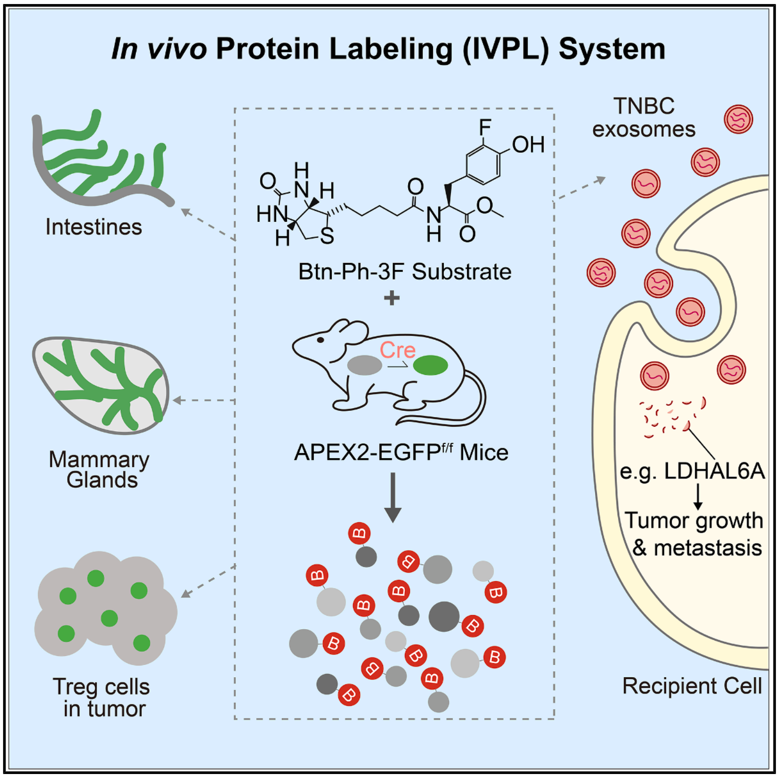
Molecular Cell

In vivo proteomic labeling reveals diverse proteomes for therapeutic targets

Graphical abstract



Highlights

- IVPL enables cell-type-specific *in vivo* proteomics via new substrates and APEX2 mice
- Maps *in situ* proteomes of intestinal epithelium, mammary gland, and tumor T_{reg} cells
- Traces patient-derived exosomal proteins in recipient organs in live mice
- Identifies LDHAL6A as a pro-metastatic, targetable vulnerability in TNBC

Authors

Qilong Wang, Yuening Jiang,
Meiyu Bi, ..., Hang Yin, Peng Zou, Mo Li

Correspondence

limo@hsc.pku.edu.cn

In brief

Wang et al. establish an *in vivo* proteomic labeling (IVPL) system combining the Btn-Ph-3F substrate with APEX2-EGFP$^{f/f}$ Cre crosses to map cell-type-specific proteomes in intact organs and trace patient-derived exosomal proteins. IVPL identifies LDHAL6A as a pro-metastatic driver and highlights its potential as a therapeutic target in TNBC.

Wang et al., 2025, Molecular Cell *85*, 4651–4666
December 18, 2025
https://doi.org/10.1016/j.molcel.2025.11.008

CellPress

Molecular Cell

Article

In vivo proteomic labeling reveals diverse proteomes for therapeutic targets

Qilong Wang,[1,2,3,4,14] Yuening Jiang,[5,14] Meiyu Bi,[1,2,3,4,14] Gang Wang,[6,7] Yinan Xiao,[1,2,3,4] Hailian Zhao,[8,9] Yuhan Guo,[1,2,3,4] Xinyu Li,[1,2,3,4] Wei Yue,[1,2,3,4] Na Zhang,[1,2,3,4] Bingteng Xie,[10] Yuanchao Xue,[8,9] Hang Yin,[11,12,13] Peng Zou,[6,7] and Mo Li[1,2,3,4,15,*]

[1]State Key Laboratory of Female Fertility Promotion, Center for Reproductive Medicine, Department of Obstetrics and Gynecology, Peking University Third Hospital, Beijing 100191, China
[2]Department of Human Anatomy, Histology and Embryology, School of Basic Medical Sciences, National Clinical Research Center for Obstetrics and Gynecology, Third Hospital, Peking University Health Science Center, Beijing 100191, China
[3]Key Laboratory of Assisted Reproduction (Peking University), Ministry of Education, Peking University Medical and Health Analysis Center, Beijing 100191, China
[4]Beijing Key Laboratory of Reproductive Endocrinology and Assisted Reproductive Technology (Peking University Third Hospital), Beijing 100191, China
[5]Medical Center for Human Reproduction, Beijing Chao-Yang Hospital, Capital Medical University, Beijing 100020, China
[6]College of Chemistry and Molecular Engineering, Peking-Tsinghua Center for Life Sciences, PKU-IDG/McGovern Institute for Brain Research, Peking University, Beijing 100871, China
[7]Chinese Institute for Brain Research (CIBR), Beijing 102206, China
[8]Key Laboratory of RNA Biology, Institute of Biophysics, Chinese Academy of Sciences, Beijing 100101, China
[9]University of Chinese Academy of Sciences, Beijing 100049, China
[10]Key Laboratory of Molecular Medicine and Biological Diagnosis and Treatment (Ministry of Industry and Information Technology), School of Life Science, Beijing Institute of Technology, Beijing 100081, China
[11]School of Pharmaceutical Sciences, Tsinghua University, Beijing 100084, China
[12]Tsinghua-Peking Center for Life Sciences, Tsinghua University, Beijing 100084, China
[13]Beijing Frontier Research Center for Biological Structure, Tsinghua University, Beijing 100084, China
[14]These authors contributed equally
[15]Lead contact
*Correspondence: limo@hsc.pku.edu.cn

https://doi.org/10.1016/j.molcel.2025.11.008

SUMMARY

Proteomics is transforming medical sciences, but bridging isolated samples with intact *in vivo* microenvironments remains a major hurdle. We present an *in vivo* proteomic labeling (IVPL) platform built on a new substrate, Btn-Ph-3F, and engineered ascorbate peroxidase (APEX2)-EGFP$^{f/f}$ mice. Btn-Ph-3F shows high stability in organs possessing complex microenvironments, while APEX2-EGFP$^{f/f}$ mice readily cross with commercial Cre lines, enabling specific proteomic labeling for customized cell groups in distant organs. IVPL robustly profiles *in situ* proteomes of intestinal epithelium, mammary gland, and tumor-infiltrating T$_{reg}$ cells, and, critically, labels trace exogenous proteomes from patient-derived exosomes in live mice. We identify lactate dehydrogenase A-like 6A (LDHAL6A) as a persisting exosomal effector that promotes malignant programs in recipient cells. Inhibition of LDHAL6A combined with paclitaxel treatment markedly suppresses triple-negative breast cancer growth and metastasis. Collectively, our work not only establishes an advanced model for IVPL but also profiles ultimately exosomal actors in recipient organs for targeted therapy.

INTRODUCTION

Proteomics, a pivotal discipline in life and medical sciences, has transformed our understanding of biological processes.[1,2] Yet conventional approaches that profile isolated samples inherently overlook the contextual cues and spatial relationships *in vivo*. Given the complexity and dynamism of living systems, proteins are best examined in their native microenvironments to reveal authentic functions and behaviors.[3,4] Lacking this context blunts mechanistic insight and impedes discovery of clinically relevant biomarkers and drug targets. Bridging the divide between *ex vivo* catalogs and intact *in vivo* contexts is essential for the next phase of proteomics and for translating its advances into more precise diagnostics and therapeutics.

