# A proximity proteomics pipeline with improved reproducibility and throughput

Xiaofang Zhong[1,2,3,†], Qiongyu Li[1,2,3,†], Benjamin J. Polacco[1,2,3], Trupti Patil[1,2,3], Aaron Marley[4], Helene Foussard[1,2,3], Prachi Khare[1,2,3], Rasika Vartak[1,2,3], Jiewei Xu[1,2,3], Jeffrey F. DiBerto[5], Bryan L. Roth[5], Manon Eckhardt[1,2,3], Mark Von Zastrow[1,3,4], Nevan J. Krogan[1,2,3], Ruth Hüttenhain[1,2,3,6]

[1]Quantitative Biosciences Institute (QBI), University of California, San Francisco, San Francisco, CA 94158, USA.
[2]J. David Gladstone Institutes, San Francisco, CA 94158, USA.
[3]Department of Cellular and Molecular Pharmacology, University of California, San Francisco, San Francisco, CA 94158, USA.
[4]Department of Psychiatry and Behavioral Sciences, University of California, San Francisco, CA 94158, USA.
[5]Department of Pharmacology, School of Medicine, University of North Carolina at Chapel Hill, Chapel Hill, NC 27599, USA.
[6]Department of Molecular and Cellular Physiology, Stanford University, Stanford, CA 94305, USA.

[†]Authors contributed equally

## Abstract

Proximity labeling (PL) through biotinylation coupled with mass spectrometry (MS) has emerged as a powerful technique for capturing spatial proteomes within living cells. Large-scale sample processing for proximity proteomics requires a workflow that minimizes hands-on time while enhancing quantitative reproducibility. Here, we present a scalable PL pipeline integrating automated enrichment of biotinylated proteins in a 96-well plate format. By combining this pipeline with an optimized quantitative MS acquisition method based on data-independent acquisition (DIA), we not only significantly increased sample throughput but also improved the reproducibility of protein identification and quantification. We applied this pipeline to delineate subcellular proteomes across various cellular compartments, including endosomes, late endosomes/lysosomes, the Golgi apparatus, and the plasma membrane. Moreover, employing $5HT_{2A}$ serotonin receptor as a model, we investigated temporal changes of proximal interaction networks induced by the receptor's activation with serotonin. Finally, to demonstrate the applicability of our PL pipeline across multiple experimental conditions, we further modified the PL pipeline for reduced sample input amounts to accommodate CRISPR-based gene knockout, and assessed the dynamics of the $5HT_{2A}$ network in response to the perturbation of selected proximal interactors. Importantly, the presented PL approach is universally applicable to PL proteomics using biotinylation-based PL enzymes, increasing both throughput and reproducibility of standard protocols.

## Introduction

The subcellular localization of proteins and their interactions with other proteins play a pivotal role in protein function. Both protein localization and protein-protein interactions (PPIs) are dynamic processes, with proteins navigating between subcellular neighborhoods and altering their interaction networks to carry out their functions in response to extracellular signals. Therefore, systematic characterization of protein localization and dynamic protein networks is essential for elucidating the molecular mechanisms underlying fundamental cellular processes.

Proximity labeling (PL) has emerged as a robust and versatile proteomics approach for globally mapping the spatial organization of proteins (Rhee *et al*, 2013; Go *et al*, 2021) and their molecular interactions, including protein-protein (Go *et al*, 2021; Lobingier *et al*, 2017; Paek *et al*, 2017), protein-RNA (Ramanathan *et al*, 2018; Qin *et al*, 2021b), and protein-DNA interactions (Gao *et al*, 2018; Myers *et al*, 2018). PL relies on genetically fusing an enzyme to a target protein or a protein localization domain that subsequently catalyzes the proximity-dependent biotinylation of biomolecules in live cells. Various PL enzymes have been developed representing two main enzyme classes: (1) peroxidases, including horseradish peroxidase (Li *et al*, 2014; Martell *et al*, 2016) and engineered ascorbate peroxidase 2 (APEX2) (Lam *et al*, 2015) and (2) biotin ligases including BioID (Roux *et al*, 2012), UltraID (Kubitz *et al*, 2022), TurboID and miniTurbo (Branon *et al*, 2018), and others (Samavarchi-Tehrani *et al*, 2020; Qin *et al*, 2021a).

The underlying principle of all these PL approaches involves expressing the PL enzyme fusion construct in cells. To mitigate artifacts from overexpression and variability in fusion construct expression levels, most studies aim for moderate and inducible expression in isogenic cell clones (Samavarchi-Tehrani *et al*, 2018) or fusing the PL enzyme to the endogenously expressed proteins (Vandemoortele *et al*, 2019; Gupta *et al*, 2018), requiring the characterization of multiple cellular clones for each PL construct. Once a suitable cellular clone is selected, the biotinylation reaction might be performed under various experimental conditions to label proteins proximal to the protein of interest. Subsequently, biotinylated proteins are enriched for each sample using streptavidin beads, followed by a series of washing steps to remove non-biotinylated proteins and proteolytic digestion to prepare samples for quantitative mass spectrometric (MS) analysis. However, this multi-step process is time-consuming and lacks scalability, thereby limiting sample throughput in current PL proteomics applications. Moreover, sample preparation across multiple batches may introduce variability, compromising reproducibility in quantifying biotinylated proteins across batches. Although PL represents a powerful technology to study dynamics in subcellular localizations as well as protein interactions across diverse conditions, throughput can be a limiting factor due to the multistep nature of the standard protocol (Go *et al*, 2021; Antonicka *et al*, 2020; Göös *et al*, 2022; Piette *et al*, 2021).

To address these issues, we present a scalable and reproducible pipeline for PL-based proteomics (**Figure 1**), incorporating automated enrichment of biotinylated proteins in a 96-well plate format and an optimized proteomics method for reproducible protein quantification. To facilitate PL studies with higher throughput, we developed an automated 96-well format platform for enriching biotinylated proteins using magnetic streptavidin beads on the KingFisher Flex platform, followed by enzymatic digestion and sample clean up in a 96-well format manually.

While PL samples have commonly been analyzed using data-dependent acquisition (DDA)-based MS, the preference for high-abundance peptides in DDA can result in stochastic and irreproducible measurements (Aebersold & Mann, 2003; Michalski *et al*, 2011; Hu *et al*, 2016). In contrast, owing to the systematic and unbiased fragmentation of all precursor ions within a predefined mass range, data-independent acquisition (DIA) achieves higher quantitative consistency at comparable proteome coverage (Gillet *et al*, 2012; Collins *et al*, 2017; Ludwig *et al*, 2018). Therefore, we optimized a DIA-MS method for samples derived from PL experiments based on several parameters, including precursor scan range, normalized collision energy, maximum injection time, and the precursor isolation window overlap. The combination of automated enrichment for biotinylated proteins with the DIA-MS method not only increased throughput but also provided higher reproducibility of protein identification and quantification compared to manual sample processing.

To evaluate our strategy, we selected two common PL applications: mapping of subcellular proteomes and quantitatively analyzing protein interaction network dynamics. First, we used APEX2-based PL to identify compartment-specific proteins for several subcellular compartments including plasma membrane, endosome, Golgi apparatus, and lysosome. Our approach successfully identified compartment-specific proteins aligned with their expected locations based on a comparison with the Human Cell Map (Go *et al*, 2021). Next, we applied our PL strategy to map temporal protein interaction network changes of the $5HT_{2A}$ serotonin receptor, a G protein-coupled receptor (GPCR), following activation with its endogenous ligand serotonin (5-hydroxytryptamine; 5HT). Our PL approach not only achieved high reproducibility in quantifying ligand-induced proximal proteome changes over a time course but also identified well-known regulators of GPCR function, such as G protein-coupled receptor kinase 2 (GRK2), beta-arrestin 2, as well as subunits of protein kinase C and D (PKC and PKD, respectively). Overall, the activity-dependent protein interaction changes for the $5HT_{2A}$ receptor could be clustered into rapid events with transient changes and later events with sustained differences in biotin labeling.

Finally, we assessed the applicability of our PL pipeline for capturing proximal protein network dynamics across multiple experimental conditions. As high-throughput PL studies are not only limited by the enrichment of biotinylated proteins, but also by the scale of the cell culture required to generate sufficient sample input amounts, we initially adapted our PL pipeline to accommodate low sample input amounts. We demonstrated that PL experiments could be performed in a 6-well plate format without compromising performance substantially compared to the high-input PL. To test the potential of our method for investigating interaction dynamics across multiple conditions, we then utilized CRISPR-based gene knockout to perturb selected components of the $5HT_{2A}$ interaction network and to monitor resulting activation-dependent PL dynamics with the low-input PL. Our data revealed that the knockout of both, Annexin A2 (ANXA2) and S100-A10 (S100A10), two components of the $5HT_{2A}$ network, resulted in reduced biotinylation of ANXA2, S100A10, and AHNAK, which is consistent with previous findings demonstrating that these proteins physically interact (Benaud *et al*, 2004; Chen *et al*, 2022).

In summary, we described a versatile PL pipeline applicable across a wide range of PL applications, various experimental scales, and multiple experimental conditions. Although we

focus on APEX2-based PL in our study, the automated enrichment of biotinylated proteins can be widely utilized to accelerate essentially any PL strategy reliant on biotinylation.

## Results

### *Generating monoclonal cell lines for proximity labeling*

To set up a PL proteomics experiment, the PL enzyme is typically genetically fused to a protein localization domain (PLD) or a protein of interest (POI) and expressed in cells. Here, we present considerations for designing PL constructs and outline the process of generating and characterizing monoclonal cell lines tailored for PL proteomics. To generate cell lines with stable and inducible expression of the PL constructs, we introduced APEX2 constructs into a lentiviral backbone under a Doxycycline-inducible promoter containing Puromycin resistance for antibiotic selection (**Figure 2A**). Additionally, the constructs were appended with a FLAG epitope and/or GFP for detection via fluorescence-activated cell sorting (FACS), imaging, and western blot analysis. For spatial proteomics, a PLD was selected to direct the PL construct to the subcellular compartment of interest. For protein interaction mapping, the APEX2 enzyme was fused to the POI and separated by a flexible linker to minimize interference of the PL enzyme with the protein's function (**Figure 2A**). After generating polyclonal cells by lentiviral transduction and puromycin selection, the expression of PL constructs was validated using FACS and immunofluorescence microscopy. To achieve consistent and homogeneous expression of the APEX2-tagged constructs and to mitigate shifts in the cellular populations between biological replicates, monoclonal cell lines were generated by single cell sorting. To identify optimal clones after sorting, multiple clones were initially screened using flow cytometry to assess uniformity of the fusion construct expression. Promising clones were further evaluated by confocal immunofluorescence imaging to validate subcellular localization of the fusion construct, followed by western blot analysis to evaluate their biotinylation capacity (**Figure 2A**).

As an illustration, we detailed the generation and characterization of monoclonal cell lines in HEK293T cells for a plasma membrane (PM) localized APEX2 construct, for which the enzyme was directed to the PM using a Lyn11 motif appended with GFP for visualization (**Appendix Figure S1, Dataset EV1**). Following lentiviral transduction and puromycin selection, the polyclonal cell line exhibited heterogeneous expression of the PL construct, with varied GFP expression patterns observed, such as GFP expression in cytoplasm and nucleus in addition to the PM (**Figure 2B**). Following single cell sorting, 24 monoclonal PM-APEX2 cell lines were subjected to flow cytometry analysis to assess expression levels and consistency of the APEX2-tagged PM construct based on GFP fluorescence. Monoclonal cell lines containing over 85% GFP-positive cells were prioritized for further characterization by confocal immunofluorescence microscopy (**Figure 2C**). Confocal imaging revealed distinct localization patterns for different PM-APEX2 clones, with clones 7, 8, and 11 exhibiting concentration at the PM, while clone 1 displayed mainly diffuse cytoplasmic distribution (**Figure 2D**). Finally, western blot analysis was performed to evaluate the biotinylation efficiency of the monoclonal PM-APEX2 cell lines. Biotinylation was initiated by addition of biotin-phenol and $H_2O_2$, with the sole biotin-phenol treatment as the negative control. All PM-APEX2 clones demonstrated varying levels of biotin labeling (**Figure 2E**). Considering unspecific labeling associated with higher biotinylation

efficiency, clones 8 and 11, exhibiting moderate biotinylation efficiency, were selected for PL studies.

Following this workflow, monoclonal cell lines could be generated and characterized within a timeframe of three weeks. Although APEX2-based PL was employed in our study, these experimental procedures are adaptable to any other PL enzyme.

### *A scalable biotin enrichment method coupled to quantitative MS*

The enrichment of biotinylated proteins represents a key step in PL workflows. However, manual enrichment using streptavidin beads, followed by multiple washing steps to remove non-biotinylated proteins, are not only time-consuming but can also introduce variability between samples, thereby compromising quantitative reproducibility. To address these challenges and facilitate large-scale PL studies while maintaining high reproducibility and quantitative accuracy, we devised a PL strategy combining automated enrichment of biotinylated proteins in 96-well plate format with MS analysis using data-independent acquisition (DIA) (**Figure 3A**). We developed an automated enrichment protocol using magnetic streptavidin beads to bind biotinylated proteins, adapted for use on the KingFisher Flex system (**Figure 3B**). The protocol, outlined in **Dataset EV2**, requires less than 5 minutes to load the corresponding plates into the system.

To maximize the enrichment efficiency, we evaluated the impact of varying protein input amounts. Using the PM-APEX2 cell line, we compared protein input amounts ranging from 0.5 to 2 mg for the enrichment. Our results indicated minor differences comparing 0.5, 1, and 2 mg in the number of peptides and proteins identified, with 1 mg yielding slightly better performance (**Appendix Figure S2A**), suggesting the protocol's applicability across a range of protein input amounts. Additionally, we tested different volumes of streptavidin beads (60, 80, and 100 µl) while maintaining a constant protein input amount of 1 mg. To evaluate nonspecific protein binding to beads, we performed the APEX experiment for the PM-APEX cell line in the presence or absence of $H_2O_2$. 60 µl of streptavidin bead slurry resulted in significantly lower peptide and protein identifications compared to 80 and 100 µl beads (**Figure 3C**). Notably, similar numbers of proteins were identified for 80 µl and 100 µl bead volumes with the presence or absence of $H_2O_2$ (**Figure 3C, Appendix Figure S2B**), with the 80 µl bead volume condition sacrificing peptide and protein IDs only minimally while achieving slightly lower nonspecific protein binding for the smaller bead volume. These results suggest that there is a wider optimum for the bead volume to be used for PL experiments. However, with a higher volume of beads, the intensities of streptavidin peptides released during proteolytic digestion also increased (**Appendix Figure S2C, Dataset EV3**), which might interfere with identification of biotinylated proteins. Based on these observations, we opted for 80 µL of streptavidin beads in subsequent experiments, considering both cost-effectiveness and performance. It is important to note that the optimal starting protein lysate amount and bead volume may depend on factors such as PL construct expression levels, cell types, and bead types.

We further aimed to enhance the throughput by testing shorter binding times for biotinylated proteins to streptavidin beads. While 1 and 2 hrs binding time seemed insufficient, minor differences were observed between 4 and 18 hrs in the number of peptides and proteins

identified, indicating the feasibility of completing the enrichment of biotinylated proteins within a single day (**Figure 3D**).

To maximize protein identifications and quantitative accuracy of the PL strategy, we fine-tuned several parameters for the MS analysis of PL samples using DIA, including precursor scan range, normalized collision energy, maximum injection time, and the overlap width of the precursor isolation windows. Optimized parameters for the DIA method consisted of one full MS scan over a range of 350-1050 *m*/*z* (**Appendix Figure S2D**), $MS^2$ scans with a normalized HCD (higher-energy collisional dissociation) collision energy of 30% (**Appendix Figure S2E**), a maximum injection time of 32 ms (**Appendix Figure S2F**), and isolation windows of 20 *m*/*z* over 350-1050 *m*/*z* range with 2 *m*/*z* overlap between windows (**Appendix Figure S2G**).

