, K. F. et al. Proximity labeling in mammalian cells with TurboID and split-TurboID. *Nat. Protoc*. https://doi.org/10.1038/s41596-020-0399-0 (2020).
123. Minde, D. P., Ramakrishna, M. & Lilley, K. S. Biotin proximity tagging favours unfolded proteins and enables the study of intrinsically disordered regions. *Commun. Biol*. https://doi.org/10.1038/s42003-020-0758-y (2020).

124. Cole, A. et al. Inhibition of the mitochondrial protease ClpP as a therapeutic strategy for human acute myeloid leukemia. *Cancer Cell* **27**, 864–876 (2015).
125. Shin, J. J. H., Gillingham, A. K., Begum, F., Chadwick, J. & Munro, S. TBC1D23 is a bridging factor for endosomal vesicle capture by golgins at the *trans*-Golgi. *Nat. Cell Biol.* **19**, 1424–1432 (2017).
126. Liu, L., Doray, B. & Kornfeld, S. Recycling of Golgi glycosyltransferases requires direct binding to coatomer. *Proc. Natl Acad. Sci. USA* **115**, 8984–8989 (2018).
127. Jia, J. et al. Galectins control mTOR in response to endomembrane damage. *Mol. Cell* **70**, 120–135.e8 (2018).
128. Tenenbaum, S. A., Carson, C. C., Lager, P. J. & Keene, J. D. Identifying mRNA subsets in messenger ribonucleoprotein complexes by using cDNA arrays. *Proc. Natl Acad. Sci. USA* **97**, 14085–14090 (2000).
129. McMahon, A. C. et al. TRIBE: hijacking an RNA-editing enzyme to identify cell-specific targets of RNA-binding proteins. *Cell* **165**, 742–753 (2016).
130. Lapointe, C. P., Wilinski, D., Saunders, H. A. J. & Wickens, M. Protein–RNA networks revealed through covalent RNA marks. *Nat. Methods* **12**, 1163–1170 (2015).

Acknowledgements

This work was supported by the NIH (R01-DK121409 to A.Y.T.) and Stanford Wu Tsai Neurosciences Institute. K.F.C. was supported by NIH Training Grant 2T32CA009302-41 and the Blavatnik Graduate Fellowship. P.E.C. was supported by the NSF Graduate Research Fellowship. A.Y.T. is an investigator of the Chan Zuckerberg Biohub.

Competing interests

The authors declare no competing interests.

Additional information

Correspondence should be addressed to A.Y.T.

Peer review information Rita Strack was the primary editor on this article and managed its editorial process and peer review in collaboration with the rest of the editorial team.