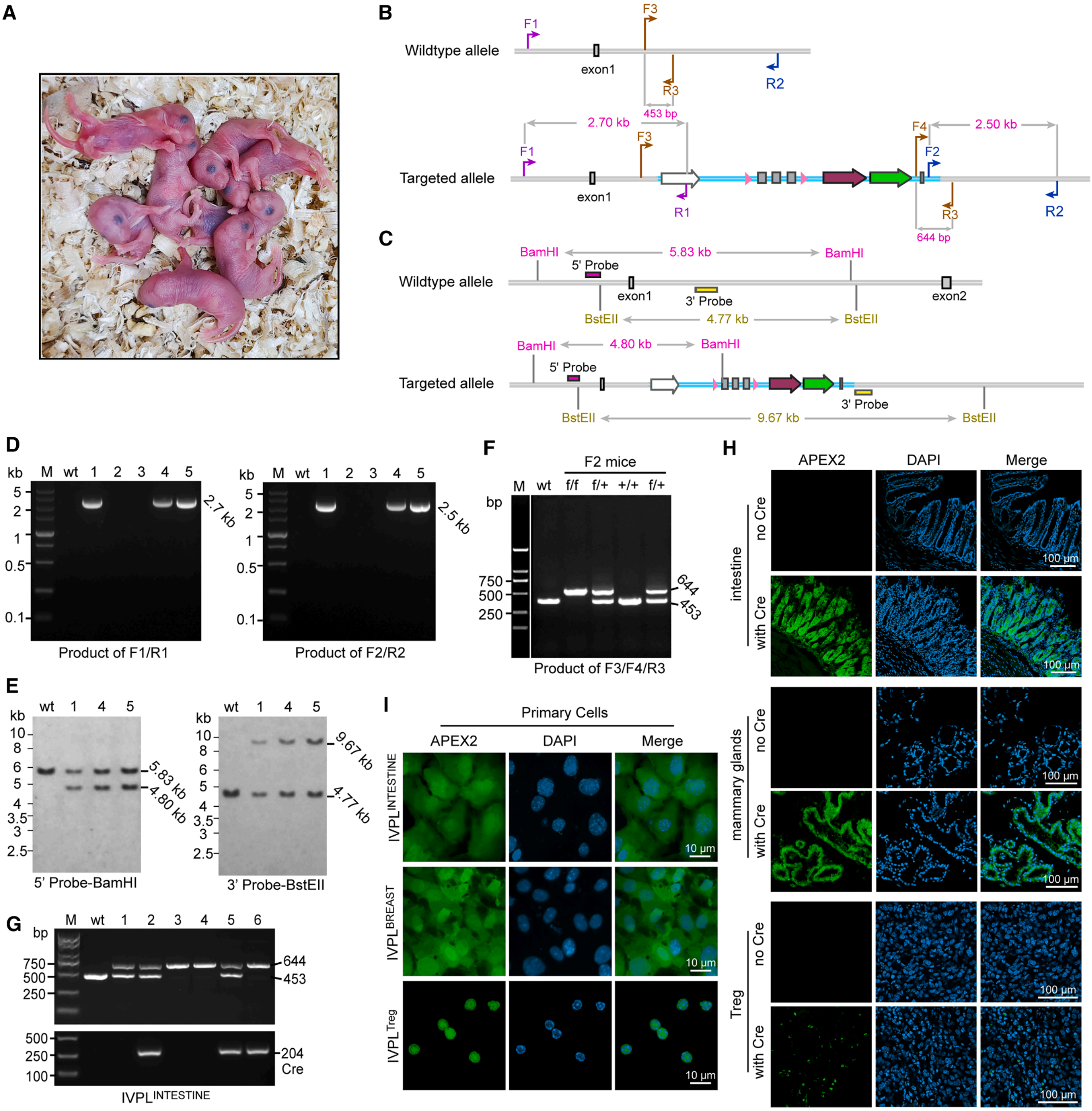


Figure 1. IVPL mouse generation

(A) F2 IVPL pups.

(B) Schematic of PCR-based strategy for genotyping wild-type (WT) and targeted alleles. In F1 pups, targeted alleles yield 2.7 kb (F1/R1) and 2.5 kb (F2/R2) bands. In F2 pups, WT alleles yield a 453 bp band (F3/R3), targeted alleles yield a 644 bp band (F4/R3), and heterozygotes show both bands.

(C) Southern blot analysis strategy for the inserted sequence in the genome.

(D) Genotyping of F1 IVPL mice by PCR. PCR amplification of the left and right arms of the inserted transgene generated 2.7 and 2.5 kb bands, respectively.

(E) Southern blot analysis of BamHI- or BstEII-digested tail DNA using Rosa26-specific 5′-hybridization and 3′-hybridization probes shows the predicted 4.80 and 9.67 kb bands, respectively, in addition to the 5.83 and 4.77 kb fragments from the Rosa26 WT locus.

(F) Genotyping of F2 pups from (A). PCR of WT samples generated a single 453-bp band, homozygous F2 samples generated a single 644-bp band, and heterozygous F2 samples generated 644- and 453-bp bands.

(G) Genotyping of Villin-Cre; IVPL (namely IVPL^{INTESTINE}) mice produced by mating between IVPL and Villin-Cre mice.

(legend continued on next page)

As a promising method for proteomic labeling, engineered ascorbate peroxidase (APEX2) technology enables rapid proximity-dependent protein biotinylation in living cells *in vitro*.[5,6] The peroxidase-mediated biotinylation based on the substrate of biotin-phenol (BP) occurs in close proximity to the enzyme, enabling precise spatial resolution within the subcellular or molecular level.[7] It is also encouraging that recent studies have utilized adeno-associated virus (AAV)-mediated expression of APEX2 or genetically targeted biotin ligase (BioID2) for proteomic labeling in specific contexts, such as neuronal and astrocyte proteomics in the mouse brain,[8,9] complementary to the earlier approach of non-canonical amino acid tagging for newly synthesized proteins specifically.[10,11] However, due to the lack of dedicated mouse models and novel reaction substrates effective for complex and heterogeneous mammalian tissues, comprehensive proteomic labeling of customized organs *in vivo* remains blank. Further, *in situ* labeling of trace amounts of exogenous proteins that induce pathological changes in organisms poses an even greater challenge.

In this study, we generated an *in vivo* protein labeling (IVPL) system based on new labeling substrates and homemade floxed APEX2 (APEX2-EGFP$^{f/f}$) mice. The Btn-Ph-3F substrate exhibited robust stability in organs possessing complex microenvironments and diverse degrading enzymes, while APEX2-EGFP$^{f/f}$ mice could be crossbred with any commercial Cre mice, enabling specific proteomic labeling for customized cell groups of intestinal epithelia, mammary glands, and T_{reg} cells within solid tumors *in vivo*. Moreover, although thousands of proteins have been identified in isolated exosomes *in vitro* over the past decade, the trace amounts of proteomic players that eventually persist and function in recipient cells *in vivo* remain unknown because exosomes absorbed by recipient cells often undergo the endocytic pathway and are likely to be directed to lysosomes for degradation.[12–14] Taking advantage of the IVPL system, we achieved *in situ* labeling of trace amounts of exogenous proteomics sourced from intravenously injected exosomes (namely "uptake exo-proteome") from luminal-A breast cancer (LABC) and triple-negative breast cancer (TNBC) patients in mammary glands of live mice. Different from the LABC uptake exo-proteome, the TNBC uptake exo-proteome exhibited a high prevalence of proteins related to glycolysis, adherens junctions, central carbon metabolism, and biosynthesis of amino acids. Accordingly, the proteomic features of recipient cells were profoundly reshaped, implying strong interactions between these exogenous proteins and the proteins in the recipient cells. Notably, we found that lactate dehydrogenase A-like 6A (LDHAL6A), a type of lactate dehydrogenase, emerged as the most abundant protein within the TNBC uptake exo-proteome. Knockdown of LDHAL6A weakened proliferative and metastatic abilities of TNBC cells. Targeting LDHAL6A combined with paclitaxel treatment effectively suppressed TNBC tumor growth and metastasis in both patient-derived xenograft (PDX)- and 231-mouse models. Collectively, this work develops an advanced platform for *in vivo* proteomic labeling (IVPL) and identifies exosomal factors within recipient organs that serve as actionable targets for tumor therapy.