Finally, to assess the protein identification and quantification capabilities of the optimized DIA method coupled with the automated enrichment method (Auto_DIA), we compared it with manual biotinylation enrichment coupled with DIA analysis (Manual_DIA) as well as a data-dependent analysis (DDA) method with automated enrichment (Auto_DDA). Auto_DIA outperformed Manual_DIA and DDA analysis (Auto_DDA) based on the following results: (1) The number of precursor ions identified across all the three replicates were 36,992 (95%) for Manual_DIA, 31,743 (78.7%) for Auto_DDA, and increased to 40,211 (97.1%) for Auto_DIA (**Figure 3E)**. Similar trends were also observed for the peptides and proteins identified across three independent replicates (**Appendix Figure S3A-C**, **Table EV1**); (2) Auto_DIA achieved a narrower distribution of coefficients of variation (CV) for protein intensities across three independent replicates with median CV of 7%, while Manual_DIA and Auto_DDA resulted in median CVs of 13% and 11%, respectively (**Figure 3F**); (3) The correlation coefficients comparing peptide intensities of the three replicates ranged from 0.91-0.948 for Manual_DIA, 0.917-0.95 for Auto_DDA, but improved to 0.941-0.967 for Auto_DIA (**Appendix Figure S3D**).

Collectively, we established a PL strategy integrating an automated protocol for biotinylated proteins enrichment coupled with optimized DIA MS analysis, enabling high-throughput sample processing in 96-well plate format, significantly reduced manual processing time, and enhanced reproducibility in terms of protein identification and quantification across replicates (**Figure 3G**).

### *Mapping subcellular proteomes using the proximity proteomics strategy*

PL is a premier method for spatial proteomics and delineating the proteome composition of subcellular structures. To evaluate the performance of our optimized proximity proteomics strategy in subcellular proteome mapping, we generated and characterized monoclonal HEK293T cell lines expressing APEX2 proximity labeling constructs targeted to the endosome, Golgi apparatus, lysosome, and plasma membrane using the step-by-step approach outlined above. The protein localization domains included: Lyn11 for the plasma membrane, 2xFYVE for the endosome, LAMTOR1 targeting sequence and full length LAMP1 for the late endosome/lysosome, β-1,4-galactosyltransferase (GalT) for the Golgi apparatus, and nuclear export signal (NES) as control (**Appendix Figure S1, Dataset EV1**). The final selected clones exhibited homogenous expression of the APEX2 construct with over 92% GFP-positive cells, except for the NES-APEX2 construct (**Appendix Figure S4A**). Furthermore, the APEX2 constructs localized to the corresponding subcellular compartments of interest were confirmed

by co-localization analysis with markers, including Vimentin for cytosol, E-cadherin for plasma membrane, Rab5 for early endosome, Rab9 for late endosome/lysosome, and Golgin-97 for Golgi apparatus (**Figure 4A**). Lastly, the final selected clones demonstrated median biotinylation intensities and minimal background biotinylation among other monoclonal cell lines (**Appendix Figure S4B**), which varied among different clones, likely due to the distinct expression levels of APEX2-tagged localization constructs (**Appendix Figure S4C**). Following biotinylation, proteins underwent automated enrichment, on-bead proteolytic digestion, and quantitative DIA-MS. All cell lines showed successful biotinylation and efficient enrichment, as assessed by western blot analysis (**Appendix Figure S5**).

Principal component analysis (PCA) comparing the biotinylated proteins for each localization construct showed separation of each subcellular structure, but clustering of their biological replicates (**Figure 4B**). To determine compartment-specific proteins, we performed a two-step filtering process. First, protein abundances for each location construct were compared to the control (NES-APEX2) using MSstats (Choi *et al*, 2014) (**Figure 4C, Appendix Figure S6**). All proteins with a $\log_2$ fold change >1 and adjusted p-value <0.05 were considered candidate compartment-specific proteins (**Dataset EV4**). Second, significant proteins were required to be exclusively present in one comparison, except for the proximal proteomes of the endosome (2xFYVE) and lysosome/late endosome (LAMP1, and LAMTOR1) due to an expected significant overlap in their subcellular proteomes. Based on these criteria, 337, 55, and 228 location-specific proteins were identified for the endosome and lysosome/late endosome (Endo-Lyso), Golgi apparatus, and plasma membrane, respectively (**Figure 4D**). Based on the median normalized $\log_2$FC (fold change) across all samples, the location-specific proteins were divided into three groups (**Figure 4E**). Gene ontology enrichment analysis demonstrated that, as expected, proteins specific for 2xFYVE, LAMP1, and/or LAMTOR1 in Group 1 were enriched for the early endosome, endosome, and lysosome; Group 2, comprising proteins specific for GalT, showed enrichment for Golgi apparatus subcompartment and Golgi membrane, while proteins in group 3, specific for Lyn11, were enriched in actin filament-based process, anchoring junction, and plasma membrane region.

Finally, we compared the location-specific proteins derived from our PL analysis to the proteins defined for the respective compartments in the Human Cell Map (Go *et al*, 2021), a proximity map of the HEK293 proteome generated using BioID-based PL. As anticipated, proteins specific for 2xFYVE, LAMP1, and LAMTOR1 were enriched for the endosome and lysosome proteome determined by the Human Cell Map, while proteins specific for GalT were enriched in the Golgi apparatus, and Lyn11-specific proteins were enriched in the plasma membrane and cell junction (**Figure 4F**). Notably, our dataset also uncovered location-specific proteins not identified in the Human Cell Atlas, such as Golgin-45 (Short *et al*, 2001), sphingomyelin synthases 2 (Villani *et al*, 2008), and palmitoyltransferase ZDHHC14 (Adachi *et al*, 2016) for Golgi apparatus; cell surface hyaluronidase (Yamamoto *et al*, 2017), pleckstrin homology domain-containing family O member 1 (Olsten *et al*, 2004), and guanine nucleotide-binding protein G(i) subunit alpha-1 (Cho & Kehrl, 2007) for plasma membrane; and proteins involved in endosomal trafficking, like Ras-related protein Rab-4A (RAB4A) (Nag *et al*, 2018), charged multivesicular body protein 2a (CHMP2A) (Yang *et al*, 2022), sorting nexin-11 (SNX11) (Xu *et al*, 2013), sorting nexin-16 (SNX16) (Hanson & Hong, 2003), and vacuolar protein sorting-associated protein 41 homolog

(VPS41) for endosome/lysosome (Bowers & Stevens, 2005). Taken together, employing our automated PL strategy enabled the identification of location-specific proteins in various cellular compartments, underscoring its applicability to spatial proteomics.

### *Mapping the activity-dependent proximal interaction network of the $5HT_{2A}$ receptor*

Since PL is also applicable to map proximal protein networks of a protein of interest (POI), we sought to evaluate our PL pipeline for studying dynamics of proximal protein networks. As a proof of concept, we selected the $5HT_{2A}$ serotonin receptor, a member of the G protein-coupled receptor (GPCR) family, to map its proximal interaction network changes upon activation by its endogenous ligand, serotonin (5-hydroxytryptamine; 5HT). Despite the therapeutic relevance of $5HT_{2A}$ (Kwan *et al*, 2022; McClure-Begley & Roth, 2022), beyond a subset of known signal transducers and regulatory molecules, little is known about the interaction network in response to $5HT_{2A}$ activation. APEX-based PL allows the systematic mapping of its proximal interaction network upon receptor activation, which could unveil novel proteins regulating the receptor's activity that might in the future provide additional targets to finetune receptor activity.

To generate the $5HT_{2A}$ PL construct, we placed the APEX2 enzyme within the C-terminal tail of the receptor to avoid disturbing the C-terminal PSD-95/Discs-large/ZO-1 (PDZ)-binding domain, which is essential for receptor-mediated signaling and trafficking (Xia *et al*, 2003) (**Appendix Figure S7A**). Following monoclonal HEK293T cell line generation, we validated that the $5HT_{2A}$-APEX2 construct localized to the plasma membrane (**Appendix Figure S7B**) and retained signaling activity determined based on Gq heterotrimer dissociation (**Appendix Figure S7C**). For PL, the receptor was activated with 5HT over a 30-minute time course in the presence of biotin phenol and $H_2O_2$, followed by automated enrichment of biotinylated proteins and DIA-MS (**Figure 5A**). Our optimized DIA-MS method quantified 6429 proteins across the dataset. To identify proteins with agonist-dependent labeling changes, we performed an analysis of variance (ANOVA) by fitting a polynomial curve for each protein across the time course (**Figure 5A**), identifying 744 proteins with significant differences following 5HT treatment (adjusted p-value <0.05) (**Figure 5B, Dataset EV5**). An overrepresentation analysis of location-specific proteins for plasma membrane, endosome, late endosome/lysosome (**Figure 4D**) suggested a decrease in biotinylation of plasma membrane-specific proteins and an increase of endosome and lysosome-specific proteins across the time-course, indicating ligand-induced receptor trafficking (**Figure 5B**). To distinguish between (1) location-specific proteins residing in the receptor's local environment but do not participating in its function and (2) protein networks potentially regulating receptor signaling and trafficking, we examined the fold change (FC) distribution of the location specific proteins and further filtered the significant proteins with FC > 1.5 (**Figure 5C**). Based on the temporal profiles, the remaining 142 proteins formed two distinct clusters: a cluster representing proteins with rapid kinetics and transient changes in biotin labeling at 1 minute after agonist treatment and the second cluster with later, sustained differences in labeling across 30 minutes of 5HT treatment (**Figure 5D**).

The transient cluster was enriched for proteins functioning in GPCR signaling and Rho GTPases signaling (**Figure 5E**), including well-characterized interactors of $5HT_{2A}$ such as G protein-coupled receptor kinase 2 (GRK2), which phosphorylates $5HT_{2A}$, and beta-arrestin 2 (ARRB2), which binds to the phosphorylated receptor to meditate receptor desensitization and

internalization (Allen *et al*, 2008). We also observed increased transient labeling of downstream $5HT_{2A}$ signaling targets such as protein kinase C (PKCD, PKCA, and PKCQ; isoforms of protein kinase C), protein kinase D (PRKD2), and diacylglycerol kinase zeta (DGKZ), which have long been implicated in $5HT_{2A}$ signaling (Roth & Chuang, 1987; Allen *et al*, 2008). Moreover, $5HT_{2A}$ activation triggered the signaling by Rho GTPases, which play a key role in the regulation of actin dynamics through activation of WASPs (Wiskott-Aldrich syndrome proteins) and WAVEs (WASP family verprolin-homologous) (Sit & Manser, 2011). Our dataset recovered four proteins of WASP/WAVE complex with 5HT-dependent increase in PL. While the transient cluster contained mainly proteins involved in receptor signaling, the sustained cluster showed enrichment of proteins involved in membrane trafficking (**Appendix Figure S7D**). Specifically, we identified regulators of clathrin-mediated endocytosis including BMP-2 inducible kinase (BMP2K), the mannose 6-phosphate receptor (M6PR), as well as the clathrin adaptor proteins Epsin-1 (EPN1) and Epsin-2 (EPN2) in our dataset (McMahon & Boucrot, 2011; Wang *et al*, 2016; Mettlen *et al*, 2018). Previous studies have demonstrated the involvement of clathrin components in $5HT_{2A}$ receptor endocytosis (Bhatnagar *et al*, 2001). The results from the $5HT_{2A}$ receptor study suggest that our PL strategy can capture transient and sustained dynamics of proximal protein networks. We envision that automated enrichment in 96-well plate format combined with DIA-based MS accelerates PL studies across many conditions in parallel and, given its quantitative reproducibility, allows for sensitive detection of PL changes.

### ***Exploring proximal interaction network perturbations induced by CRISPR-based gene knockout using low-input PL pipeline***

Next, we sought to assess whether our PL pipeline is amenable to monitor interaction network changes across multiple conditions. As a proof of concept, we wanted to examine how CRISPR-based knockout (KO) of individual network components affect the overall activity-dependent $5HT_{2A}$ interaction network (**Figure 6A**). High-throughput PL studies across various conditions are not only limited by the enrichment of biotinylated proteins but also by the scale of the cell culture required to generate sufficient sample input amounts. Given the stable performance of the automated enrichment of biotinylated proteins across a range of protein input amounts and its high reproducibility (**Figure 3, Appendix Fig. S2**), we reasoned that our PL strategy (high-input PL) could be applicable to lower input amounts (low-input PL). To modify the high-input PL pipeline for lower sample inputs, we adjusted several parameters, including reducing cell culture scale from a 10-cm dish to a 6-well plate format, decreasing bead and reagent volumes throughout the enrichment and protein digestion, as well as increasing sample injection volume for MS analysis. All parameters for the low-input PL pipeline are listed in **Table EV2**. Next, we compared the performance of the high- and low-input PL pipeline based on reproducibility in protein identification and quantification following PL of the activated $5HT_{2A}$-APEX. While the low-input PL protocol identified overall less peptide features, the protein identifications for the low-input protocol decreased by only 8.6% compared to the high-input (**Figure 6B**). Moreover, the quantitative reproducibility was only marginally affected by the low-input PL protocol increasing the average coefficient of variation from 7.6% to 10.5% (**Figure 6C**). These results suggest that it is indeed feasible to utilize the low-input protocol to perform PL studies without compromising performance substantially.

To examine how CRISPR-based KO of individual network components affects activation-dependent changes in the $5HT_{2A}$ interaction network, we selected two proximal $5HT_{2A}$ interactors from the transient cluster (**Figure 5D**), ANXA2 (Annexin A2) and S100A10 (also known as p11). S100A10 and ANXA2 form a heterotetramer (Réty *et al*, 1999), which has been shown to bind to membrane proteins, such as the metabotropic glutamate receptor 5 (mGluR5) (Lee *et al*, 2015), L-type voltage-gated calcium channel (VGCC) (Jin *et al*, 2020), and a subset of the 5HT receptor family (Svenningsson *et al*, 2006; Warner-Schmidt *et al*, 2009) assisting their trafficking to the plasma membrane (Chen *et al*, 2022). To deplete endogenous ANXA2 and S100A10 by CRISPR-Cas9 KO, two guide RNAs (gRNA) were designed per gene, complexed with Cas9 and electroporated into $5HT_{2A}$-expressing HEK293T cells (**Figure 6A**). Two non-targeting control (NTC) gRNAs were included as negative controls. We demonstrated high KO efficiency in the polyclonal KO cell lines for both, S100A10 and ANXA2, as evaluated by inference of CRISPR edits or Western Blot analysis, respectively (**Figure 6D-E**). Using the low-input PL protocol we performed APEX-based PL of the 5HT-activated $5HT_{2A}$ receptor across multiple time points in different KO cell lines.

The majority of proteins in the transient (38/40) (**Figure 6F**) and the sustained (89/102) (**Appendix Figure S8**) clusters of the 5HT-evoked $5HT_{2A}$ proximal proteome discovered with the high-input PL protocol were also captured by the low-input protocol. Notably, comparable kinetics were observed across the timepoints of 5HT treatment when comparing the high- and low-input PL protocols. Moreover, the data suggested a high level of consistency and reproducibility across knockout (KO) conditions (**Figure 6F, Appendix Figure S8**). In ANXA2-KO cells, there was a significant reduction in S100A10 protein levels compared to NTC cells, both with and without $5HT_{2A}$ receptor activation. However, the converse knockout of S100A10 did not significantly affect ANXA2 protein levels (**Figure 6G**). Additionally, in both S100A10-KO and ANXA2-KO cells, the protein AHNAK exhibited reduced proximal labeling across all conditions (**Figure 6G**). Neither ANXA2 nor S100A10 KO resulted in significant changes in sustained, activity-dependent proximal interactions (**Appendix Figure S8**). Finally, to investigate the effect of ANXA2 and S100A10 KO on cell surface expression of the $5HT_{2A}$ receptor, we quantified the receptor pool at the plasma membrane using flow cytometry in KO and non-targeting control cells. The results indicated a small but significant decrease in the expression of the $5HT_{2A}$ receptor at the plasma membrane in both ANXA2 and S100A10 KO cells (**Appendix Figure S9**). This validates the potential of our method for identifying biologically relevant interactors and, specifically, suggests a role of the ANXA2-S100A10 complex in determining $5HT_{2A}$ receptor trafficking or stability, but further functional validation is necessary to test this hypothesis and explore the functional significance of the novel interactions detected.