RESULTS

Generation of IVPL mice for *in situ* proteomic labeling in multiple organs

Establishing IVPL requires (1) a base mouse enabling multi-organ protein labeling and (2) an *in vivo*-efficient chemical reaction. Considering the advantage of APEX2 peroxidase in rapidly biotin-tagging proteins *in vitro*, we hypothesized as a first step that constructing a whole-cell expression APEX2 flox mouse, which could mate with any Cre recombinase mouse targeting customized organs, may be feasible for the IVPL system. Under this framework, a Cre-inducible APEX2-EGFP cassette was inserted into Rosa26 of C57BL/6 zygotes via CRISPR-Cas9 to produce APEX2-EGFP$^{f/f}$ (IVPL) mice (Figures 1A, S1A, and S1B). Correct targeting in F1 and F2 animals was verified by PCR, southern blot, and genomic sequencing (Figures 1B–1F and S1C). For proteomic labeling in specific cell groups in diverse organs (or pathological tissues such as tumors), homozygous IVPL mice were crossed with Villin-Cre (intestine epithelium), mouse mammary tumor virus (MMTV)-Cre (mammary gland), and Foxp3-Cre (T_{reg}), yielding IVPL^{INTESTINE}, IVPL^{BREAST}, and IVPL^{Treg} mice, respectively (Figures 1G and S1D). For IVPL^{Treg}, colon adenocarcinoma MC38 cells were subcutaneously implanted into the mouse to profile the *in situ* proteome of infiltrating T_{reg} cells. This trio of mice models also corresponds to the three subsequent methods of substrate administration, namely oral gavage (PO), nipple injection, and intratumoral injection. Immunofluorescence confirmed specific APEX2-EGFP expression in target tissues without detectable signal in heart, liver, lung, kidney, brain, or ovary (Figures 1H and S1E), and APEX2-positive cells sorted from targeted organs exhibited both nuclear and cytoplasmic localization (Figure 1I). Survival, fertility, and open-field behavior were within normal ranges in IVPL^{INTESTINE}, IVPL^{BREAST}, and IVPL^{Treg} mice (Figures S1F–S1H). These results collectively indicate that IVPL mice enable specific APEX2 expression in customized organs, offering potential for applications across multiple scenarios.

Synthesis of new compounds for the IVPL system

Given the challenges of IVPL, we evaluated five substrates, two newly designed (Btn-O-Ph and Btn-Ph-3F) and three *in vitro* reagents (Btn-An, Btn-Nap,[15] and BP), to compare their *in vivo* performances and establish IVPL for specific organs (Figure 2A). Synthesis of Btn-O-Ph and Btn-Ph-3F was confirmed by $^{1}H/^{13}C$ NMR (Figures S2A–S2F). Next, we established three methods of substrate administration, namely, oral gavage for

---

(H) Immunofluorescence detection of APEX2 in the intestines, mammary glands, and T_{reg} cells in tumors from indicated Cre; IVPL mice. IVPL mice that didn't crossbreed with Cre mice were used as controls. Scale bar, 100 μm.

(I) Immunofluorescence of APEX2 in sorted primary cells isolated from intestine epithelium, mammary glands, and T_{reg} cells of IVPL^{INTESTINE}, IVPL^{BREAST}, and IVPL^{Treg} mice, respectively. Scale bar, 10 μm.

See also Figure S1.

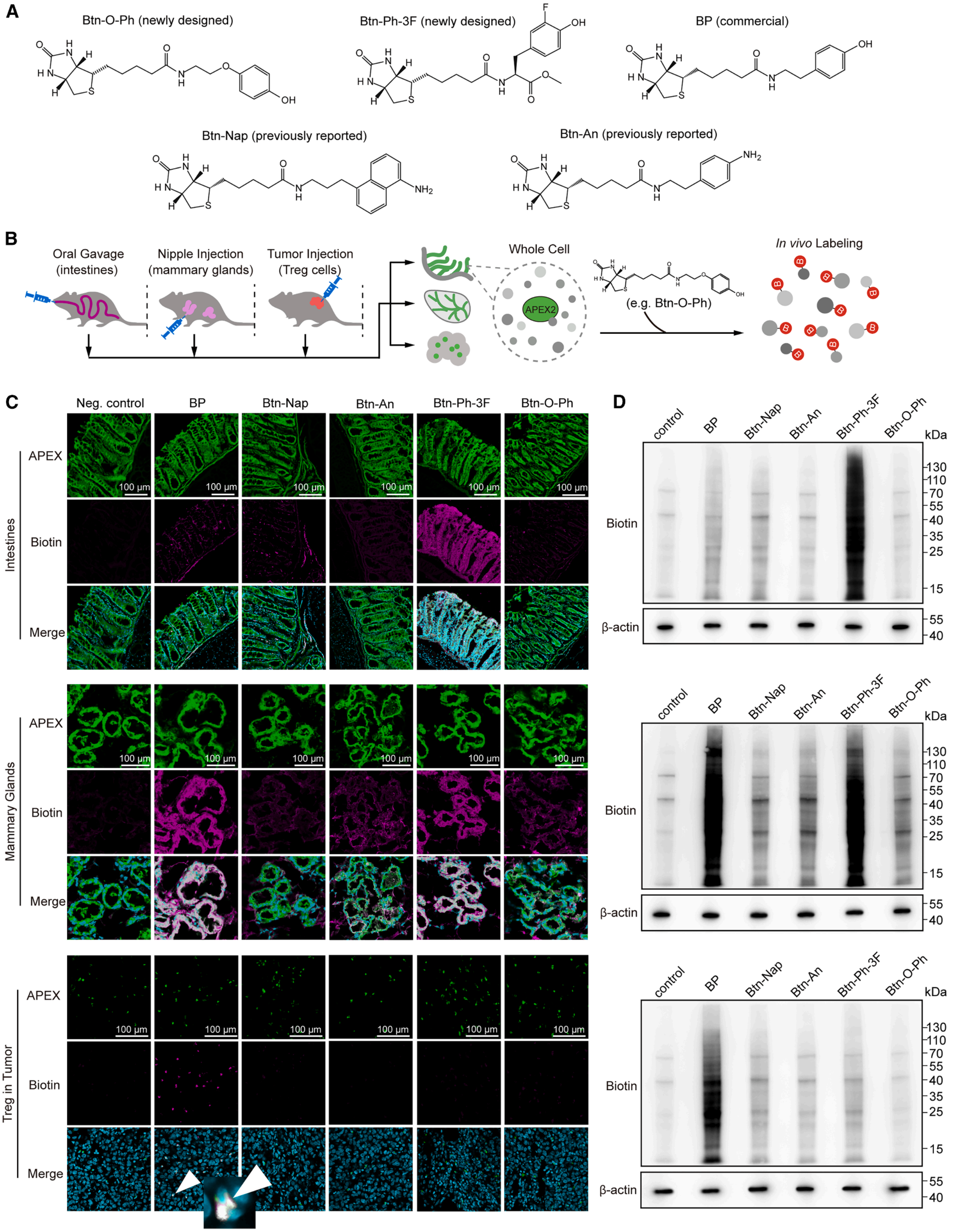


Figure 2. Design of new substrates and establishment of IVPL system

(A) Chemical structures of biotin-phenol (BP) and BP derivatives with substituents at the para-position (Btn-O-Ph), ortho-position (Btn-Ph-3F), and with aromatic amine structures (Btn-An and Btn-Nap).