Together, our study demonstrates the efficacy of the automated, low-input PL protocol in delivering reproducible results at a smaller scale and in monitoring interaction network perturbations across multiple conditions with CRISPR KO of proximal interactors.

## Discussion

In this study, we present a scalable and reproducible approach for PL-based proteomics aimed at accelerating large-scale PL investigations (**Figure 1**). Initially, we outlined a detailed protocol for the generation and characterization of monoclonal cell lines for PL (**Figure 2**), a crucial step for ensuring consistent PL levels across diverse experimental conditions. Subsequently, we implemented an automated 96-well format method on the KingFisher Flex instrument for enriching biotinylated proteins, thus minimizing variability inherent in manual sample processing (**Figure 3**). Compared to the labor-intensive 3-hour manual enrichment process capable of handling up to around 16 samples, our automated approach enables processing of 96 samples in parallel with only 30 minutes of manual intervention. To optimize the automated protocol, we fine-tuned various parameters, including the amounts of protein lysate input and streptavidin beads. While these parameters may vary depending on factors such as the PL enzyme used, the expression level of the PL construct, and the cell type, our study provides a guide for parameter optimization. Additionally, we evaluated different binding times for the biotinylated proteins to streptavidin beads. While an 18-hours incubation exhibited best performance in terms of protein identification, shortening the binding time to 4 hours resulted in only a marginal decrease in protein identification, suggesting the possibility of shortening the protocol to one day. Furthermore, in addition to improving reproducibility in sample preparation, we adopted a DIA-based MS method for the measurement of biotinylated proteins, offering enhanced reproducibility and consistency in protein identification and quantification (**Figure 3**). The consistency in quantification is of particular importance for studies across multiple conditions, for example receptor activation over a time course. In previous APEX2 studies we used a combination of DDA-based and targeted proteomics approaches to ensure consistent quantification (Lobingier *et al*, 2017; Polacco *et al*, 2024), which duplicated the measurement time compared to using the DIA approach itself. Integration of manual enrichment with DIA-based protein quantification already improved the reproducibility in protein identification from 78.7% to 95% overlap among three replicates. The incorporation of automated enrichment not only boosted throughput but also reduced variability in protein quantification from a median CV of 13% to 7%. While our strategy automated the enrichment of biotinylated proteins, including binding to streptavidin beads followed by rigorous washing steps, on the KingFisher Flex platform, subsequent proteolytic digestion of biotinylated proteins and clean-up for MS analysis were performed manually in 96-well plate format. Several recent studies have proposed methodologies to automate these additional steps (Fu *et al*, 2018; Liu *et al*, 2021; Müller *et al*, 2020), promising further reduction in hands-on time for sample preparation and increased reproducibility, albeit requiring an additional liquid handling platform.

We showcased the versatility of our automated proximity proteomics approach in mapping subcellular proteomes by targeting the APEX2 PL enzyme to early and late endosomes, the Golgi apparatus, lysosomes, and plasma membrane (**Figure 4**). The inter-compartmental crosstalk and membrane contact sites pose challenges in deciphering compartment-specific proteins solely based on existing knowledge (Go *et al*, 2021; Eisenberg-Bord *et al*, 2016). Hence, we used the PL proteome generated by cytosolic APEX2-construct (NES-APEX2) as control and required proteins that were significantly overrepresented in the subcellular compartments compared to the control to be exclusively present in one compartment. Using

these stringent filtering criteria, we identified 337, 55, and 228 location-specific proteins for endosome and lysosome/late endosome, Golgi, and plasma membrane, respectively. The majority of our compartment-specific proteins aligned with their location defined in the Human Cell Map (Go *et al*, 2021), a proximity map of the HEK293 proteome generated using PL. Discrepancies between our dataset and the Human Cell Map could be attributed to the usage of different proximity labeling enzymes (APEX2 versus BioID), different proteins or protein localization domains employed to target the enzyme to a subcellular location, as well as variances data analysis criteria. Notably, expression of an APEX2 construct (or other proximity labeling constructs) at a subcellular compartment could potentially impact the function of the compartment. We envision that our scalable proximity labeling strategy will facilitate the investigation of dynamics in subcellular protein localization across various experimental conditions.

As a second application of the automated proximity proteomics strategy, we employed it to investigate temporal changes in protein interactions. Previously, we demonstrated that APEX2-based PL could delineate protein interaction networks of GPCRs and their agonist dependent changes with temporal and spatial precision (Lobingier *et al*, 2017; Paek *et al*, 2017). To ensure consistent and reproducible quantification over the time course, we used a two-step quantitative MS workflow consisting of DDA for identification followed by targeted proteomics for quantification of biotinylated proteins in our prior work (Lobingier *et al*, 2017). Here, we demonstrated the capability of our automated PL approach combined with DIA-MS to yield high reproducibility in quantifying agonist-induced proximal proteomes over a time course. We delineated the changes in the proximal protein interaction network of $5HT_{2A}$ upon activation with its endogenous ligand, 5HT (**Figure 5**). The 744 proteins exhibiting significant changes in the $5HT_{2A}$ PL dataset displayed a down-regulation of plasma membrane proteins and up-regulation of endosomal proteins, indicative of receptor internalization and trafficking following activation. To account for location-specific proteins that reside in the local environment of $5HT_{2A}$ and enrich for functionally relevant proteins in the receptor's proximal interaction network, we filtered the dataset for significant proteins using a defined fold change cutoff. This filtering criteria retained proteins with well-characterized functions in GPCR signaling and trafficking, such as GRK2, ARRB2, PKC and PKD. Notably, the transient nature of these interactions has traditionally posed challenges for their capture using conventional affinity purification coupled to MS. Furthermore, although HEK293T cells serve as an excellent model system for demonstrating the efficacy of the automated PL approach to detect PPI dynamics with high sensitivity, these cells do not endogenously express the $5HT_{2A}$ receptor. Therefore, we emphasize the necessity of conducting PPI mapping in other cell types expressing endogenous $5HT_{2A}$, ideally by endogenously tagging the receptor with the APEX enzyme to avoid artifacts from overexpression, and/or conducting functional characterization to validate any hypotheses derived from the proximal interaction networks.

Finally, we assessed the feasibility of using our PL pipeline for monitoring proximal interaction network perturbations induced by CRISPR-based gene knockout (KO) (**Figure 6**). To achieve this, the PL pipeline was modified to accommodate lower sample inputs, which compared to the high-input PL pipeline maintained similar performance based on protein identification and

quantitative consistency. Subsequently, we examined the impact of CRISPR-based KO of individual network components on the activity-dependent $5HT_{2A}$ interaction network. Specifically, ANXA2 and S100A10, two proteins in the proximal interaction network of $5HT_{2A}$, were selected for KO, with the effects assessed using the low-input PL pipeline. In ANXA2-KO cells, a significant reduction in S100A10 protein levels was observed. This is consistent with prior studies showing reduced protein levels of p11 in various tissues from Anxa2-KO mice (He *et al*, 2008), suggesting a role for ANXA2 interaction in the stabilization of S100A10 (P11). The converse knockout of S100A10 in our study did not affect ANXA2 levels. Furthermore, both S100A10-KO and ANXA2-KO cells showed reduced labeling of the protein AHNAK across all conditions. This observation also aligns with previous studies demonstrating AHNAK's interaction with the ANXA2-S100A10 complex (Jin *et al*, 2020), and the requirement of ANXA2 and S100A10 for its recruitment to the plasma membrane (Benaud *et al*, 2004). Notably, investigation into the effect of ANXA2 and S100A10 KO on $5HT_{2A}$ receptor, specifically receptor expression at the plasma membrane, revealed a slight decrease in both KO cell lines. Though the interaction between the S100A10/ANXA2 heterotetramer and AHNAK has been well characterized (Benaud *et al*, 2004; Jin *et al*, 2020; Chen *et al*, 2022), it is the first time to suggest their potential participation in the $5HT_{2A}$ interaction network. The S100A10/ANXA2 heterotetramer interacts with a consensus binding motif in the C-terminal region of AHNAK (Chen *et al*, 2022; Jin *et al*, 2020). AHNAK consists of a PDZ domain in the N terminal region, so possibly the binding of AHNAK/S100A10/ANXA2 is mediated by the PDZ binding motif in 5-HT2A. However, further functional validation is needed to confirm our findings and their exact mechanism. Overall, the results highlight the efficacy of the automated, low-input PL pipeline in conjunction with CRISPR KO of proximal interactors for delivering reproducible results and monitoring interaction network perturbations across multiple conditions.

Collectively, we have developed a PL proteomics approach enabling the mapping of subcellular proteomes and the characterization of spatially and temporally resolved proximal interaction networks with high throughput and reproducible quantification. The method presented here holds great promise for large-scale proximity labeling studies. While we demonstrated this automated proximity proteomics for APEX2-based PL and implemented it on a KingFisher Flex platform, we anticipate its broad applicability in two aspects: (1) its potential extension to various other PL enzymes that function by biotinylating proteins, and (2) its seamless adaption to other automated platforms capable of handling magnetic beads.

## Materials and Methods

### *Reagents and Tools*

| Reagent/resource | Reference or source | Identifier or catalog number |
|---|---|---|
| **Cell lines** | | |
| HEK293T/17 cells | ATCC | CRL-11268 |
| **Antibodies** | | |
| Mouse anti-FLAG (M1) | Sigma | F-3040 |
| Mouse anti-Vimentin | Invitrogen | MA3-745 |
| Rabbit anti-E Cadherin | CST | 3195T |
| Mouse anti-RAB5A | Fisher | 89333555 |
| Mouse anti-RAB9 | Invitrogen | MA3-067 |
| Rabbit anti Golgin-97 | CST | 13192 |
| Chicken anti-GFP | VWR | RL600-901-215 |
| AF488-labeled goat anti-Chicken | Invitrogen | A11039 |
| AF555-labeled goat anti-Mouse | Invitrogen | A21422 |
| AF647-labeled goat anti-Rabbit | Invitrogen | A21244 |
| Rabbit HRP anti-GAPDH | Bio-legend | 607904 |
| HRP anti-Streptavidin | VWR | N100 |
| **Recombinant DNA** | | |
| NES-APEX2 | This study | N/A |
| 2xFYVE-APEX2 | This study | N/A |
| GalT-APEX2 | This study | N/A |
| LAMP1-APEX2 | This study | N/A |
| LAMTOR1-APEX2 | This study | N/A |
| Lyn11-APEX2 | This study | N/A |
| $5HT_{2A}$-APEX2 | This study | N/A |

| | | |
|---|---|---|
| $G\alpha_q$-RLuc8 | Addgene | 140982 |
| $G\beta_3$ | Addgene | 140988 |
| GFP2-$G\gamma_9$ | Addgene | 140991 |
| $5HT_{2A}$ | (Kim *et al*, 2020) | N/A |
| **Chemical, enzymes and other reagents** | | |
| Dulbecco's Modification of Eagle's Medium (DMEM) | Corning | 10-013-CV |
| Fetal bovine serum (FBS) | Gibco | A31605-01 |
| Dulbecco's Phosphate-Buffered Saline (DPBS) | Corning | 21-031-CV |
| DPBS, calcium, magnesium | Gibco | 14040182 |
| 0.05% Trypsin-EDTA | Gibco | 25300-054 |
| Penicillin streptomycin solution | Corning | 30-002-CI |
| PolyJet | SignaGen | SL100688 |
| 0.45 µm PVDF filter unit | Millipore | SE1M003M00 |
| Cellstripper | Corning | 25-056-CI |
| Attune performance tracking beads | Thermo Fisher | 4449754 |
| ECL western blotting substrate | Pierce | 32106 |
| Streptavidin magnetic beads | Pierce | 88817 |
| Biotin phenol (Biotin Tyramide) | Iris Biotech | LS-3500 |
| Hydrogen peroxide | Sigma | H1009-100ML |
| Sodium azide | Sigma | S2002 |
| Sodium ascorbate | Spectrum | S1349 |
| Trolox (6-hydroxy-2,5,7,8-tetramethylchromane-2-carboxylic acid) | Sigma | 238813 |
| 16% Formaldehyde (w/v) | Pierce | 28908 |
| EverBrite mounting medium with DAPI | Biotium | 23002 |

| | | |
|---|---|---|
| cOmplete protease inhibitor cocktail tablets mini, EDTA-free | Roche | 11846170001 |
| 660nm protein assay kit | Pierce | 22662 |
| TCEP | Pierce | 20490 |
| Dithiothreitol (DTT) | Sigma | D0632 |
| Iodoacetamide (IAA) | Sigma | I1149 |
| Sequencing-grade modified trypsin | Promega | V5111 |
| Lysyl endopeptidase | VWR | 100369-822 |
| BioPureSPE 96-well C18 plate | NEST | HNS S18V |
| Ascorbic Acid | Sigma | A5960 |
| BSA, Free Fatty Acid | Akron | AK8909 |
| Coelenterazine 400a | Nanolight | 340-1 |
| Dialyzed FBS | Gibco | A3382001 |
| Hank's Balanced Salt Solution | Gibco | 14065056 |
| Sodium chloride | Sigma | S7653-250G |
| Tris HCl pH 7.5 | Corning | 46-030-CM |
| Tris HCl pH 8.0 | Corning | 46-031-CM |
| Sodium deoxycholate | Sigma | D6750-500G |
| Triton-x100 | Sigma | T9284-100ML |
| Sodium dodecyl sulfate | Fisher BioReagents | BP166-500 |
| Urea | Promega | V3171 |
| Potassium chloride | Sigma | P9541-1KG |
| Sodium carbonate | Sigma | S2127-500G |
| 5-hydroxytryptamine hydrochloride | TCI | S0370 |
| Pierce 660 nm Protein Assay Reagent | Thermo Scientific | 22660 |
| SF cell line 96-well nucleofector kit | Lonza Walkersville | V4SC2096 |
| HEPES buffer | Corning | 25-060-CI |

| **Software** | | |
|---|---|---|
| Spectronaut (version 16.0) | Biognosys | https://biognosys.com/shop/spectronaut |
| MSstats (version 4.4.0) | Bioconductor | https://www.bioconductor.org/packages/release/bioc/html/MSstats.html |
| Attune NxT software | Thermo Fisher Scientific | |
| FlowJo | FlowJo, LLC | https://www.flowjo.com/ |
| Fiji | imageJ | https://imagej.net/software/fiji/ |
| NIS-Elements software (v. 5.30.01 build 1541) | Nikon | |
| Prism (v8.0) | GraphPad Software | |
| **Other** | | |
| Orbitrap Exploris 480 MS with internal calibration option | Thermo Fisher Scientific | BRE725533 |
| EASY-nLC 1200 system | Thermo Fisher Scientific | |
| Nikon Ti2-E microscope | Nikon | |
| Attune NxT Flow Cytometer | Thermo Fisher Scientific | |
| KingFisher Flex system | Thermo Fisher Scientific | |
| 4D-Nucleofector | Lonza Bioscience | |

***Methods and Protocols***

***Mammalian Cell Culture Conditions***

HEK293T cells (CRL-1583, ATCC) were maintained in Dulbecco's Modified Eagle's Medium (DMEM, GIBCO or Fisher Scientific) supplemented with 10% (v/v) fetal bovine serum (FBS) (GIBCO) and 1% penicillin/streptomycin solution at 37 $^{o}$C in a 5% $CO_2$ humidified incubator.

***cDNA constructs***

Standard cloning techniques were used to generate all constructs. Vectors were digested using restriction enzymes (EcoRI and BamHI) and PCR fragments were amplified using Q5 DNA

polymerase (NEB). NES-APEX2, Lyn11-APEX2, 2xFYVE-APEX2, LAMTOR1-APEX2, LAMP1-APEX2, GalT-APEX2, and $5HT_{2A}$-APEX2 were cloned into pLVX-TetOne-Puro by in-Fusion HD cloning. Ligated plasmid products were transformed into stellar competent cells (*E. coli* HST08 strain, Takara Bio). See **Dataset EV1** for full descriptions of constructs.