(legend continued on next page)

IVPL^{INTESTINE} mice, nipple injection for IVPL^{BREAST} mice, and intratumoral injection for IVPL^{Treg} mice, to assess the protein labeling efficiency of five substrates in live animals. The individual substrate was incubated in the body for 45–90 min (depending on different organs), followed by the injection of H_2O_2 using the same administration to rapidly trigger protein labeling. After 90 s, the entire organ (tumor) was removed from the anesthetized mouse and then subjected to rapid quenching and lysis for protein harvesting or to fixation for staining (Figure 2B). Another set of IVPL mice that underwent the same treatment but without the substrate was used as the negative control. In intestines, immunofluorescence showed efficient biotinylation of nearly all epithelial cells in both nucleus and cytoplasm with Btn-Ph-3F, whereas the other substrates yielded minimal or unsatisfactory labeling (Figure 2C). Streptavidin pull-down of whole-intestine lysates corroborated these results: abundant banded biotinylated proteins with Btn-Ph-3F, but far fewer with other substrates or controls (Figure 2D). In IVPL^{BREAST}, Btn-Ph-3F and BP achieved higher, comparable efficiencies, while the remaining three compounds were inferior (Figures 2C and 2D). Unlike in the mammary glands or intestines, Btn-Ph-3F failed to efficiently label proteins in tumor-infiltrating T_{reg} cells, whereas BP achieved high labeling efficiency (Figures 2C and 2D). We reasoned that this organ-dependent difference arises from distinct local microenvironments. The intestine, rich in degradative enzymes, rapidly metabolizes small molecules, favoring the stability of fluorinated substrates such as Btn-Ph-3F due to strong bonds formed by fluorine atoms and high electronegativity.[16,17] By contrast, solid tumors have fewer degradative enzymes but much higher tissue density,[18–20] conditions that favor substrates with superior membrane permeability.

To validate this, we assessed substrate stability in intestinal lysate (37°C, 90 min) and found Btn-Ph-3F markedly more stable than the other four substrates, with the advantage persisting in the lysate of the liver, an organ rich in metabolic enzymes (Figure S2G), further supporting the notion that Btn-Ph-3F is well-suited for application in organs with diverse degrading enzymes and complex microenvironments. We then compared pharmacokinetics (PKs) of Btn-Ph-3F vs. BP across intravenous (i.v.), intraperitoneal (i.p.), and PO routes, quantifying C_{max}, T_{max}, area under the concentration-time curve from 0 to 6 h ($AUC_{0–6\ h}$), and plasma half-life in eight tissues and plasma. Btn-Ph-3F showed greater tissue-specific accumulation, e.g., liver C_{max} 1,500.00 vs. 495.93 ng/g (i.v.), small intestine 3,335.37 vs. 814.43 ng/g (PO), and colon 2,512.73 vs. 622.13 ng/g (PO). $AUC_{0–6\ h}$ corroborated higher exposure (small intestine, PO: 9,735.52 vs. 1,430.33 ng·h/g), despite similar plasma half-lives (1.45 vs. 0.94 h) (Figures S2H and S2I; Table S1). These data underscore the superior bioavailability, prolonged retention, and organ-selective distribution of Btn-Ph-3F, particularly in metabolically active or enzymatically harsh environments such as the liver and gastrointestinal tract. To link PK profiles with labeling efficiency, we further performed PK/pharmacodynamic (PK/PD) relationship analysis in representative organs with distinct physiology: liver (highly perfused, metabolically active), small intestine (mucosal epithelium-lined, tubular structures), and ovary (solid intra-abdominal, connective tissue-rich organs). Streptavidin blots showed *in vivo* biotinylation only above a minimum effective concentration (MEC): $\geq$541.43 ng/g for Btn-Ph-3F in liver (i.v., 5 mg/kg) vs. $\geq$500.83 ng/g for BP (i.v., 20 mg/kg), and $\geq$690.03 ng/g for Btn-Ph-3F in small intestine (PO, 2 mg/kg) vs. $\geq$837.53 ng/g for BP (PO, 50 mg/kg) (Figures S2J and S2K). This quantitative relationship allowed extrapolation of MEC values to other tissues with comparable anatomical and physiological features (Table S1), providing a practical reference for compound selection, dosing route, and expected labeling performance (Table S1). Together, the favorable PK characteristics and *in situ* labeling performance of Btn-Ph-3F support its broader applicability in diverse tissue labeling contexts *in vivo*.

The IVPL system achieves high-quality *in situ* proteomics for distinct organs

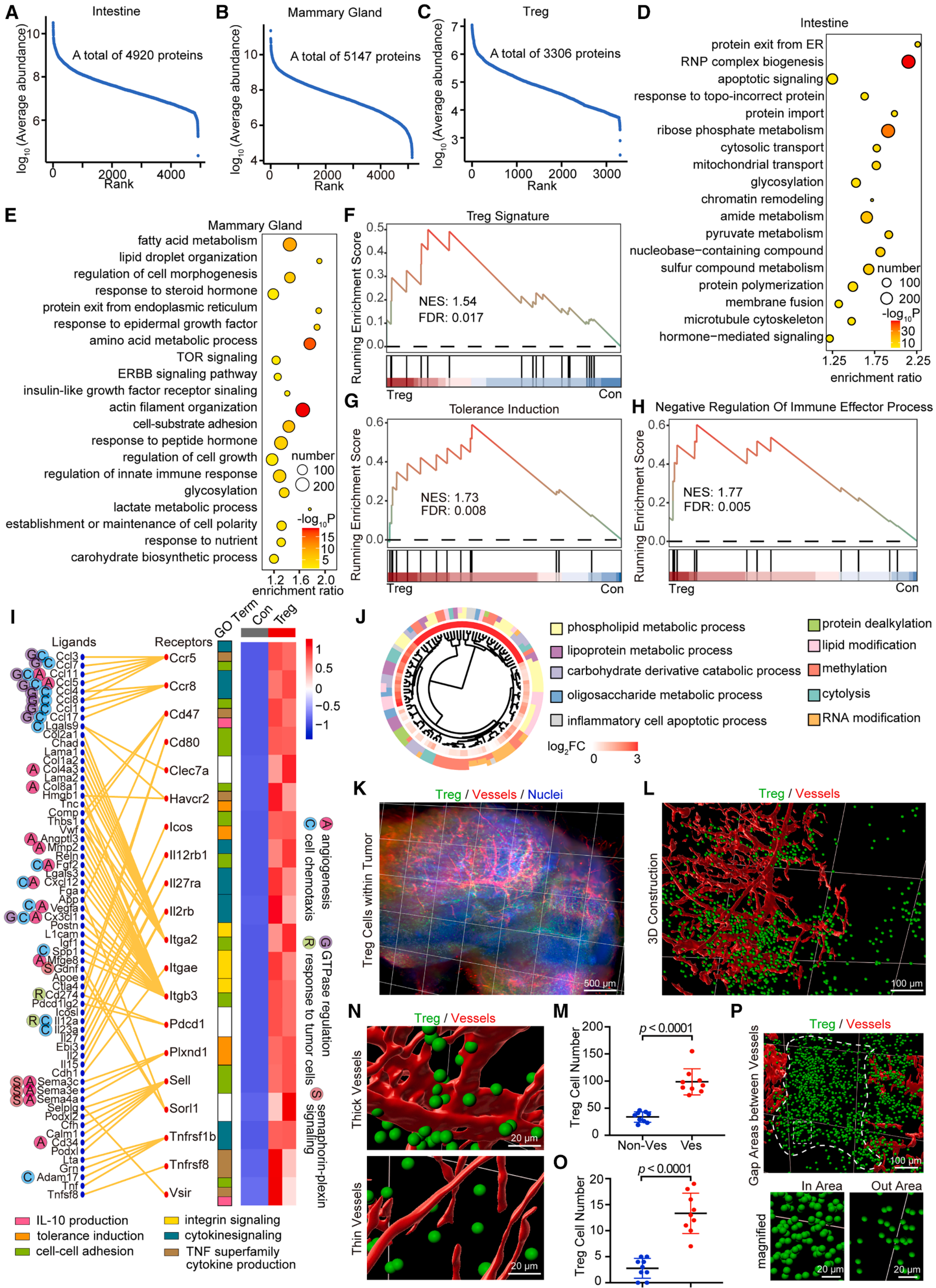
The IVPL system achieved deep, quantitative coverage with strong overlap across samples. At a protein-level false discovery rate (FDR) of <1%, we quantified 4,920 proteins in intestinal epithelial cells, 5,147 in mammary gland cells, and 3,306 in T_{reg} cells *in situ* (Figures 3A–3C; Table S2). Correlation analysis and violin plots showed replicate consistency and protein abundances in each group (Figures S3A and S3B). To exclude treatment artifacts, paired mice received H_2O_2 or Btn-Ph-3F as described above, and whole-tissue proteomics showed no appreciable changes vs. untreated controls by correlation and differential analyses (Figures S3C–S3F; Table S2). Negative controls (Cre-negative tissues, no-substrate, or no-H_2O_2) ruled out background biotinylation and non-specific streptavidin binding (Figure S3G; Table S2). Gene Ontology (GO) analysis of intestinal epithelial cells highlighted pathways consistent with lower gastrointestinal biology, including ribose-phosphate and pyruvate metabolism, protein import, cellular amide metabolism, and apoptotic signaling. The latter is proposed to result from the rapid turnover rate of intestinal epithelial cells, which typically renew every 4 to 5 days (Figure 3D). Similarly, the proteome of mammary gland cells also exhibited organ specificity, with typical biological processes including lipid droplet organization, amino acid metabolism, fatty acid metabolism, and the ERBB signaling pathway (Figure 3E). For the T_{reg} proteome within tumors, we found that the *in situ* proteome encompasses nearly all core T_{reg} signatures, including forkhead box P3 (FOXP3),