### *Lentiviral generation and transductions*

To generate lentiviruses, 8 x $10^6$ HEK293T cells were seeded in T175 flask. The next day, cells were transfected with 5 µg of each of the pLVX-TetOne-Puro lentiviral plasmids containing the gene of interest, 3.33 µg of Gag-Pol-Tat-Rev packaging vector, and 1.66 µg of VSV-G envelope vector with 30 µL PolyJet reagent (3 µL per µg DNA) in 250 µL serum-free DMEM (25 µL per µg DNA). After 24 hrs, the culture media was replaced with fresh DMEM supplemented with 10% FBS and 1% Penicillin-Streptomycin solution. Virus was harvested at 48 hrs post media change. Briefly, the media was collected and centrifuged at 2500 rpm for 5 minutes. The supernatant was filtered through a 0.45-µm PVDF filter and mixed with the PEG-salt solution (a final 8.5% PEG and 0.3 M NaCl). The virions were aggregated and precipitated at 4 $^{o}$C for at least 2 hrs, followed by centrifugation at 3500 rpm for 20 minutes. Immediately after spinning, the virus pellet was resuspended with 500 µL sterile PBS. For generation of stable APEX2-tagged cell lines, HEK293T cells were infected with lentivirus at ~50% confluency and selected by 2 µg/mL puromycin in growth medium.

### *Flow cytometric analysis*

To create monoclonal stable cell lines, polyclonal cells were single cell sorted in a 96-well plate by fluorescence-activated cell sorting (FACS) using BD FACS Aria or BD FACS Fusion. Polyclonal cells were washed once with DPBS, collected by using Cellstripper (Corning). In the case of non-GFP tagged expression construct ($5HT_{2A}$-APEX2), cells were incubated with Alexa 647-conjugated M1-anti-FLAG antibody (1:1000 dilution) at 4 $^{o}$C for 40 min. After spinning down, cells were washed with ice-cold PBS and resuspended in 1mL of sample buffer (PBS/1mM EDTA/25 mM HEPES/1% FBS) for sorting. 24 clones were selected and expanded in a 24-well plate for characterization by flow cytometry using the Attune NxT Flow Cytometer. The cellular morphology was first gated based on an FSC-A vs. SSC-A (forward scatter - area vs. side scatter - area) plot. Next, gating was generated based on the FSC-A vs. FSC-H (area vs. height) plot to select singlets. These singlets were further analyzed to identify the cells expressing $5HT_{2A}$ or the localization domains based on Alexa 647 or GFP fluorescence signal, respectively. The data was processed in FlowJo.

To quantify the $5HT_{2A}$ receptor at the plasma membrane, the non-targeting controls (NTC), ANXA2 KO, and S100A10 KO cells were washed with an ice-cold sample buffer (PBS/1mM EDTA/25mM HEPES/1% dialyzed FBS) and lifted in Cellstripper for 5 min at room temperature. After neutralization with an equal volume of cell culture media (DMEM/10% dialyzed FBS/ 1% penicillin/streptomycin), the cells were collected and incubated with Alexa 647-conjugated M1-anti-Flag antibody (1:1000 dilution) at 4 $^{o}$C for 40 min. Median fluorescence intensity of 10,000 cells per condition was measured using Attune NxT Flow Cytometer.

### *Confocal fluorescence imaging*

HEK293T cells expressing NES-APEX2, Lyn11-APEX2, 2xFYVE-APEX2, LAMTOR1-APEX2, LAMP1-APEX2, and GalT-APEX2 were plated onto 12 mm collagen coated glass coverslips in 12-well plate and treated with 1 µg/mL doxycycline for 24 hours. For initial polyclonal and monoclonal cell lines selection, after aspirating culture media, cells were fixed with 4% paraformaldehyde in PBS at room temperature (RT) for 15 min. Cells were then mounted onto glass slides with EverBrite mounting medium with DAPI (Biotium Cat #23002), and ready for imaging. For the final monoclonal cell line characterization, the localization of all spatial APEX2 constructs were determined by the co-localization analysis with subcellular location markers. Cells were first washed with PBS twice and fixed in 4% paraformaldehyde in PBS for 15 min at room temperature (RT). After washing with PBS, cells were permeabilized with 0.2% Triton X-100 in PBS for 10 min at RT and blocked in 5% BSA (w/v) in PBS with 0.1% Tween-20 (PBS-T) for 1h at RT. Incubation with primary antibodies diluted in 1% BSA in PBS-T against GFP (1:500) and Vimentin (1:100) for NES, GFP (1:500) and E-cadherin (1:500) for Lyn11, Rab 5 (1:200) for 2xFYVE , Rab9 (1:200) for LAMTOR1 and LAMP1, Golgin-97 (1:200) for GalT, and Alexa 647-conjugated M1-anti-FLAG antibody (1:200) for $5HT_{2A}$ was performed at 4 $^{o}$C overnight. After washing with PBS-T three times, cells were incubated with AF488-labeled goat anti-Chicken, AF555-labeled goat anti-Mouse, and AF647-labeled goat anti-Rabbit secondary antibodies diluted in 1% BSA in PBS-T (1:1000) for 1 hour at RT. Cells were then washed in PBS-T twice and one final wash in PBS, mounted onto glass slides using EverBrite mounting medium with DAPI (Biotium Cat #23002), and imaged by confocal fluorescence microscopy. Images were generated using a Nikon Ti2-E microscope equipped with a Crest X-Light-V2 spinning disk confocal (Crest Optics), emission filters, 438/24 (DAPI), 511/20 (GFP), 560/25 (RFP), and 685/40 (Cy5), and Celeste Light Engine excitation lasers 405/477/546/638 nm used respectively (Lumencor), Piezo stage (Mad City Labs), and a Prime 95B 25mm CMOS camera (Photometrics) using a Plan Apo VC 100x/1.4 Oil (Nikon). The data was captured with NIS-Elements software (v. 5.41.01 build 1709, Nikon) and processed in Fiji/ImageJ2 (Schindelin *et al*, 2012).

### *BRET assay*

Bioluminescence resonance energy transfer (BRET) assays were performed to validate the functionality of the $5HT_{2A}$-APEX2 construct. Here, the TRUPATH platform for measuring G protein dissociation was used as previously described (Olsen *et al*, 2020; Kim *et al*, 2020; DiBerto *et al*, 2022). In brief, wildtype or APEX2 constructs of $5\text{-}HT_{2A}$ were co-transfected with $Ga_q$-RLuc8, $Gb_3$, and GFP2-$Gg_9$ in a 1:1:1:1 ratio in HEK293T cells maintained in Dulbecco's Modified Eagle Media (DMEM) supplemented with 10% fetal bovine serum (FBS) and 1% penicillin-streptomycin (pen-strep). After 8 hours, the media were exchanged for DMEM supplemented with 1% dialyzed FBS and 1% pen-strep to minimize receptor desensitization by serum serotonin, and at least 12 hours before the experiment, cells were plated in 96-well microplates in DMEM supplemented with 1% dialyzed FBS and 1% pen-strep. To conduct the experiments, plates were first vacuum aspirated and 60 mL of assay buffer (1x Hank's Balanced Salt Solution in phosphate buffered saline, 20 mM HEPES, pH 7.4) were added to the wells. Next, 10 mL of coelenterazine 400a diluted in the assay buffer (50 mM) were added to the wells,

and allowed to incubate for 5 minutes. Next, 30 mL of 3x drug diluted in drug buffer (assay buffer supplemented with 3 mg/mL fatty acid-free bovine serum albumin and 0.3 mg/mL ascorbic acid) were added to the wells and allowed to incubate for 5 minutes. Last, plates were read using a BMG Labtech PHERAstar FSX with $BRET^2$ plus optic module. All data were analyzed using GraphPad Prism (V8.0).

***APEX proximity labeling***

HEK293T cells stably expressing $5HT_{2A}$-APEX2 and protein localization domains (NES-APEX2, Lyn11-APEX2, 2xFYVE-APEX2, LAMTOR1-APEX2, LAMP1-APEX2, and GalT-APEX2) were seeded with 350K cells/well in 6-well plate for testing biotinylation by western blot, and $3.5x10^6$ cells/dish in 10cm dish for enrichment of biotinylated proteins followed by quantitative MS. Following 24 hrs doxycycline induction, cells were incubated with 500 µM biotin phenol at 37 $^oC$ for 30 min in complete medium (DMEM/10% FBS/1% pen/strep), but $5HT_{2A}$-APEX2 cells were seeded with 10% dialyzed FBS/DMEM/1% pen/strep and changed to 1% dialyzed FBS/DMEM/1% pen/strep during doxycycline induction. Meanwhile, 0.1 µM 5HT was added for incubation of 1, 5, 10, and 30 min in cells expressing $5HT_{2A}$-APEX2 to capture the protein interaction dynamics. There was no agonist incubation for all APEX2-tagged localization domains. 2 mM $H_2O_2$ diluted in complete medium was freshly prepared prior to use. The $H_2O_2$ containing media was mixed with biotin phenol containing media at 1:1 ratio (*v*/*v*) to initiate APEX labeling. The labeling reaction was quenched after 45 sec by removing the medium, washing cells three times with ice-cold quenching solution (PBS supplemented with 10 mM sodium ascorbate, 5 mM Trolox, and 10 mM sodium azide). Cells were collected in a quenching solution (1 mL for 6-well plate and 8 mL for 10cm dish) and pelleted by centrifugation at 3,000 g for 10 min at 4 $^oC$. The supernatant was removed and the cell pellet was resuspended in lysis buffer (50 mM Tris, 150 mM NaCl, 0.5% sodium deoxycholate, 0.1% SDS, 1% Triton X-100, 10 mM sodium ascorbate, 5 mM Trolox, and 10 mM sodium azide, 1 mM DTT, and cOmplete protease inhibitor) supplemented with 10 mM sodium ascorbate, 5 mM Trolox, and 10 mM sodium azide, 1mM DTT, and protease inhibitor (100 µL for small-scale analysis and 1 mL for large-scale analysis). With a freeze-thaw cycle, each sample was sonicated for 5 seconds twice, centrifuged at 13,000 g for 10 min at 4 $^oC$. The supernatant was subjected to the automated enrichment of biotinylated proteins. The manual enrichment was performed in parallel with the automated enrichment protocol, but using a magnetic rack (DynaMag-2, Thermo Fisher). The protein concentration was determined using Protein assay 660 (Pierce). For each proteomic sample, 25 µL lysate before and after binding was saved for western blot analysis to evaluate biotinylation and enrichment efficiency of biotinylated proteins.

***Western blot analysis***

Cell lysates were mixed with equal volume of sample loading buffer (200 mM β-mercaptoethanol in NuPAGE LDS sample buffer), boiled at 95 $^oC$ for 20 min, separated on 4%-12% SDS-PAGE gels (BioRad), and transferred to PVDF membranes (Trans-blot turbo transfer system, BioRad). The blots were blocked in 5% milk in TBS-T (Tris-buffered saline with 0.1% Tween 20) at 4 $^oC$ overnight and washed three times with TBS-T for 10 min. Blots were incubated with horseradish peroxidase (HRP)-conjugated anti-Streptavidin antibody (1:5000

dilution using 2% BSA in TBS-T) or HRP-conjugated anti-GAPDH antibody (1:10000 dilution) at room temperature for 1 hour. After washing three times in TBS-T for 10 min, the blots were incubated with ECL western blotting substrate (Pierce) for 1 min. Chemiluminescent signals were captured on the Azure 400 (Azure biosystem).

***Automated enrichment protocol for biotinylated proteins***

For the automated biotinylation enrichment protocol, the Kingfisher Flex system (Thermo Fisher) is programmed to simultaneously process a maximum of 96 samples. This protocol below includes two parts, where part 1 (Plate 1-3) is for washing magnetic streptavidin beads and binding of the biotinylated proteins to beads, and part 2 (Plate 4-12) is for washing and collecting beads prior to Lys-C/trypsin digestion. The enrichment protocol is conducted in the cold room using deep-well plates. See **Dataset EV2** for full description of the Kingfisher program. A detailed step-by-step protocol is provided on protocols.io (dx.doi.org/10.17504/protocols.io.yxmvm3jbbl3p/v1).

Plate 1. Add 100 µL of 50% streptavidin magnetic bead slurry (Pierce) to each well of a 96-well plate and add 900 µL RIPA buffer (50 mM Tris, 150 mM NaCl, 0.5% sodium deoxycholate, 0.1% SDS, 1% Triton X-100, pH 7.4) to each well. The beads are mixed for 2 seconds and slowly washed for 30 seconds. At the end of this step, the beads are collected for 5 counts with collection time of 10 seconds/count. A total of 4 min for washing.

Plate 2. Add 1 mL RIPA buffer. The beads are transferred to this plate and slowly washed for 4 min.

Plate 3. Add 1 mg protein from cell lysate to each well and top-up to 1 mL with RIPA buffer. The beads are transferred to this plate for overnight binding (~12 hrs).

Plate 4-6. Add 1 mL RIPA buffer to each plate. The beads with bound biotinylated proteins are collected from Plate 3 and transferred to Plate 4-6 for washing three times with RIPA buffer. Each wash step takes 7 min.

Plate 7. Add 1 mL 1 M KCl solution. Wash the beads for 4 min.

Plate 8. Add 1 mL 0.1 M $Na_2CO_3$ solution. Wash the beads for 4 min.

Plate 9. Add 1 mL freshly prepared 2 M urea in 50 mM Tris-HCl (pH 8.0). Wash the beads for 4 min.

Plate 10-11. Add 1 mL 50 mM Tris-HCl (pH 8.0). Wash the beads twice for 5 min.

Plate 12. Add 200 µL of freshly prepared 2 M urea in 50 mM Tris-HCl buffer (pH 8.0). The beads with bound biotinylated proteins are collected in this plate for on-bead digestion.

***Sample preparation for MS analysis***

Samples were reduced with 5 mM TCEP at 37 $^{o}$C for 30 min by shaking at 1000 rpm, followed by alkylation with 5 mM iodoacetamide (IAA) at room temperature (RT) for 30 min. Extra IAA was quenched by 5 mM dithiothreitol (DTT) for 10 min at RT. Samples were then digested with 1 µg Lys-C and 1 µg trypsin at 37 $^{o}$C for 6 hrs and 25 $^{o}$C for 13 hrs. 0.5 µg trypsin was added to

each sample for an additional 2 hrs incubation at 37 $^{o}$C. Following digestion, the supernatant was transferred to a new 96-well plate and 10 µL of 10% trifluoroacetic acid (TFA) was added to each sample to a final pH 2-3. The peptide samples were desalted using C18 96-well plate (BioPureSPE, HNS S18V-20mg, the Nest group) by centrifugation at 1,600 rpm for 2 min at each step. The plate was activated with 100 µL methanol, washed three times with 100 µL 80% acetonitrile (ACN)/0.1% TFA, and equilibrated three times with 100 µL 2% ACN/0.1% TFA. After loading samples, the plate was washed three times with 100 µL 2% ACN/0.1% TFA and eluted twice with 55 µL 50% ACN/0.25% formic acid (FA). Samples were dried by vacuum centrifugation, resuspended in 20 µL 0.1% formic acid, and injected 1 µL for MS analysis.

***Mass spectrometric data acquisition***

Digested peptide mixtures were analyzed on an Orbitrap Exploris 480 MS system equipped with an Easy nLC 1200 ultra-high pressure liquid chromatography system. Samples were injected on a C18 reverse phase column (15 cm x 75 µm I.D. packed with BEH C18 1.7 µm particles, Waters) in 0.1% formic acid (FA). Mobile phase A consisted of 0.1% FA and mobile phase B consisted of 80% acetonitrile (ACN)/0.1% FA. Peptide mixtures were separated by mobile phase B ranging from 4% to 16% over 40 min, followed by an increase to 28% B over 26 min and 44% B over 4 min, then held at 95% B for 10 min at a flow rate of 300 nL/min.

To generate a spectral library, one of three biological replicates in each experiment was acquired using data-dependent acquisition (DDA). DDA analysis consisted of one full scan over a *m*/*z* range of 350-1050 in the Orbitrap at a resolving power (RP) of 120K with an RF lens of 40%, a normalized automatic gain control (AGC) target of 300%. The 20 most intense precursors at charge states of 2-6 from the full scan were selected for higher energy collisional dissociation (HCD) FTMS$^2$ analysis at RP 15K with an isolation width of 1.6 Da, a normalized collision energy (NCE) of 30, an AGC target of 200%, and a maximum injection time of 22 ms. Dynamic exclusion was enabled for 30 seconds with a repeat count of 1.