(B) Schematic of IVPL system design. See also STAR Methods.

(C) Immunofluorescence staining of intestines, mammary glands, and tumor masses harvested from IVPL^{INTESTINE}, IVPL^{BREAST}, and IVPL^{Treg} mice, respectively, after *in situ* reaction with different substrates. Scale bar, 100 μm.

(D) Lysates of the intact intestines, mammary glands, and tumor mass, taken from anesthetized IVPL^{INTESTINE}, IVPL^{BREAST}, and IVPL^{Treg} mice after *in situ* reaction with different substrates, were incubated with streptavidin-coated magnetic beads, and the precipitated biotinylated proteins were subjected to SDS polyacrylamide gel electrophoresis. Biotin was then detected by western blot analysis. Reactions performed in the absence of substrates were used as controls. β-actin was used as a loading control.

See also Figure S2 and Table S1.



(legend on next page)

inducible T cell COStimulator (ICOS), and IKAROS family zinc-finger 2 (IKZF2) (Figure 3F). Benchmarking against fluorescence-activated cell sorting (FACS)-sorted T_{reg} cells from the same tumors (Table S3) showed high concordance (>87% protein overlap) while revealing additional immunoregulatory proteins that are frequently under-represented after isolation (e.g., LAG3, CCR8, TIGIT, and ITGA2) (Figures S3H–S3K). GO analyses were concordant and significantly enriched for tolerance induction and negative regulation of immune effector processes (Figures 3G and 3H). Ligand-receptor mapping detected receptors including Ccr5/Ccr8, Cd47/Cd80, Clec7a, Havcr2, Il12rb1/Il27ra/Il2rb, integrins (Itga2/Itgae/Itgb3), Pdcd1, Sell, Tnfrsf1b/Tnfrsf8, and Vsir, with their binding ligands predicted. Beyond canonical T_{reg} ligands (e.g., angiogenic and GTPase-regulatory cues), we identified previously unreported candidates, Col2a1, Chad, Lama1/2, and Podxl2 (Figure 3I). To further clarify the cellular context of these ligand-receptor interactions, we mapped the predicted ligands to their known cell-type-specific expression profiles using the PanglaoDB database, a curated single-cell transcriptomic resource comprising over 1,000 murine datasets.[21] This allowed us to assign each ligand to its most likely source cell type and reconstruct a directional interaction network from predicted ligand-expressing cells (e.g., macrophages, natural killer [NK] cells, fibroblasts) to T_{reg} cells (Figures S3L and S3M). By comparing the *in situ* T_{reg} proteome with the dataset downloaded from the *in vitro* T_{reg} proteome,[22] we filtered out 991 proteins that were not previously identified (Table S4). Interestingly, cytolysis, which is not typically exhibited as part of standard T_{reg} functions, was found to be enriched in the *in vivo* proteome (Figure 3J), suggesting an unusual role for these cells that warrants further exploration. Leveraging the visualization advantage of the targeted cell population, we performed whole-tissue imaging and 3D reconstruction to map T_{reg} cells co-stained with the vascular marker CD31 (Figure 3K). T_{reg} cells were significantly enriched in vascular-dense vs. vascular-sparse regions (Figures 3L and 3M), consistent with their pro-angiogenic roles. Larger, mature vessels recruited more T_{reg} cells, suggesting stronger endothelial ligand cues for adherence (Figures 3N and 3O). Notably, T_{reg} cells also clustered in vessel-scarce junctions between two vascular-rich zones (Figure 3P), potentially priming neovascularization. These spatial readouts may forecast dynamic vascular remodeling and inform evaluation of anti-angiogenic therapies.

IVPL applicability in barrier-protected organs was tested by generating $\text{IVPL}^{\text{BRAIN}}$ mice ($\text{APEX2-EGFP}^{f/f}$ × Camk2a-Cre) for excitatory neuron-specific labeling across forebrain regions. Btn-Ph-3F was delivered intracerebroventricularly into the lateral ventricle, yielding robust, spatially confined biotinylation in the cortex, hippocampus, and striatum by immunofluorescence and streptavidin staining (Figure S4). Mass spectrometry identified a rich brain-specific proteome, including canonical neuronal markers and regionally enriched proteins (Figure S4; Table S2). The results demonstrate that IVPL is compatible with brain tissue through localized delivery, achieving both anatomical precision and proteomic depth. Together with systemic labeling results, these findings establish a general framework for IVPL delivery: systemic routes suffice for most organs, whereas local injection is reserved for anatomically protected tissues such as the brain.

The IVPL system enables *in vivo* labeling for trace amounts of exogenous proteins

Next, we wondered whether this model could label trace amounts of exogenous proteins *in vivo*. As a class of extracellular vesicles (EVs) that play a crucial role in the communication between cells, exosomes transport various cargoes, including proteins, across the body for physiological and pathological

Figure 3. *In situ* proteomics across distinct organs achieved by IVPL system

(A–C) Ranking plots based on the $\log_{10}$-transformed protein abundances of $\text{IVPL}^{\text{INTESTINE}}$, by using the substrate of Btn-Ph-3F (A); $\text{IVPL}^{\text{BREAST}}$, by using the substrate of Btn-Ph-3F (B); and $\text{IVPL}^{\text{Treg}}$, by using the substrate of BP (C) proteomes.

(D) GO analysis for the proteomes of intestinal epithelial cells. The node size indicates the number of overlapped proteins enriched in the specific term. The color schema from yellow to red represents $-\log_{10}$ (enrichment p value) from low to high.

(E) GO analysis for the proteomes of mammary gland cells. The top 20 terms of biological processes are shown.

(F–H) GSEA of three representative gene sets for T_{reg} cells, including T_{reg} signature, tolerance induction, and negative regulation of immune effector process, all showing significant enrichments in $\text{IVPL}^{\text{Treg}}$ proteomes. NES, normalized enrichment score.

(I) Interaction analysis between receptors (right) identified in T_{reg} cells and predicted binding ligands (left). Dots adjacent to ligands and colored squares adjacent to receptors indicate the associated GO terms. The heatmap on the right shows the expression levels of receptors. The color key from blue to red indicates the expression levels from low to high.

(J) Circular dendrogram of GO terms shows the functional clustering of 991 proteins uniquely identified in the *in situ* $\text{IVPL}^{\text{Treg}}$ proteome compared with the *in vitro* proteome. The inner ring displays the $\log_2$ fold change (FC), indicated by different shades of red, while the outer ring shows GO terms assigned to each protein.

(K) 3D reconstruction of intact tumor mass taken from $\text{IVPL}^{\text{Treg}}$ mice. The tumor mass was surgically isolated from anesthetized $\text{IVPL}^{\text{Treg}}$ mice, followed by whole-tissue fluorescent imaging. T_{reg} cells, green; blood vessels labeled by CD31, red; nuclei labeled by Hoechst33342, blue. Scale bar, 500 μm.