The parameters evaluated for data-independent acquisition (DIA) method include mass scan range of 350-1050 and 390-1010, MS$^2$ scans with a normalized HCD (higher-energy collisional dissociation) collision energy of 27%, 30% and 33%, maximum injection time of 22 ms, 32 ms, 42 ms, dynamic 6-point, dynamic 8-point, and Auto, and window overlap width of 0, 1, and 2 *m*/*z*. In the end, all samples were analyzed using DIA-MS by collecting one full scan over a range of 350-1050 *m*/*z* in the Orbitrap at RP 120K with an AGC target of 300% and an automatic maximum injection time, followed by DIA MS$^2$ scans over 350-1050 *m*/*z* at RP 15K using an isolation window of 20 *m*/*z*, an overlapping isolation window of 2 *m*/*z*, a normalized HCD collision energy of 30%, an AGC target of 200%, and a maximum injection time of 32 ms.

***Protein identification and quantification***

The DDA raw data for automated enrichment combined with data-dependent acquisition MS (Auto_DDA) were analyzed using MaxQuant (version 1.6.12.0) for identification of peptides and proteins. The DIA raw data for manual enrichment protocol combined with data-independent acquisition MS (Manual_DIA), automated enrichment protocol with data-independent analysis (Auto_DIA), and DIA-MS method optimization were analyzed using Spectronaut (version 16.0)

with direct DIA search. All the raw files were searched against the SwissProt Human database (downloaded 10/2020) using default settings with variable modification of methionine oxidation and protein N-termini acetylation, and fixed modification of cysteine carbamidomethylation. The data were filtered to obtain a false discovery rate of 1% at the peptide spectrum match and the protein level.

In the data analysis of methods optimization, the DIA datasets were searched against the SwissProt Human database (downloaded 10/2020) by using a spectral library-free approach (direct-DIA) in Spectronaut (version 16.0). We observed a slightly higher consistency in precursor identifications across runs when comparing library-free and library-based analysis of subcellular proteomes, so in the data analysis of mapping subcellular proteomes or protein interaction dynamics, the DDA datasets were collected to construct a spectral library. All the raw data files were searched against the SwissProt Human database (downloaded 10/2020) by using the Pulsar search engine integrated into Spectronaut (version 16.0). The raw mass spectra files were processed using the default BGS settings with digestion enzyme of trypsin, fixed modification of cysteine carbamidomethylation, variable modification of methionine oxidation and protein N-termini acetylation and filtered to a 1% false discovery rate (FDR) at the peptide spectrum match (PSM), peptide, and protein level. The DIA datasets were searched against the generated DDA spectral libraries in Spectronaut with default BGS setting, but the cross-run normalization was disabled.

***Statistical analysis of proteomic data and data visualization***

Proteomic statistical analyses were completed using the R programming language, version 4.2.0. Peptide ion intensities from Spectronaut were summarized to protein intensities using the package MSstats (version 4.4.0) and its function dataProcess enabling options to turn off model-based imputation, and turning on removal of uninformative features and outliers (Tsai *et al*, 2020). All downstream analyses were performed on log2 transformed protein intensities.

Subcellular proteomes were compared for each protein or protein group using the built-in R function lm to estimate additive effects for both subcellular location and replicate batch, and the emmeans function from package emmeans (version 1.7.3) to calculate pairwise subcellular location $\log_2$ fold changes and t-test statistics including p values. Adjusted p values were calculated as false discovery rates using the Benjamini-Hochberg method available in R built-in function p.adjust. Selection of location-specific proteins was made by identifying proteins with significant labeling increases when compared to the cytosolic NES-APEX2 sample ($\log_2 FC > 1$, unadjusted p.value < 0.005), and requiring the increase to be exclusive to GalT, Lyn11, or to the combined 2xFYVE, LAMTOR1, LAMP1 set. Gene ontology enrichment analysis of location-specific protein sets was performed using the enricher function from package clusterProfiler (version 4.4.1) and human gene ontology annotations in the Bioconductor annotation package org.Hs.eg.db (version 3.15.0). The Human Cell Map data were downloaded on 2022-July-06.

Proteomic labeling changes over 5 time points (0, 1, 5, 10, and 30 minutes) were scored per protein by fitting a cubic polynomial curve through the data using R built in functions lm and poly. A categorical term for replicate batch was included in the model, and time was continuous but transformed from minutes to ranks (1 through 5). P values were calculated using F-tests

comparing the models with and without the polynomial time terms. The time point with the largest deviation of mean intensity from time point 0 was used to calculate the log2FC as the difference in mean log2 intensities. Time courses of protein labeling changes were clustered by starting with a protein-by-time matrix of differences in mean log2 intensity from time 0, calculating protein-protein distances as *1 - R*, where R is Pearson's correlation coefficient, and defining clusters with the pam function in package cluster (version 2.1.3), with option nstart set to 20.

Networks for proteins represented in the transient and sustained cluster were generated extracting interactions between proteins in each cluster with a confidence above 0.4 derived from STRING (https://string-db.org/) (Szklarczyk *et al*, 2011). Interactions were important into Cytoscape (version 3.8.1) for visualization (Shannon *et al*, 2003). The maximum log2-fold change compared to time point 0 for each protein was visualized in the network as node color. Enrichment of Reactome pathways (https://reactome.org/) (Gillespie *et al*, 2022) was performed using Enrichr (https://maayanlab.cloud/Enrichr/) (Kuleshov *et al*, 2016; Xie *et al*, 2021; Chen *et al*, 2013).

### *CRISPR knockouts*

Polyclonal KO cells were generated by electroporation of Cas9 ribonucleoprotein complexes (Cas9 RNPs) into HEK293T cells stably expressing 5HT2A-APEX. Electroporation was performed using the SF Cell Line 4D-Nucleofector X Kit S (Lonza) and 4D-Nucleofector (Lonza). Recombinant S. pyogenes Cas9 protein used in this study contains two nuclear localization signal (NLS) peptides that facilitate transport across the nuclear membrane. The protein was obtained from the QB3 Macrolab, University of California, Berkeley. Purified Cas9 protein was stored in 20 mM HEPES at pH 7.5 plus 150mM potassium chloride, 10% glycerol, and 1mM tris(2-carboxyethyl)phosphine (TCEP) at −80°C. Each crRNA and the tracrRNA was chemically synthesized (Dharmacon/Horizon) and suspended in 10 mM Tris-HCl pH 7.4 to generate 160 µM RNA stocks. To prepare Cas9 RNPs, crRNA and tracrRNA were first mixed 1:1 and incubated 30 min at 37°C to generate 80 µM crRNA:tracrRNA duplexes. An equal volume of 40 µM S. pyogenes Cas9-NLS was slowly added to the crRNA:tracrRNA and incubated for 15 min at 37°C to generate 20 µM Cas9 RNPs. crRNAs targeting selected proteins and non-targeting controls were designed by Dharmacon. For each reaction, roughly $5x10^5$ HEK293 cells were pelleted and suspended in 20 µL nucleofection buffer. 4 µL 20 µM Cas9 RNP mix was added directly to these cells and the entire volume transferred to the bottom of the reaction cuvette. Cells were electroporated using program CM-130 on the Amaxa 4D-Nucleofector (Lonza). 80 µL pre-warmed complete DMEM was added to each well and the cells were allowed to initially recover for 10 min at 37°C followed by diluting with DMEM media, plating into 24-well plates and further recovering for 3 days. KO for selected genes was confirmed by Inference of CRISPR Edits (ICE).

### *Data availability*

The mass spectrometry proteomics data have been deposited to the ProteomeXchange Consortium via the PRIDE (https://www.ebi.ac.uk/pride/) (Deutsch *et al*, 2023) partner repository

with the dataset identifier PXD040762. The reviewer account details are shown below.
Username: reviewer_pxd040762@ebi.ac.uk
Password: poggC9R8

The source data of confocal images in Figure 2 and Figure 4 are submitted to BioImage Archive with the accession number of S-BIAD1088.

The source data of western blots in Figure 2E and Figure 6E are submitted to BioStudies with the accession number of S-BSST1404.

## Acknowledgements

This work was supported by funding from the Defense Advanced Research Projects Agency (DARPA) under the Cooperative Agreements HR0011-19-2-0020 (to N.J.K., R.H., B.L.R.) and HR0011-20-2-0029 (to N.J.K., R.H., B.L.R.). The views, opinions, and/or findings contained in this material are those of the authors and should not be interpreted as representing the official views or policies of the Department of Defense or the U.S. Government. This work further received funding from the NIH (R01DA056354 to R.H. and M.v.Z.; P01HL146366 to N.J.K.; R37DA045657 to B.L.R.). The work was carried out in the Thermo Fisher Scientific Mass Spectrometry Facility for Disease Target Discovery at the J. David Gladstone Institutes and the UCSF Center for Advanced Technology. The imaging data for this study were acquired at the Center for Advanced Light Microscopy-Nikon Imaging Center at UCSF on instruments obtained using grants from the UCSF Program for Breakthrough Biomedical Research funded in part by the Sandler Foundation and the UCSF Research Resource Fund Award. We thank Ajda Rojc, Prachi Khare, and Michael McGregor for the help with flow cytometry analysis and members of Krogan lab for helpful advice and comments.

## Author Contributions

X.Z., Q.L., and R.H. conceived and directed the study with input from N.J.K., M.v.Z., and B.L.R. Proximity proteomics: X.Z. and Q.L.; Data analysis: B.J.P., X.Z., Q.L., and R.H.; Cell line development: T.P., J.F.D., Q.L., R.H., and J.X.; Confocal fluorescence imaging: H.F., X.Z., R.V., Q.L., and A.M.; Flow cytometry: T.P., X.Z., Q.L., P.K., and A.M.; BRET assay: J.F.D.; CRISPR knockout: P.K. and X.Z.; X.Z., Q.L., R.H., and M.E. wrote the manuscript with input from all authors.

## Disclosure & Competing Interests Statement

The Krogan laboratory has received research support from Vir Biotechnology, F. Hoffmann-La Roche and Rezo Therapeutics. N.J.K. has financially compensated consulting agreements with Maze Therapeutics and Interline Therapeutics. He is on the Board of Directors and is President of Rezo Therapeutics and is a shareholder in Tenaya Therapeutics, Maze Therapeutics, Rezo Therapeutics, GEn1E Lifesciences and Interline Therapeutics. B. L. R. is a member of the Scientific Advisory Boards of Septerna Pharmaceuticals, Escient Pharmaceuticals, Onsero Pharmaceuticals, and Levator.

## Figure Legends

***Figure 1. A schematic overview of universal workflow for proximity proteomics.***

This workflow includes (1) PL construct design as well as generation and characterization of monoclonal cell lines, (2) proximity labeling, (3) automated enrichment of biotinylated proteins, (4) quantitative mass spectrometry data acquisition, and (5) the application of this workflow to map subcellular proteomics as well as protein-protein interactions and their dynamics.

***Figure 2. Step-by-step guide for generating monoclonal cell lines for proximity labeling.***

A. Schematic overview of PL construct design and generation and stepwise characterization of monoclonal cell lines for PL (PLD refers to protein localization domain and POI to protein of interest).
B. Characterization of polyclonal proximity labeling cell lines by flow cytometry (left) and confocal imaging (right). The APEX2 enzyme was targeted to the plasma membrane using a Lyn11 protein localization domain (PM-APEX2). GFP was used to evaluate expression and localization of the APEX2 construct. The non-Doxycyline induced cell line was used as control for flow cytometry. Scale bar represents 20µm.
C. Flow cytometry analysis of four monoclonal PM-APEX2 cell lines.
D. Confocal imaging of four monoclonal PM-APEX2 cell lines. Cells were stained with DAPI and Hoechst 33342 for the nucleus and the APEX2 construct location was indicated by GFP. Scale bar represents 20µm.
E. Western blot analysis of whole cell lysate derived from four monoclonal PM-APEX2 cell lines to evaluate the biotin labeling efficiency of APEX2 using streptavidin antibody. GAPDH is used as loading control.

***Figure 3. A scalable biotin enrichment method coupled to quantitative MS.***

A. Schematic overview of the development and optimization of the automated enrichment method for biotinylated proteins and quantitative mass spectrometry method.
B. Overview of the automated biotin enrichment protocol established on the KingFisher Flex.
C. Peptides and proteins identified for varying amounts of streptavidin beads (60, 80, and 100ul) in the presence and absence of $H_2O_2$ treatment to assess nonspecific protein binding to beads depending on the bead amount, n = 3, statistical significance was determined using unpaired t-test in Prism (GraphPad).(Peptides: 60uL vs. 100uL, $p<0.0001$; 80uL vs. 100 uL, $P<0.0001$. Proteins: 60uL vs. 100uL, $p<0.0001$; 80uL vs. 100 uL, $P<0.0001$.
D. Peptides and proteins identified for varying binding times (1, 2, 4, and 18 hrs) of biotinylated proteins to streptavidin beads, n = 3, statistical significance was determined using unpaired t-test in Prism (GraphPad) .
E. Performance comparison of manual enrichment protocol combined with data-independent acquisition MS (Manual_DIA), automated enrichment combined with data-dependent acquisition MS (Auto_DDA), and automated enrichment protocol with data-

independent analysis (Auto_DIA). Performance was evaluated comparing the precursor ions being identified from three replicates (n = 3).

F. Distribution of the coefficient of variation of protein intensities from three replicates comparing Manual_DIA, Auto_DDA, and Auto_DIA approaches. Numbers depicted on the graph represent median CVs.

G. Summary comparison of the manual and automated enrichment method for biotinylated proteins.

***Figure 4. Mapping subcellular proteomes using the automated proximity proteomics strategy.***

A. Confocal fluorescence imaging of monoclonal cell lines of all spatial APEX2 constructs to co-localize with a subcellular location marker. APEX2-fusion constructs were stained using a GFP antibody. Location markers of vimentin, E-cadherin, Rab5, Rab9, and Golgin-97 were used for NES (cytosol), Lyn11 (plasma membrane), 2xFYVE (endosome), LAMTOR1 and LAMP1 (late endosome/lysosome), and GalT (Golgi apparatus), respectively. (Scale bars, 10 µm.)

B. Principal-component analysis (PCA) comparing proximal proteomes of each replicate of the spatial APEX2 constructs, n = 3 (except for 2xFYVE, n = 2).

C. Volcano plots comparing proximal proteome of each location specific construct to the control (NES). Proteins with a $\log_2$ fold change >1 and adjusted p-value <0.05 were considered significant and colored. Data were collected from three independent biological replicates (except 2xFYVE), n = 3 (except for 2xFYVE, n = 2), statistical analysis was performed using the statistical framework implemented in MSstats (Choi *et al*, 2014), p-values were adjusted using Benjamini-Hochberg.(Benjamini & Hochberg, 1995)

D. Venn diagram comparing the location specific proteins.

E. Heatmap of location-specific proteins was highlighted as three groups based on the median normalized $\log_2$ fold change. Gene ontology (GO) enrichment analysis was performed for each cluster and the matched genes for each GO term are indicated in dark purple, n = 3 (except for 2xFYVE, n = 2).

F. Comparison of the location proteins candidates identified in this study with the protein localizations defined in Human Cell Map (Go *et al*, 2021). For the analysis we used the following number of proteins representing the different subcellular compartments: 700 for endosome/lysosome, 543 for Golgi apparatus, 1352 for plasma membrane, and 953 for cell junction.

***Figure 5. Mapping the activity-dependent proximal interaction network of the $5HT_{2A}$ receptor.***

A. Experimental design to map proximal interaction networks of $5HT_{2A}$. APEX2-fused $5HT_{2A}$ receptor was stably expressed in HEK293T cells, activated with 5HT and at indicated time points after agonist treatment proteins proximal of the receptor were biotinylated by addition of $H_2O_2$. Following automated enrichment of biotinylated proteins, the samples were analyzed using DIA.

B. Heatmap depicting proteins that significantly change in biotinylation after agonist addition. Overrepresentation of location-specific proteins for plasma membrane (Lyn 11), endosome (2xFYVE), late endosome/lysosome (LAMTOR1), and Golgi (GalT) was shown in the right heartmap. Data were collected from three independent biological replicates, n = 3, statistical analysis was performed using ANOVA.

C. $Log_2$ fold change distribution of proteins in Figure B (Control: Grey). Endosomal/lysosomal proteins (2xFYVE/LAMTOR1) and plasma membrane proteins (Lyn11) were highlighted in orange and yellow, respectively. The time point of each protein used for analysis is corresponding to the maximum $log_2$ fold change. The intensity at each timepoint was included in **Dataset EV5**.

D. Line charts of proteins filtered out from Figure C with a fold change cutoff of 1.5 clustered into proteins with transient and sustained responses to 5HT treatment. 40 and 102 proteins are grouped into transient and sustained clusters, respectively.

E. Protein interaction network connecting proteins with transient agonist-dependent changes in the proximity of $5HT_{2A}$. Proteins shown as nodes were colored according to their $log_2$ fold change. The edges connecting the proteins were derived from STRING (Szklarczyk *et al*, 2011) and the edge width was scaled according to the interaction confidence. Proteins corresponding to Reactome pathways (Jassal *et al*, 2020) that were enriched within the transient cluster were indicated with different colored node borders.

***Figure 6. Exploring proximal interaction network dynamics induced by CRISPR-based gene knockout using low-input PL.***

A. Schematic of PL with CRISPR-based gene knockout (KO). Polyclonal knockout cell lines were generated by electroporation of Cas9 ribonucleoprotein complexes into HEK293T cells expressing the APEX2-fused $5HT_{2A}$ receptor. The KO cells were seeded in a 6-well plate for proximity proteomics with receptor activation by 5HT for 0, 1, and 30 min.

B. Performance comparison of $5HT_{2A}$ PL in 10cm dish and 6-well plate format based on feature and protein identification. Two independent non-targeting controls (NTC) were used for the 6-well plate format. All numbers in the figures represent the sum of identifications of feature and protein level across three replicates, n = 3.

C. Distribution of the coefficient of variation (CV) of protein intensities from three replicates comparing $5HT_{2A}$ PL in 10cm dish and 6-well plate format. Two independent non-targeting controls (NTC) were used for the 6-well plate format. Numbers depicted on the graph represent median CVs. Data were collected from three independent biological replicates, n = 3.

D. Knockout validation for S100A10 by Inference of CRISPR Edits (ICE).

E. Knockout validation for ANXA2 by Western blot analysis.

F. Heatmap depicting proteins from 5HT2A-APEX2 experiment with transient responses to 5HT treatment (**Figure 5D**) compared across NTCs and knockouts of S100A2 and ANXA2 in 6-well plate format and the 5HT2A APEX2 data from 10cm dishes. Data were collected from three independent biological replicates (n = 3) for NTCs and S100A10 and two independent biological replicates (n = 2) for ANXA2.

G. Temporal profile for regulated proteins in $5HT_{2A}$ receptor proximal interaction network in KO vs. control (non-targeting) cells. Line charts represent the log2 fold change over the

time course of receptor activation with 0.1 μM 5HT. Red: protein AHNAK, Green: protein ANXA2, Blue: protein S100A10, Solid circle: KO using gRNA 1, Solid triangle: KO using gRNA 2. Data were collected from three independent biological replicates (n = 3) for NTCs and S100A10 and two independent biological replicates (n = 2) for ANXA2.

## References


Adachi N, Hess DT, McLaughlin P & Stamler JS (2016) S-Palmitoylation of a Novel Site in the β2-Adrenergic Receptor Associated with a Novel Intracellular Itinerary. *J Biol Chem* 291: 20232–20246

Aebersold R & Mann M (2003) Mass spectrometry-based proteomics. *Nature* 422: 198–207

Allen JA, Yadav PN & Roth BL (2008) Insights into the regulation of 5-HT2A serotonin receptors by scaffolding proteins and kinases. *Neuropharmacology* 55: 961–968

Antonicka H, Lin Z-Y, Janer A, Aaltonen MJ, Weraarpachai W, Gingras A-C & Shoubridge EA (2020) A High-Density Human Mitochondrial Proximity Interaction Network. *Cell Metab* 32: 479–497.e9

Benaud C, Gentil BJ, Assard N, Court M, Garin J, Delphin C & Baudier J (2004) AHNAK interaction with the annexin 2/S100A10 complex regulates cell membrane cytoarchitecture. *J Cell Biol* 164: 133–144

Benjamini Y & Hochberg Y (1995) Controlling the false discovery rate: A practical and powerful approach to multiple testing. *J R Stat Soc* 57: 289–300

Bhatnagar A, Willins DL, Gray JA, Woods J, Benovic JL & Roth BL (2001) The dynamin-dependent, arrestin-independent internalization of 5-hydroxytryptamine 2A (5-HT2A) serotonin receptors reveals differential sorting of arrestins and 5-HT2A receptors during endocytosis. *J Biol Chem* 276: 8269–8277

Bowers K & Stevens TH (2005) Protein transport from the late Golgi to the vacuole in the yeast Saccharomyces cerevisiae. *Biochim Biophys Acta* 1744: 438–454

Branon TC, Bosch JA, Sanchez AD, Udeshi ND, Svinkina T, Carr SA, Feldman JL, Perrimon N & Ting AY (2018) Efficient proximity labeling in living cells and organisms with TurboID. *Nat Biotechnol* 36: 880–887

Chen EY, Tan CM, Kou Y, Duan Q, Wang Z, Meirelles GV, Clark NR & Ma'ayan A (2013) Enrichr: interactive and collaborative HTML5 gene list enrichment analysis tool. *BMC Bioinformatics* 14: 128

Chen MX, Oh Y-S & Kim Y (2022) S100A10 and its binding partners in depression and antidepressant actions. *Front Mol Neurosci* 15: 953066

Cho H & Kehrl JH (2007) Localization of Gi alpha proteins in the centrosomes and at the midbody: implication for their role in cell division. *J Cell Biol* 178: 245–255

Choi M, Chang C-Y, Clough T, Broudy D, Killeen T, MacLean B & Vitek O (2014) MSstats: an R package for statistical analysis of quantitative mass spectrometry-based proteomic experiments. *Bioinformatics* 30: 2524–2526

Collins BC, Hunter CL, Liu Y, Schilling B, Rosenberger G, Bader SL, Chan DW, Gibson BW, Gingras A-C, Held JM, *et al* (2017) Multi-laboratory assessment of reproducibility, qualitative and quantitative performance of SWATH-mass spectrometry. *Nat Commun* 8: 291

Deutsch EW, Bandeira N, Perez-Riverol Y, Sharma V, Carver JJ, Mendoza L, Kundu DJ, Wang S, Bandla C, Kamatchinathan S, *et al* (2023) The ProteomeXchange consortium at 10 years: 2023 update. *Nucleic Acids Res* 51: D1539–D1548

DiBerto JF, Olsen RHJ & Roth BL (2022) TRUPATH: An Open-Source Biosensor Platform for Interrogating the GPCR Transducerome. *Bioluminescence*: 185–195

Eisenberg-Bord M, Shai N, Schuldiner M & Bohnert M (2016) A Tether Is a Tether Is a Tether: Tethering at Membrane Contact Sites. *Developmental Cell* 39: 395–409

Fu Q, Kowalski MP, Mastali M, Parker SJ, Sobhani K, van den Broek I, Hunter CL & Van Eyk JE (2018) Highly Reproducible Automated Proteomics Sample Preparation Workflow for Quantitative Mass Spectrometry. *J Proteome Res* 17: 420–428

Gao XD, Tu L-C, Mir A, Rodriguez T, Ding Y, Leszyk J, Dekker J, Shaffer SA, Zhu LJ, Wolfe SA, *et al* (2018) C-BERST: defining subnuclear proteomic landscapes at genomic elements with dCas9-APEX2. *Nat Methods* 15: 433–436

Gillespie M, Jassal B, Stephan R, Milacic M, Rothfels K, Senff-Ribeiro A, Griss J, Sevilla C, Matthews L, Gong C, *et al* (2022) The reactome pathway knowledgebase 2022. *Nucleic Acids Res* 50: D687–D692

Gillet LC, Navarro P, Tate S, Röst H, Selevsek N, Reiter L, Bonner R & Aebersold R (2012) Targeted data extraction of the MS/MS spectra generated by data-independent acquisition: a new concept for consistent and accurate proteome analysis. *Mol Cell Proteomics* 11: O111.016717

Go CD, Knight JDR, Rajasekharan A, Rathod B, Hesketh GG, Abe KT, Youn J-Y, Samavarchi-Tehrani P, Zhang H, Zhu LY, *et al* (2021) A proximity-dependent biotinylation map of a human cell. *Nature* 595: 120–124

Göös H, Kinnunen M, Salokas K, Tan Z, Liu X, Yadav L, Zhang Q, Wei G-H & Varjosalo M (2022) Human transcription factor protein interaction networks. *Nat Commun* 13: 766

Gupta R, Somyajit K, Narita T, Maskey E, Stanlie A, Kremer M, Typas D, Lammers M, Mailand N, Nussenzweig A, *et al* (2018) DNA Repair Network Analysis Reveals Shieldin as a Key Regulator of NHEJ and PARP Inhibitor Sensitivity. *Cell* 173: 972–988.e23

Hanson BJ & Hong W (2003) Evidence for a role of SNX16 in regulating traffic between the early and later endosomal compartments. *J Biol Chem* 278: 34617–34630

He K-L, Deora AB, Xiong H, Ling Q, Weksler BB, Niesvizky R & Hajjar KA (2008) Endothelial cell annexin A2 regulates polyubiquitination and degradation of its binding partner S100A10/p11. *J Biol Chem* 283: 19192–19200

Hu A, Noble WS & Wolf-Yadlin A (2016) Technical advances in proteomics: new developments in data-independent acquisition. *F1000Res* 5

Jassal B, Matthews L, Viteri G, Gong C, Lorente P, Fabregat A, Sidiropoulos K, Cook J, Gillespie M, Haw R, *et al* (2020) The reactome pathway knowledgebase. *Nucleic Acids Res* 48: D498–D503

Jin J, Bhatti DL, Lee K-W, Medrihan L, Cheng J, Wei J, Zhong P, Yan Z, Kooiker C, Song C, *et al* (2020) Ahnak scaffolds p11/Anxa2 complex and L-type voltage-gated calcium channel and modulates depressive behavior. *Mol Psychiatry* 25: 1035–1049

Kim K, Che T, Panova O, DiBerto JF, Lyu J, Krumm BE, Wacker D, Robertson MJ, Seven AB, Nichols DE, *et al* (2020) Structure of a Hallucinogen-Activated Gq-Coupled 5-HT2A Serotonin Receptor. *Cell* 182: 1574–1588.e19

Kubitz L, Bitsch S, Zhao X, Schmitt K, Deweid L, Roehrig A, Barazzone EC, Valerius O, Kolmar H & Béthune J (2022) Engineering of ultraID, a compact and hyperactive enzyme for proximity-dependent biotinylation in living cells. *Commun Biol* 5: 657

Kuleshov MV, Jones MR, Rouillard AD, Fernandez NF, Duan Q, Wang Z, Koplev S, Jenkins SL, Jagodnik KM, Lachmann A, *et al* (2016) Enrichr: a comprehensive gene set enrichment analysis web server 2016 update. *Nucleic Acids Res* 44: W90–7

Kwan AC, Olson DE, Preller KH & Roth BL (2022) The neural basis of psychedelic action. *Nat Neurosci* 25: 1407–1419

Lam SS, Martell JD, Kamer KJ, Deerinck TJ, Ellisman MH, Mootha VK & Ting AY (2015) Directed evolution of APEX2 for electron microscopy and proximity labeling. *Nat Methods* 12: 51–54

Lee K-W, Westin L, Kim J, Chang JC, Oh Y-S, Amreen B, Gresack J, Flajolet M, Kim D, Aperia A, *et al* (2015) Alteration by p11 of mGluR5 localization regulates depression-like behaviors. *Mol Psychiatry* 20: 1546–1556

Liu X, Gygi SP & Paulo JA (2021) A Semiautomated Paramagnetic Bead-Based Platform for Isobaric Tag Sample Preparation. *J Am Soc Mass Spectrom* 32: 1519–1529

Li X-W, Rees JS, Xue P, Zhang H, Hamaia SW, Sanderson B, Funk PE, Farndale RW, Lilley KS, Perrett S, *et al* (2014) New insights into the DT40 B cell receptor cluster using a proteomic proximity labeling assay. *J Biol Chem* 289: 14434–14447

Lobingier BT, Hüttenhain R, Eichel K, Miller KB, Ting AY, von Zastrow M & Krogan NJ (2017) An Approach to Spatiotemporally Resolve Protein Interaction Networks in Living Cells. *Cell* 169: 350–360.e12

Ludwig C, Gillet L, Rosenberger G, Amon S, Collins BC & Aebersold R (2018) Data-independent acquisition-based SWATH-MS for quantitative proteomics: a tutorial. *Mol Syst Biol* 14: e8126

Martell JD, Yamagata M, Deerinck TJ, Phan S, Kwa CG, Ellisman MH, Sanes JR & Ting AY (2016) A split horseradish peroxidase for the detection of intercellular protein-protein interactions and sensitive visualization of synapses. *Nat Biotechnol* 34: 774–780

McClure-Begley TD & Roth BL (2022) The promises and perils of psychedelic pharmacology for psychiatry. *Nat Rev Drug Discov* 21: 463–473

McMahon HT & Boucrot E (2011) Molecular mechanism and physiological functions of clathrin-mediated endocytosis. *Nat Rev Mol Cell Biol* 12: 517–533

Mettlen M, Chen P-H, Srinivasan S, Danuser G & Schmid SL (2018) Regulation of Clathrin-Mediated Endocytosis. *Annu Rev Biochem* 87: 871–896

Michalski A, Cox J & Mann M (2011) More than 100,000 Detectable Peptide Species Elute in Single Shotgun Proteomics Runs but the Majority is Inaccessible to Data-Dependent LC−MS/MS. *Journal of Proteome Research* 10: 1785–1793

Müller T, Kalxdorf M, Longuespée R, Kazdal DN, Stenzinger A & Krijgsveld J (2020) Automated sample preparation with SP3 for low-input clinical proteomics. *Mol Syst Biol* 16: e9111

Myers SA, Wright J, Peckner R, Kalish BT, Zhang F & Carr SA (2018) Discovery of proteins associated with a predefined genomic locus via dCas9-APEX-mediated proximity labeling. *Nat Methods* 15: 437–439

Nag S, Rani S, Mahanty S, Bissig C, Arora P, Azevedo C, Saiardi A, van der Sluijs P, Delevoye C, van Niel G, *et al* (2018) Rab4A organizes endosomal domains for sorting cargo to lysosome-related organelles. *J Cell Sci* 131

Olsen RHJ, DiBerto JF, English JG, Glaudin AM, Krumm BE, Slocum ST, Che T, Gavin AC, McCorvy JD, Roth BL, *et al* (2020) TRUPATH, an open-source biosensor platform for interrogating the GPCR transducerome. *Nat Chem Biol* 16: 841–849

Olsten MEK, Canton DA, Zhang C, Walton PA & Litchfield DW (2004) The Pleckstrin homology domain of CK2 interacting protein-1 is required for interactions and recruitment of protein kinase CK2 to the plasma membrane. *J Biol Chem* 279: 42114–42127

Paek J, Kalocsay M, Staus DP, Wingler L, Pascolutti R, Paulo JA, Gygi SP & Kruse AC (2017) Multidimensional Tracking of GPCR Signaling via Peroxidase-Catalyzed Proximity Labeling. *Cell* 169: 338–349.e11

Piette BL, Alerasool N, Lin Z-Y, Lacoste J, Lam MHY, Qian WW, Tran S, Larsen B, Campos E, Peng J, *et al* (2021) Comprehensive interactome profiling of the human Hsp70 network highlights functional differentiation of J domains. *Mol Cell* 81: 2549–2565.e8