(L) The distribution T_{reg} cells within the vascular network of intact tumors from $\text{IVPL}^{\text{Treg}}$ mice was 3D reconstructed using the Imaris program. T_{reg} cells were spot-modeled for easier visualization. Scale bar, 100 μm.

(M) The number of T_{reg} cells per unit volume (0.15 mm^3) in areas rich in vessels and areas lacking vessels was counted with the Imaris program. Nine units from three independent biological replicates were examined for each group.

(N) The interaction between T_{reg} cells and blood vessels within the tumor. The direct binding usually occurred between these immune cells and thick vessels rather than thin vessels. Scale bar, 20 μm.

(O) The number of T_{reg} cells adhered per 0.1 mm of blood vessel was counted using the Imaris program. Nine length units from three independent biological replicates were examined for each group.

(P) The distribution of T_{reg} cells in the vessel-scarce areas at the junctions of two vascular-rich sites was 3D reconstructed using the Imaris program. Scale bar, 100 μm. The representative aggregation patterns of T_{reg} cells within and outside the junction area are depicted in enlarged images. Scale bar, 20 μm. Data are presented as mean values ± SD (M and O). p values were calculated by unpaired two-tailed Student's t tests. $p < 0.05$ was considered statistically significant.

See also Figures S3 and S4 and Tables S2, S3, and S4.

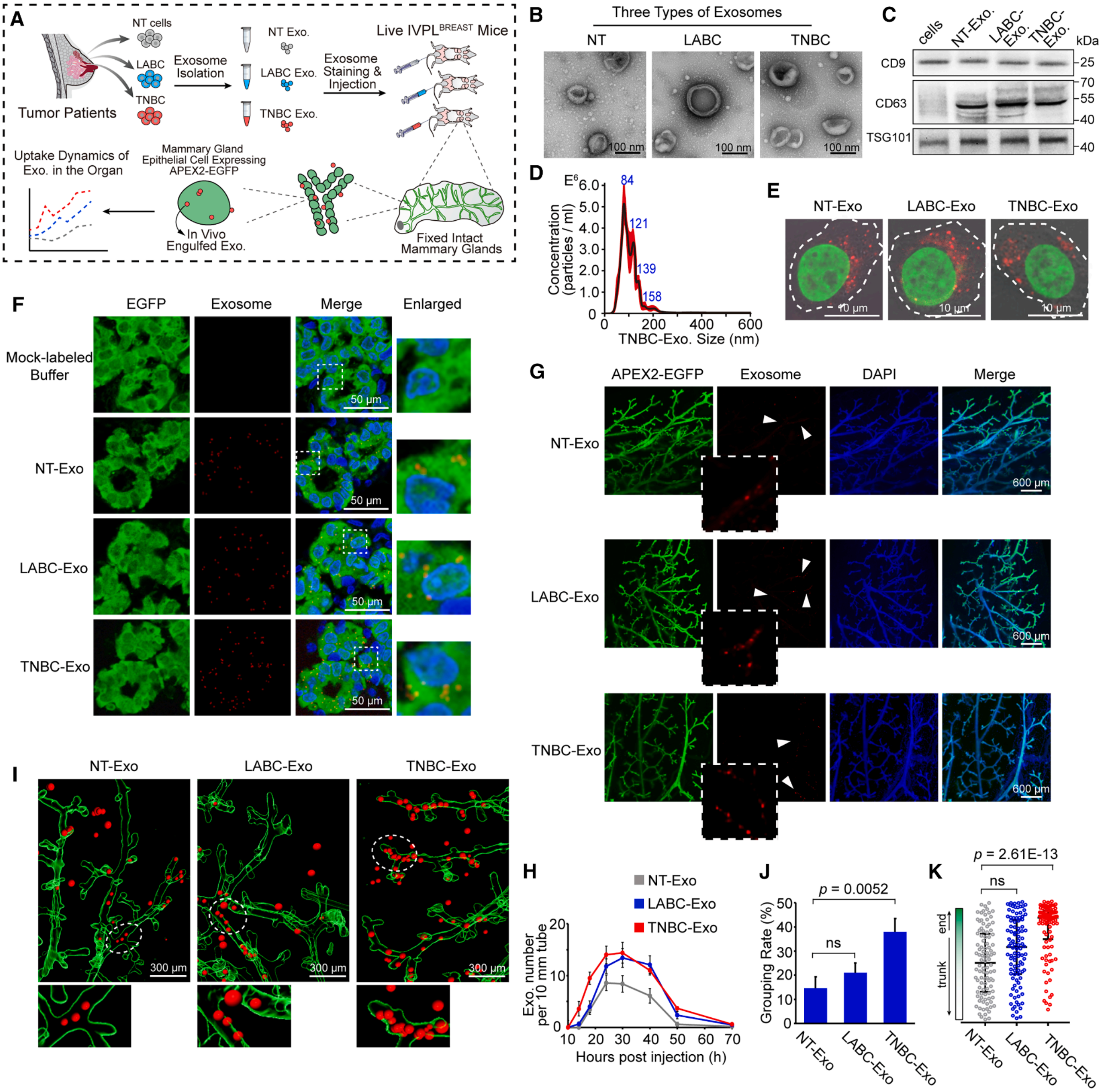


Figure 4. Uptake dynamics of PTDEs in mammary glands of IVPL^{BREAST} mice

(A) Scheme of uptake dynamics analysis of patient tumor-derived exosomes (PTDEs) in the mammary glands of IVPL^{BREAST} mice after i.v. injection of different types of exosomes.

(B) Morphological assessment of different types of exosomes by transmission electron microscopy. Scale bar, 100 nm.

(C) Exosome identification by western blot analysis of the exosomal markers CD9, CD63, and TSG101.

(D) Nanoparticle tracking analysis of the obtained vesicles from the conditioned medium of primary TNBC tumor cells. The *x* axis represents vesicle diameters, and the *y* axis represents vesicle concentration (particles mL^{-1}).

(E) Fluorescent labeling of exosomes by a lipophilic dye PKH26. The labeling enabled exosome visualization in live cells, which is shown by an example in live HeLa cells expressing EGFP-H2B. Scale bar, 10 μm.

(F) Uptake of PTDEs by mammary glands of IVPL^{BREAST} mice detected by fluorescence microscopy of frozen sections. Mock-labeled buffer, PKH26 dye subjected to the same labeling protocol but without exosomes. Scale bar, 50 μm.

(G) 3D reconstruction of intact mammary glands from IVPL^{BREAST} mice treated with different types of exosomes. The internalized exosomes within the mammary glands are indicated by arrowheads. Scale bar, 600 μm.

(H) Uptake dynamics of different types of exosomes by mammary glands of IVPL^{BREAST} mice. The number of exosomes engulfed by mammary glands was counted. The exosomes within at least 30 gland ducts per group from three independent biological replicates were examined.

(legend continued on next page)

processes.[12,23] While thousands of proteins have been identified in isolated exosomes *in vitro* over the past decade, the trace amounts of proteomic players that eventually persist and function in recipient cells *in vivo* remain largely unexplored because exosomes taken up by recipient cells usually follow the endocytic pathway and tend to enter lysosomes for degradation.[12–14] Thus, we collected primary tumor cells from 20 LABC and 20 TNBC patients (Table S5) and isolated exosomes from each. Exosomes from the same subtype were pooled to generate LABC and TNBC exosome pools, respectively. Exosomes from primary normal breast epithelial cells obtained from reduction mammoplasties served as non-tumor controls (NT exosomes) (Figure 4A).

Transmission electron microscopy confirmed the typical exosomal morphology (Figure 4B), and western blotting verified the presence of CD9, CD63, and TSG101 (Figure 4C). Nanoparticle tracking analysis showed that NT exosomes were mostly 70–135 nm in diameter, whereas LABC and TNBC exosomes were more heterogeneous, ranging from 45 to 150 nm (Figures 4D, S5A, and S5B). To monitor exosome behavior, PKH26-labeled exosomes (10 μg) were intravenously injected into IVPL^{BREAST} mice, allowing live-cell visualization (Figure 4E). Mammary glands collected 6–72 h post-injection were analyzed for uptake dynamics. Exosomes from NT, LABC, and TNBC sources were all internalized by APEX2-EGFP-positive mammary epithelial cells (Figure 4F), as confirmed by z-stack imaging (Figure S5C). Multiplex tissue imaging and 3D modeling (Figure 4G) revealed that exosomal foci predominantly emerged from large-caliber vascular networks (Figures S5D–S5F). Notably, TNBC exosomes were taken up as early as 14 h post-injection, whereas NT and LABC exosomes entered at 18–19 h, with uptake saturating at 24–26 h (Figure 4H). By 70 h, all groups showed minimal residual vesicles. Approximately 38% of TNBC exosomes formed small *in vivo* clusters (>3 foci), compared with 14.6%–21% for NT and LABC (Figures 4I–4J). Unlike the random distribution of NT and LABC exosomes, TNBC vesicles preferentially accumulated near the terminal duct lobular units, the origin of most breast cancers,[24,25] resulting in local exosome overload (Figures 4I and 4K). These findings indicate that TNBC exosomes exhibit stronger invasive activity within mammary glands *in situ*.

To profile the *in situ* uptake exosomal proteome of recipient cells, IVPL was activated 72 h after injection of patient tumor-derived exosomes (PTDEs), when no labeled vesicles were visible in the mammary gland. Biotinylated proteins from three biological replicates per group were analyzed against the *Homo sapiens* UniProt database (Figure 5A). Only peptide spectrum matches uniquely assigned to human proteins were retained to ensure species specificity.[26,27] Violin plots and correlation analyses confirmed data quality and reproducibility (Figures S5G and S5H). In total, 847, 919, and 971 human proteins were identified in NT_{EXO}-, $LABC_{EXO}$-, and $TNBC_{EXO}$-uptake proteomes, respectively (Table S6). Representative exosomal proteins, including TSG101, CLTC, RAB7A, RAB5A, and RAB5C, were shared across groups and enriched in Kyoto Encyclopedia of Genes and Genomes (KEGG) pathways related to endocytosis, cytoskeleton regulation, focal adhesion, and protein processing (Figures 5B and 5C). To validate the presence of exogenous proteins in recipient cells, six candidates from the $LABC_{EXO}$-uptake proteome were chosen based on clear molecular-weight differences between human and mouse homologs and availability of cross-reactive antibodies. Western blotting revealed dual bands for APLP2, CUL4B, RPL7, CDV3, DYNC1I2, and UBAP2L, corresponding to human and mouse forms (Figure 5D). Each protein was further overexpressed with both N-terminal EGFP and C-terminal red fluorescent protein (RFP) in LABC primary tumor cells. Two other proteins, PGK1 and GPX1, which are present in isolated exosomes[28,29] but absent in our uptake exosomal proteomes, served as negative controls. Following exosome isolation and injection (Figure 5E), recipient mammary gland cells examined at 72 h displayed merged orange fluorescence for all six proteins, confirming that these exosomal proteins remained intact in recipient cells, whereas controls showed no or sporadic EGFP or RFP signals (Figures 5F and 5G).

Next, we focused on those proteins that were highly retained in the $TNBC_{EXO}$-uptake proteome compared with the NT_{EXO}-uptake proteome. For the $TNBC_{EXO}$-uptake proteome, as many as 400 proteins were more abundant (over 2-fold, $p < 0.05$) than that in the NT_{EXO}-uptake proteome (Figure 5H). Among them, 120 proteins were uniquely found in the $TNBC_{EXO}$-uptake proteome, including LDHAL6A, zinc-finger protein 36 like 1 (ZFP36L1), solute carrier family 38 member 2 (SLC38A2), SPG matrix AAA peptidase subunit, and paraplegin (SPG7). Thirty-six proteins were more than 10-fold abundant in the $TNBC_{EXO}$-uptake proteome, including SET domain (Su[var], enhancer of zeste, Trithorax) containing 1A (SETD1A), guanosine monophosphate reductase 2 (GMPR2), armadillo repeat containing 1 (ARMC1), and heterogeneous nuclear ribonucleoprotein A3 (HNRPA3) (Table S6). KEGG analysis showed enrichment in glycolysis/gluconeogenesis, adherens junctions, central carbon metabolism, and amino acid biosynthesis (Figure 5I). In the top pathway (glycolysis/gluconeogenesis), LDHAL6A, ENO1, and ENO3 were markedly increased in the $TNBC_{EXO}$-uptake proteome (Figure 5J), suggesting metabolic reprogramming of recipient cells. Conventional proteomics of *in vitro* isolated exosomes from TNBC patient tumors identified 3,550 proteins with robust quality control (Table S6; Figures S5I–S5K), yet LDHAL6A was absent from our dataset and public breast cancer exosome

---

(I) Exosome distribution in intact mammary glands from IVPL^{BREAST} mice treated with different types of exosomes was 3D reconstructed by the Imaris program. Exosomes within glands were spots modeled for easier visualization. Scale bar, 300 μm.

(J) Aggregation of internalized exosomes in mammary glands. An aggregate consisting of more than 3 exosomes was counted as a cluster. At least 200 exosomes per group from 3 independent biological replicates were examined.

(K) Exosome distribution in the ends and trunks of mammary glands. One hundred exosomes per group from three independent biological replicates were examined. Data are presented as mean values ± SD (H, J, and K). *p* values were calculated by unpaired two-tailed Student's *t* tests. $p < 0.05$ was considered statistically significant. ns, not significant.

See also Figure S5 and Table S5.