Polacco BJ, Lobingier BT, Blythe EE, Abreu N, Khare P, Howard MK, Gonzalez-Hernandez AJ, Xu J, Li Q, Novy B, *et al* (2024) Profiling the proximal proteome of the activated μ-opioid receptor. *Nat Chem Biol*

Qin W, Cho KF, Cavanagh PE & Ting AY (2021a) Deciphering molecular interactions by proximity labeling. *Nat Methods* 18: 133–143

Qin W, Myers SA, Carey DK, Carr SA & Ting AY (2021b) Spatiotemporally-resolved mapping of RNA binding proteins via functional proximity labeling reveals a mitochondrial mRNA anchor promoting stress recovery. *Nat Commun* 12: 4980

Ramanathan M, Majzoub K, Rao DS, Neela PH, Zarnegar BJ, Mondal S, Roth JG, Gai H, Kovalski JR, Siprashvili Z, *et al* (2018) RNA-protein interaction detection in living cells. *Nat Methods* 15: 207–212

Réty S, Sopkova J, Renouard M, Osterloh D, Gerke V, Tabaries S, Russo-Marie F & Lewit-Bentley A (1999) The crystal structure of a complex of p11 with the annexin II N-terminal peptide. *Nat Struct Biol* 6: 89–95

Rhee H-W, Zou P, Udeshi ND, Martell JD, Mootha VK, Carr SA & Ting AY (2013) Proteomic mapping of mitochondria in living cells via spatially restricted enzymatic tagging. *Science* 339: 1328–1331

Roth BL & Chuang DM (1987) Multiple mechanisms of serotonergic signal transduction. *Life Sci* 41: 1051–1064

Roux KJ, Kim DI, Raida M & Burke B (2012) A promiscuous biotin ligase fusion protein identifies proximal and interacting proteins in mammalian cells. *J Cell Biol* 196: 801–810

Samavarchi-Tehrani P, Abdouni H, Samson R & Gingras A-C (2018) A Versatile Lentiviral Delivery Toolkit for Proximity-dependent Biotinylation in Diverse Cell Types. *Mol Cell Proteomics* 17: 2256–2269

Samavarchi-Tehrani P, Samson R & Gingras A-C (2020) Proximity Dependent Biotinylation: Key Enzymes and Adaptation to Proteomics Approaches. *Mol Cell Proteomics* 19: 757–773

Schindelin J, Arganda-Carreras I, Frise E, Kaynig V, Longair M, Pietzsch T, Preibisch S, Rueden C, Saalfeld S, Schmid B, *et al* (2012) Fiji: an open-source platform for biological-image analysis. *Nat Methods* 9: 676–682

Shannon P, Markiel A, Ozier O, Baliga NS, Wang JT, Ramage D, Amin N, Schwikowski B & Ideker T (2003) Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome Res* 13: 2498–2504

Short B, Preisinger C, Körner R, Kopajtich R, Byron O & Barr FA (2001) A GRASP55-rab2 effector complex linking Golgi structure to membrane traffic. *J Cell Biol* 155: 877–883

Sit S-T & Manser E (2011) Rho GTPases and their role in organizing the actin cytoskeleton. *J Cell Sci* 124: 679–683

Svenningsson P, Chergui K, Rachleff I, Flajolet M, Zhang X, El Yacoubi M, Vaugeois J-M, Nomikos GG & Greengard P (2006) Alterations in 5-HT1B receptor function by p11 in depression-like states. *Science* 311: 77–80

Szklarczyk D, Franceschini A, Kuhn M, Simonovic M, Roth A, Minguez P, Doerks T, Stark M, Muller J, Bork P, *et al* (2011) The STRING database in 2011: functional interaction networks of proteins, globally integrated and scored. *Nucleic Acids Res* 39: D561–8

Tsai T-H, Choi M, Banfai B, Liu Y, MacLean BX, Dunkley T & Vitek O (2020) Selection of Features with Consistent Profiles Improves Relative Protein Quantification in Mass Spectrometry Experiments. *Mol Cell Proteomics* 19: 944–959

Vandemoortele G, De Sutter D, Moliere A, Pauwels J, Gevaert K & Eyckerman S (2019) A Well-Controlled BioID Design for Endogenous Bait Proteins. *J Proteome Res* 18: 95–106

Villani M, Subathra M, Im Y-B, Choi Y, Signorelli P, Del Poeta M & Luberto C (2008) Sphingomyelin synthases regulate production of diacylglycerol at the Golgi. *Biochem J* 414: 31–41

Wang Y, MacDonald RG, Thinakaran G & Kar S (2016) Insulin-Like Growth Factor-II/Cation-Independent Mannose 6-Phosphate Receptor in Neurodegenerative Diseases. *Mol*

*Neurobiol* 54: 2636–2658

Warner-Schmidt JL, Flajolet M, Maller A, Chen EY, Qi H, Svenningsson P & Greengard P (2009) Role of p11 in cellular and behavioral effects of 5-HT4 receptor stimulation. *J Neurosci* 29: 1937–1946

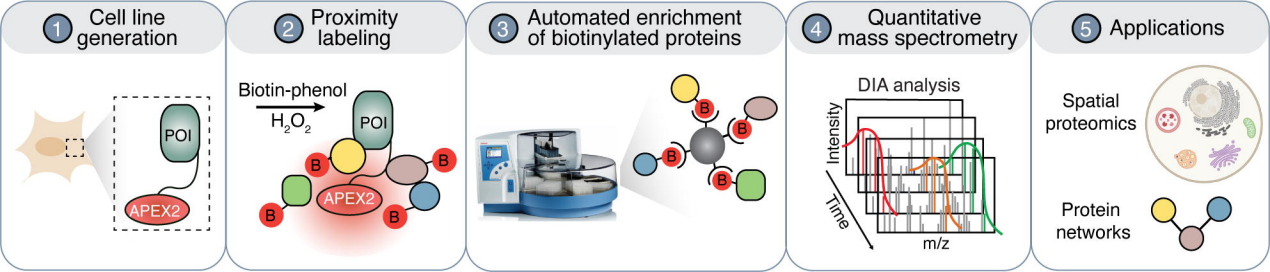
Xia Z, Gray JA, Compton-Toth BA & Roth BL (2003) A direct interaction of PSD-95 with 5-HT2A serotonin receptors regulates receptor trafficking and signal transduction. *J Biol Chem* 278: 21901–21908

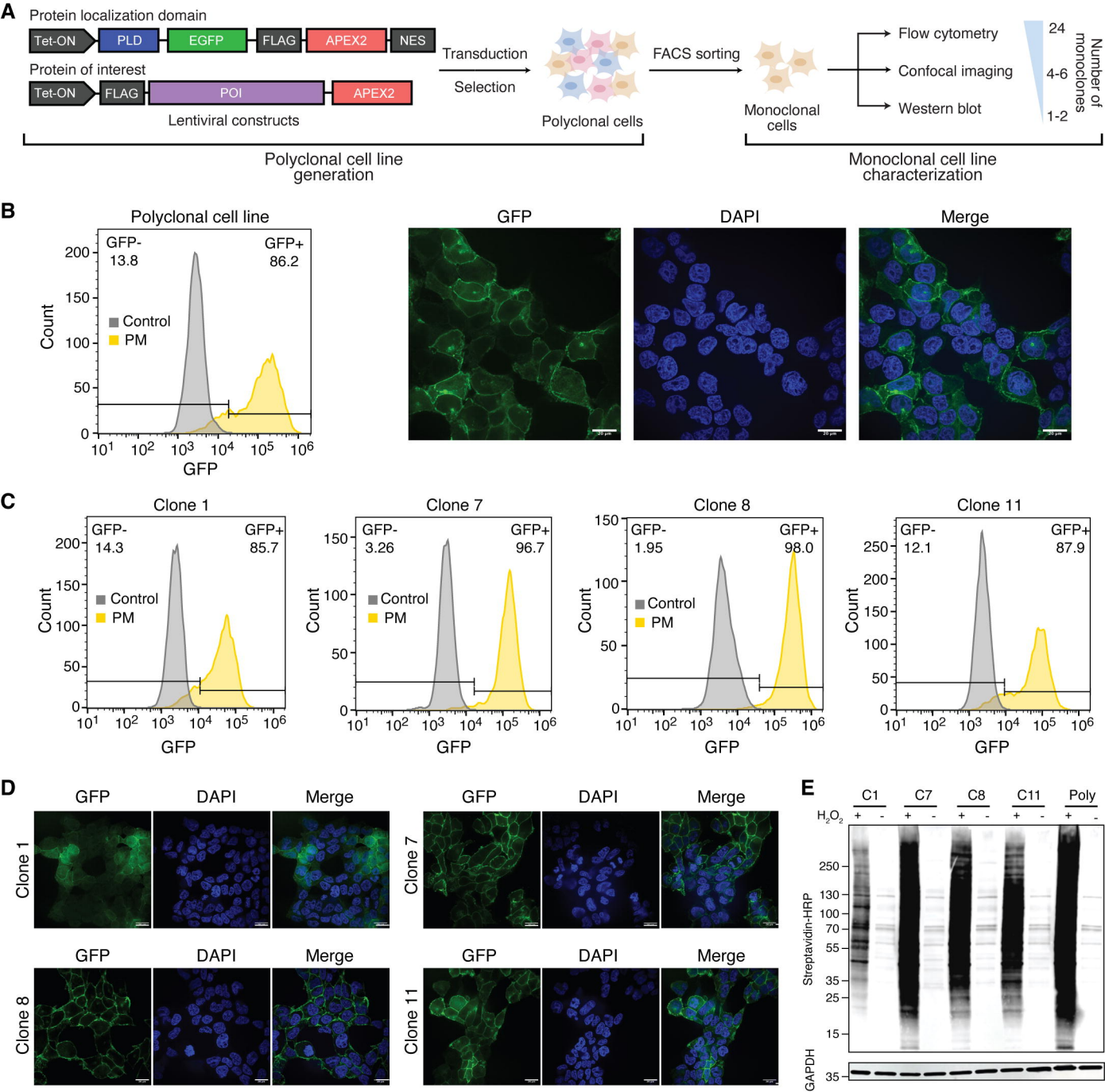
Xie Z, Bailey A, Kuleshov MV, Clarke DJB, Evangelista JE, Jenkins SL, Lachmann A, Wojciechowicz ML, Kropiwnicki E, Jagodnik KM, *et al* (2021) Gene Set Knowledge Discovery with Enrichr. *Curr Protoc* 1: e90

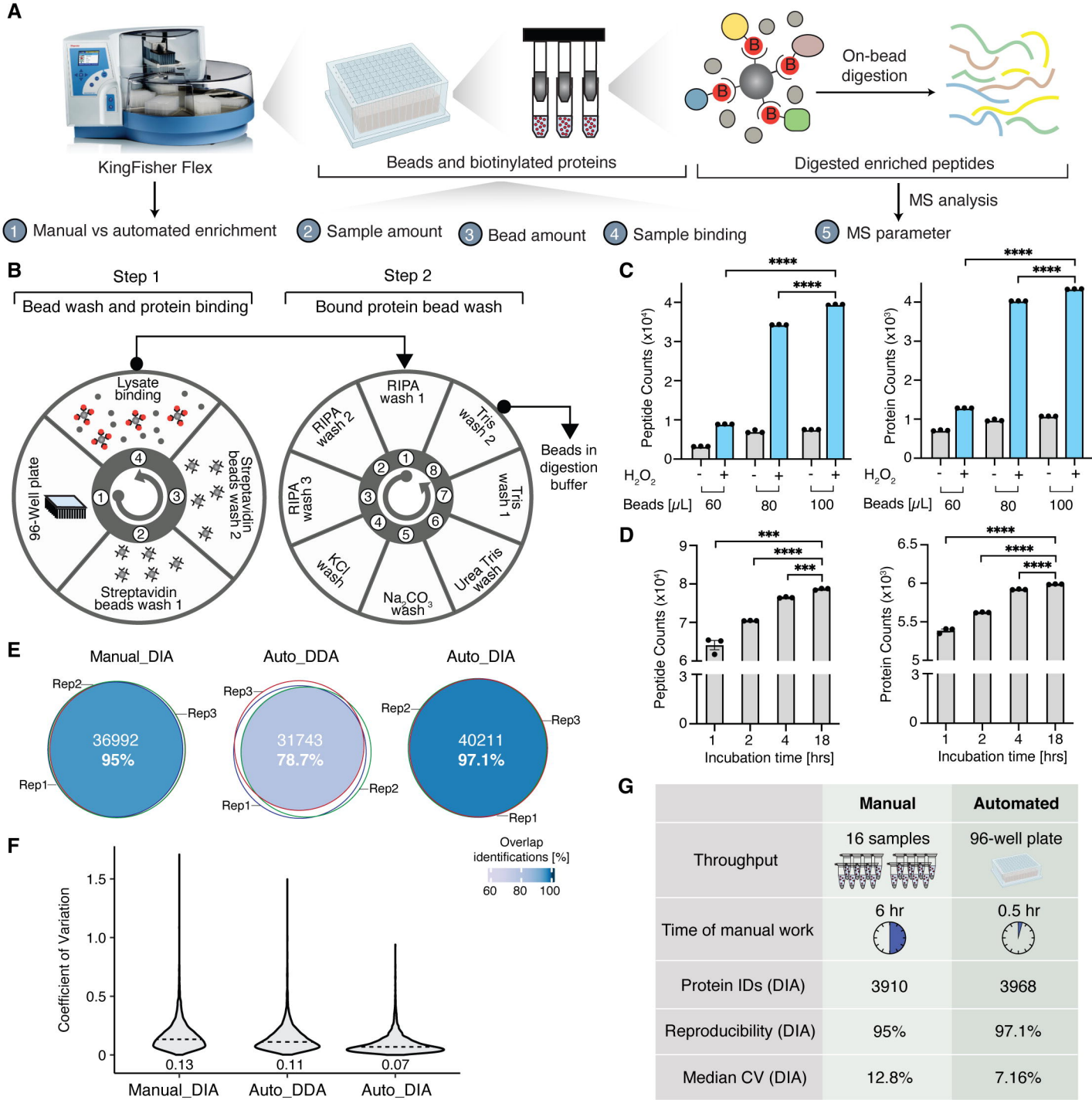
Xu J, Xu T, Wu B, Ye Y, You X, Shu X, Pei D & Liu J (2013) Structure of sorting nexin 11 (SNX11) reveals a novel extended phox homology (PX) domain critical for inhibition of SNX10-induced vacuolation. *J Biol Chem* 288: 16598–16605

Yamamoto H, Tobisawa Y, Inubushi T, Irie F, Ohyama C & Yamaguchi Y (2017) A mammalian homolog of the zebrafish transmembrane protein 2 (TMEM2) is the long-sought-after cell-surface hyaluronidase. *J Biol Chem* 292: 7304–7313

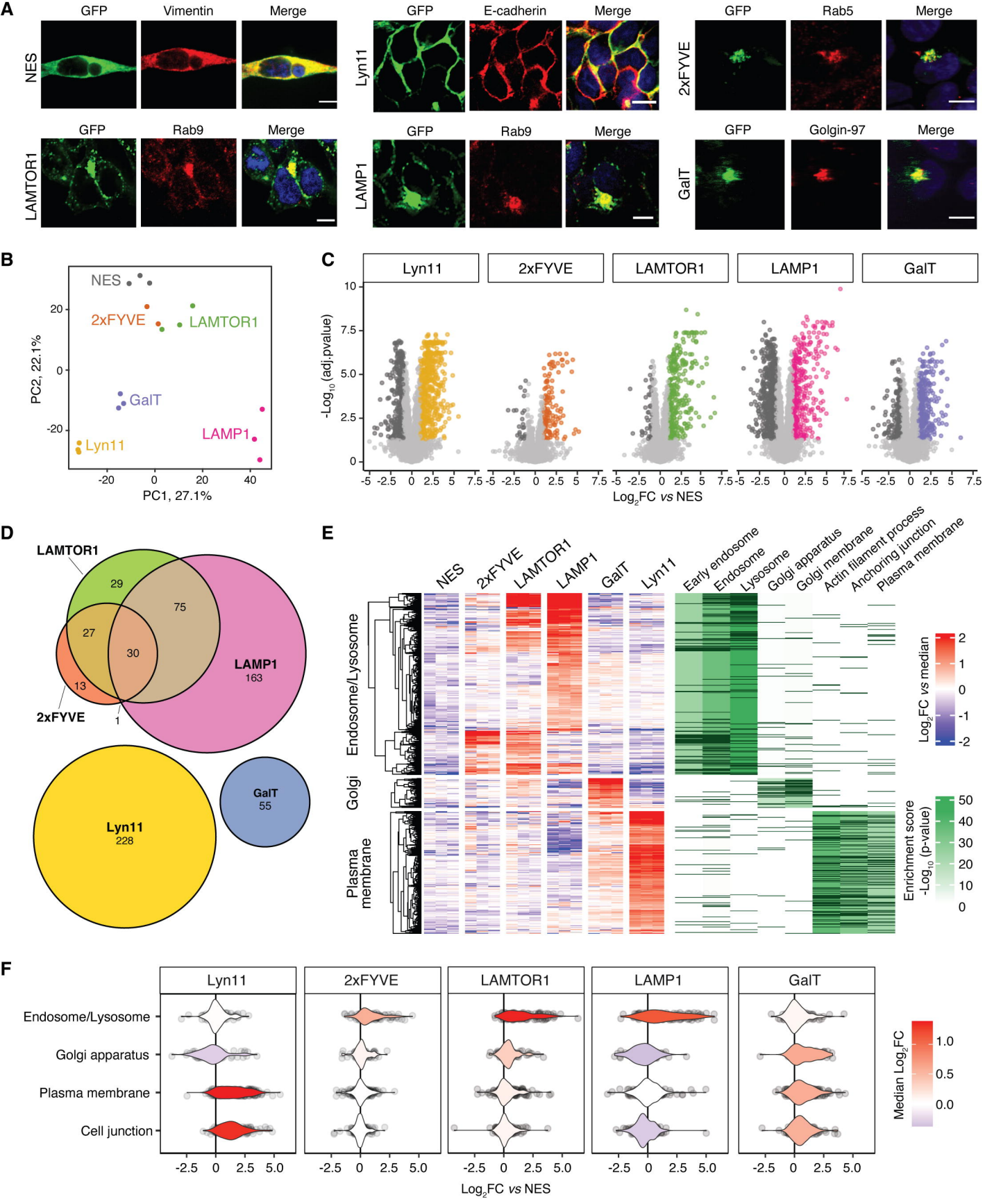
Yang Y, Wang M, Zhang Y-Y, Zhao S-Z & Gu S (2022) The endosomal sorting complex required for transport repairs the membrane to delay cell death. *Front Oncol* 12: 1007446

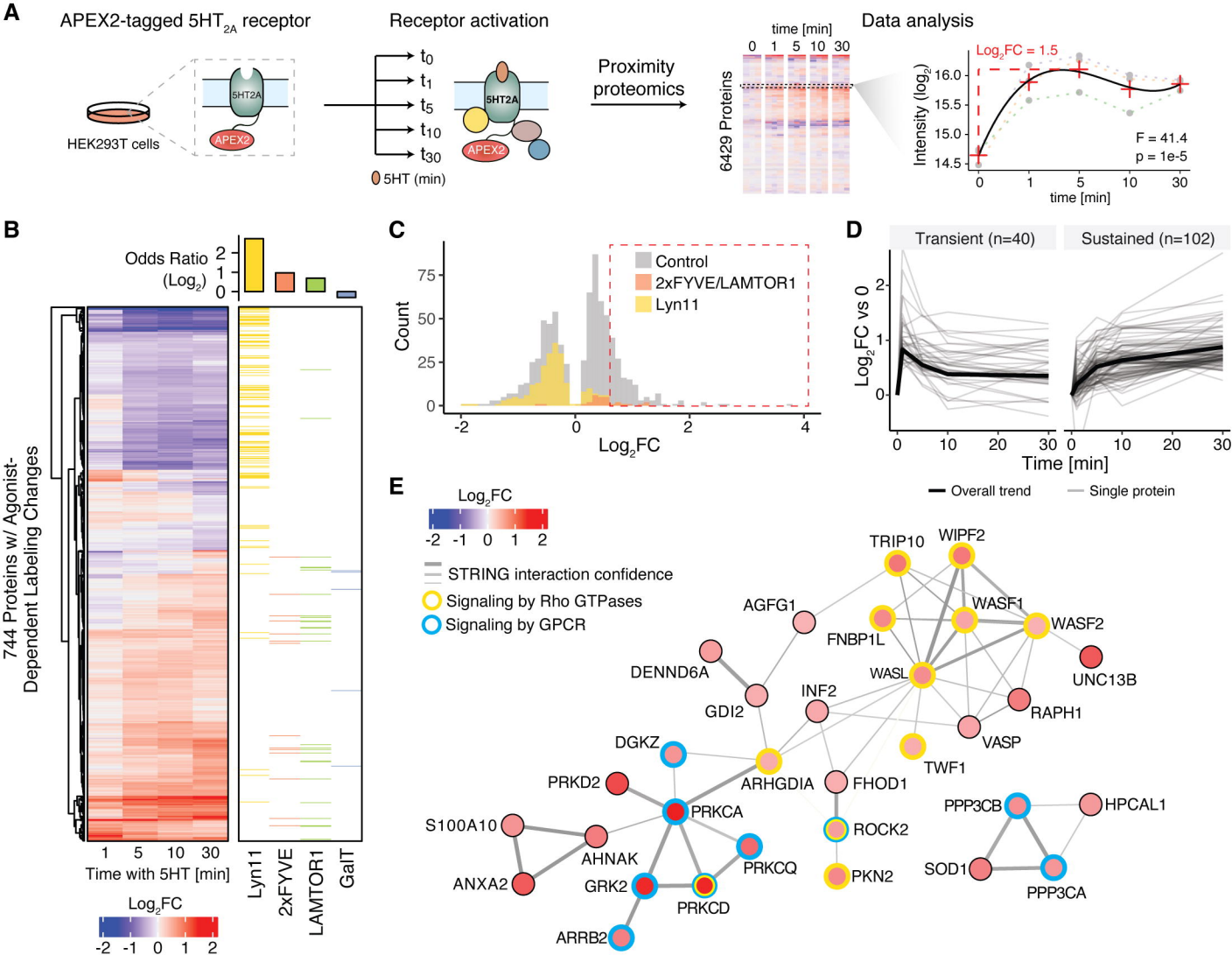

1 Cell line generation
POI
APEX2
2 Proximity labeling
Biotin-phenol
$H_2O_2$
POI
APEX2
B
3 Automated enrichment of biotinylated proteins
B
4 Quantitative mass spectrometry
DIA analysis
Intensity
Time
m/z
5 Applications
Spatial proteomics
Protein networks

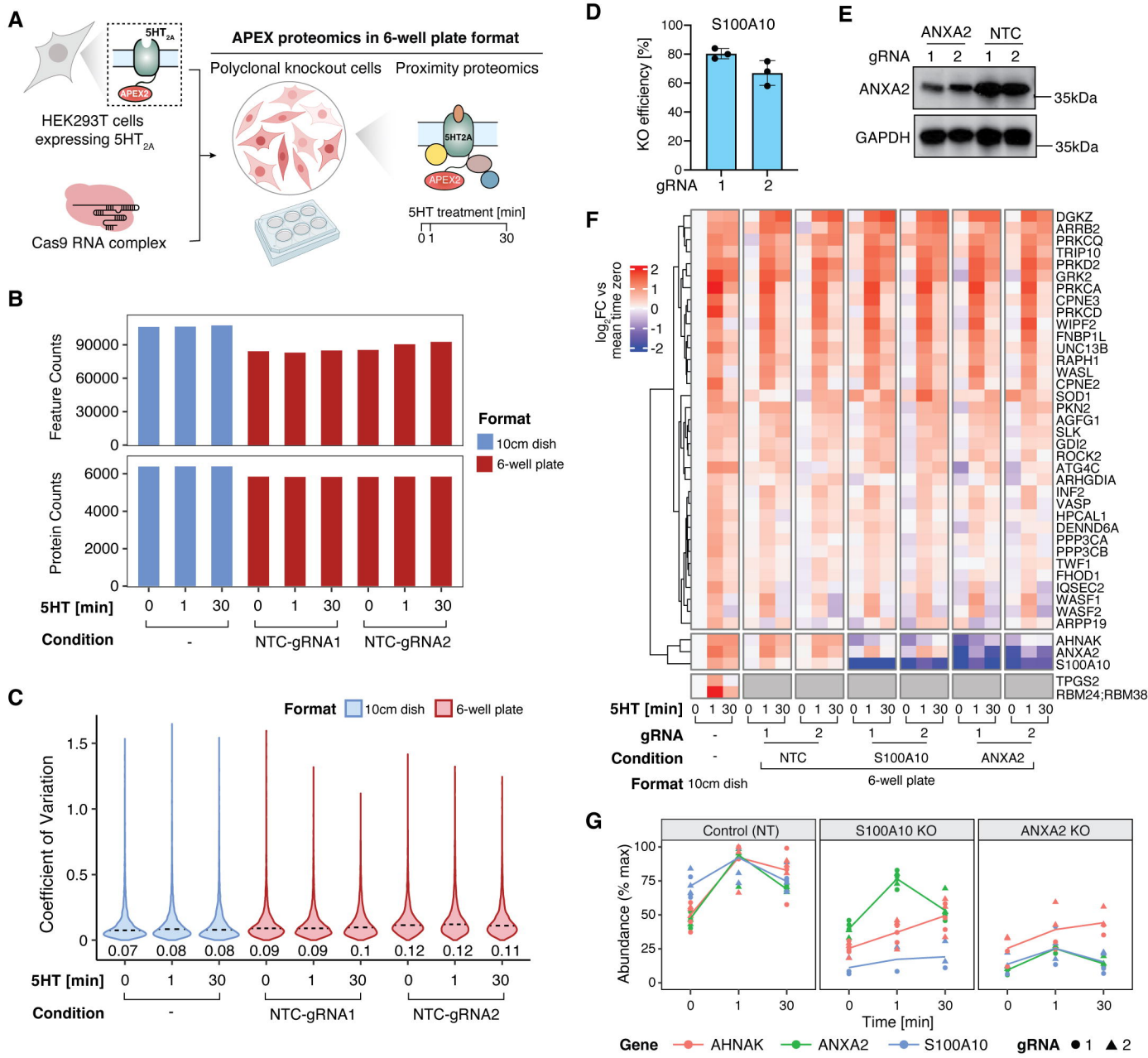

A
Protein localization domain
Tet-ON
PLD
EGFP
FLAG
APEX2
NES
Protein of interest
Tet-ON
FLAG
POI
APEX2
Lentiviral constructs
Transduction
Selection
Polyclonal cells
FACS sorting
Monoclonal cells
Flow cytometry
Confocal imaging
Western blot
24
4-6
1-2
Number of monoclones
Polyclonal cell line generation
Monoclonal cell line characterization
B
Polyclonal cell line
GFP-
13.8
GFP+
86.2
Control
PM
Count
GFP
GFP
DAPI
Merge
C
Clone 1
GFP-
14.3
GFP+
85.7
Clone 7
GFP-
3.26
GFP+
96.7
Clone 8
GFP-
1.95
GFP+
98.0
Clone 11
GFP-
12.1
GFP+
87.9
D
GFP
DAPI
Merge
Clone 1
Clone 7
Clone 8
Clone 11
E
C1
C7
C8
C11
Poly
H2O2
Streptavidin-HRP
GAPDH




| | Manual | Automated |
|---|---|---|
| Throughput | 16 samples | 96-well plate |
| Time of manual work | 6 hr | 0.5 hr |
| Protein IDs (DIA) | 3910 | 3968 |
| Reproducibility (DIA) | 95% | 97.1% |
| Median CV (DIA) | 12.8% | 7.16% |


A
GFP
Vimentin
Merge
NES
E-cadherin
Lyn11
Rab5
2xFYVE
Rab9
LAMTOR1
LAMP1
Golgin-97
GalT
B
NES
2xFYVE
LAMTOR1
GalT
LAMP1
Lyn11
PC2, 22.1%
PC1, 27.1%
C
Lyn11
2xFYVE
LAMTOR1
LAMP1
GalT
-Log10 (adj.pvalue)
Log2FC vs NES
D
LAMTOR1
29
75
27
30
13
1
LAMP1
163
2xFYVE
Lyn11
228
GalT
55
E
NES
2xFYVE
LAMTOR1
LAMP1
GalT
Lyn11
Early endosome
Endosome
Lysosome
Golgi apparatus
Golgi membrane
Actin filament process
Anchoring junction
Plasma membrane
Endosome/Lysosome
Golgi
Plasma membrane
Log2FC vs median
Enrichment score
-Log10 (p-value)
F
Lyn11
2xFYVE
LAMTOR1
LAMP1
GalT
Endosome/Lysosome
Golgi apparatus
Plasma membrane
Cell junction
Log2FC vs NES
Median Log2FC


A
APEX2-tagged 5HT2A receptor
HEK293T cells
5HT2A
APEX2
Receptor activation
t0
t1
t5
t10
t30
5HT (min)
Proximity proteomics
time [min]
0 1 5 10 30
6429 Proteins
Data analysis
Log2FC = 1.5
Intensity (log2)
F = 41.4
p = 1e-5
time [min]
B
Odds Ratio (Log2)
744 Proteins w/ Agonist-Dependent Labeling Changes
Time with 5HT [min]
Lyn11
2xFYVE
LAMTOR1
GalT
Log2FC
C
Count
Control
2xFYVE/LAMTOR1
Lyn11
Log2FC
D
Transient (n=40)
Sustained (n=102)
Log2FC vs 0
Time [min]
Overall trend
Single protein
E
Log2FC
STRING interaction confidence
Signaling by Rho GTPases
Signaling by GPCR
TRIP10
WIPF2
AGFG1
WASF1
WASF2
FNBP1L
DENND6A
WASL
UNC13B
INF2
GDI2
RAPH1
VASP
DGKZ
TWF1
ARHGDIA
FHOD1
PRKD2
PRKCA
PPP3CB
HPCAL1
S100A10
ROCK2
AHNAK
PRKCQ
SOD1
ANXA2
GRK2
PKN2
PRKCD
PPP3CA
ARRB2


A
APEX proteomics in 6-well plate format
Polyclonal knockout cells
Proximity proteomics
5HT2A
APEX2
HEK293T cells expressing 5HT2A
Cas9 RNA complex
5HT2A
APEX2
5HT treatment [min]
0 1 30
B
Feature Counts
Protein Counts
Format
10cm dish
6-well plate
5HT [min]
Condition
-
NTC-gRNA1
NTC-gRNA2
C
Coefficient of Variation
Format 10cm dish 6-well plate
0.07 0.08 0.08 0.09 0.09 0.1 0.12 0.12 0.11
5HT [min]
Condition
-
NTC-gRNA1
NTC-gRNA2
D
S100A10
KO efficiency [%]
gRNA 1 2
E
ANXA2
NTC
gRNA 1 2 1 2
ANXA2
GAPDH
35kDa
35kDa
F
log2FC vs mean time zero
DGKZ
ARRB2
PRKCQ
TRIP10
PRKD2
GRK2
PRKCA
CPNE3
PRKCD
WIPF2
FNBP1L
UNC13B
RAPH1
WASL
CPNE2
SOD1
PKN2
AGFG1
SLK
GDI2
ROCK2
ATG4C
ARHGDIA
INF2
VASP
HPCAL1
DENND6A
PPP3CA
PPP3CB
TWF1
FHOD1
IQSEC2
WASF1
WASF2
ARPP19
AHNAK
ANXA2
S100A10
TPGS2
RBM24;RBM38
5HT [min]
gRNA
Condition
NTC
S100A10
ANXA2
Format 10cm dish
6-well plate
G
Control (NT)
S100A10 KO
ANXA2 KO
Abundance (% max)
Time [min]
Gene AHNAK ANXA2 S100A10
gRNA 1 2