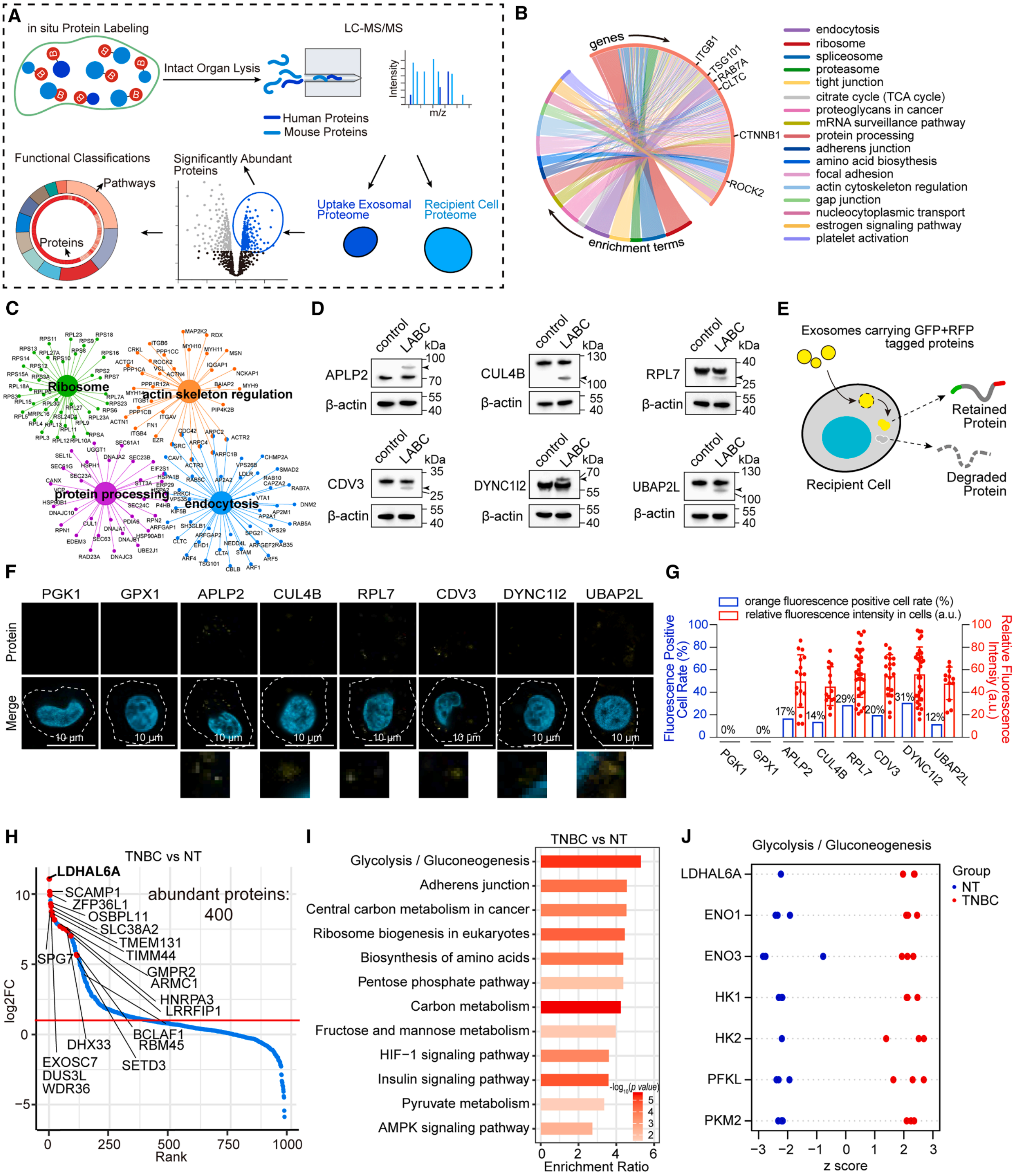


Figure 5. Analysis of uptake exosomal proteome of mammary glands in IVPL$^{\text{BREAST}}$ mice

(A) Flow diagram for profiling the uptake exosomal proteome and the endogenous recipient cell proteome. After the *in vivo* reaction, all proteins in mammary gland cells, including endogenous mouse proteins and patient-derived exosomal proteins, were labeled. The freshly harvested mammary glands were rapidly lysed for protein enrichment, followed by liquid chromatography-tandem mass spectrometry (LC-MS/MS) analysis. Proteins were identified by performing searches against *Homo sapiens* and *Mus musculus* databases.

(B) Circos plot showing enriched KEGG pathways for exosomal proteins taken up by recipient cells in IVPL$^{\text{BREAST}}$ mice.

(legend continued on next page)

resources (ExoCarta, Vesiclepedia) (Figure S5L). By contrast, IVPL uniquely captured LDHAL6A and other low-abundance, functionally relevant exosomal proteins *in vivo* (Figure S5M), highlighting its superior sensitivity over traditional approaches.

Using the IVPL system, which simultaneously decodes exogenous (human) and endogenous (murine) proteomes *in situ* (Figure 5A), we compared recipient-cell proteomes after treatment with NT, LABC, or TNBC exosomes. Across groups, 4,911–5,772 mouse proteins were detected (Table S6), with the TNBC_{EXO}-recipient proteome showing marked remodeling (Figure S5N). These cells displayed substantially more upregulated proteins than LABC_{EXO}-recipient cells (Table S6), with gene set enrichment analysis (GSEA) revealing enrichment in glycogen metabolism, post-translational modification, DNA repair, RNA metabolism, and oxidative stress response (Figure S5O). Among 52 LDHAL6A-interacting candidates expressed in TNBC_{EXO}-recipient cells, many overlapped with glycolysis/gluconeogenesis, pentose phosphate, and hypoxia-inducible factor 1 (HIF-1) signaling pathways (Figure S5P). Proteins involved in cell cycle and immune response were also elevated (Figures S5Q and S5R), and western blotting confirmed higher expression of six representative members of these pathways (Figure S5S). Given that glycolysis/gluconeogenesis was the most enriched pathway in the TNBC_{EXO}-uptake proteome, and that glycolysis-derived lactate modulates gene expression and signaling,[30–32] the reshaped proteome of TNBC_{EXO}-recipient cells mirrors the uptake profile. This finding confirms the impact of exogenous proteins on recipient cell fate and further demonstrates the advanced capabilities of the IVPL system.

Inhibition of LDHAL6A effectively suppresses TNBC growth and metastasis

As LDHAL6A was the most abundant protein in the TNBC_{EXO}-uptake proteome, we next verified its transfer to recipient cells. RFP-tagged LDHAL6A was expressed in TNBC cells (Figures 6A and 6B), and exosomes containing fluorescent LDHAL6A (exosomes$^{\text{LDHAL6A}}$) were intravenously injected into C57BL/6 mice. At 24 h post-injection, fluorescence imaging and western blotting confirmed LDHAL6A uptake by mammary gland cells (Figures 6C and 6D). Given that LDHAL6A functions as a lactate dehydrogenase,[33] we examined lactate production in recipient cells and found significantly higher levels following TNBC exosome treatment compared with NT or LABC exosomes (Figure 6E). Moreover, LDHAL6A knockdown in TNBC cells markedly reduced their proliferative and metastatic capacities (Figures 6F, 6G, S6A, and S6B). To explore whether LDHAL6A would be an efficient target for TNBC therapy, we established two TNBC mouse models: TNBC PDX in female non-obese diabetic (NOD)-severe combined immunodeficiency (SCID) mice (PDX mice) and female NOD-SCID mice implanted with MDA-MB-231 cells (231 mice). Tumors developed into palpable masses within 40–60 days, with slightly faster growth in 231 mice (Figure 6H). As expected, most tumors were Ki67-positive and apoptosis-negative (Figure 6I). Conventional paclitaxel monotherapy slowed tumor growth in both models during the first half of treatment but lost efficacy after ~35–40 days (Figures 6H and 6J). Of note, LDHAL6A knockdown via short hairpin RNA (shRNA) lentivirus from day 20 substantially inhibited tumor progression and reduced intratumoral lactate levels (Figures 6H and 6K). Importantly, combined LDHAL6A silencing and paclitaxel treatment produced a more pronounced and sustained suppression of tumor growth in most mice (12/15 PDX and 13/15 231), accompanied by a more marked decrease in Ki67 and more extensive apoptosis (Figures 6H–6K).

As distant metastasis remains the major therapeutic challenge in TNBC, we monitored metastatic progression in PDX and 231 mice. 45 days after primary tumor formation, 15 of 40 PDX mice and 17 of 40 231 mice exhibited distant metastases in lung and/or liver (Figures 6L and 6M). Paclitaxel alone modestly reduced the incidence (10/40 and 12/40, respectively), whereas LDHAL6A knockdown combined with paclitaxel markedly suppressed metastasis, detected in only 3/40 PDX and 4/40 231 mice (Figures 6L and 6M). Quantification of serial sections revealed 9–23 lung and 7–22 liver nodules in controls, slightly fewer with paclitaxel, but dramatically fewer when combined with shLDHAL6A (Figures 6N and 6O). Conversely, gain-of-function assays using LDHAL6A-rich exosomes from HEK293T cells (Figure S6C) accelerated primary tumor growth, increased Ki67 and intratumoral lactate (~2.0-fold in PDX and ~2.3-fold in 231), and exacerbated metastasis with ~90% more lung and ~70% more liver nodules (Figures S6D–S6J). Together, these results demonstrate that targeting LDHAL6A, particularly in combination with paclitaxel, effectively suppresses TNBC growth and metastasis.

(C) Functionally grouped network of the endocytosis, actin cytoskeleton regulation, protein processing, and ribosome pathways in (B).

(D) Examination of exosome-derived proteins in recipient cells by western blot analysis. Bands marked by arrowheads denote human proteins derived from LABC patient exosomes. β-actin was used as a loading control.

(E) Schematic diagram showing the approach for tracking exosome-derived proteins in recipient cells. The protein of interest was expressed with a GFP tag at the N-terminal end and RFP at the C-terminal end. The full-length protein could be detected as an orange signal in recipient cells if it did not undergo degradation.

(F) Immunofluorescence staining of six proteins as described in (D). PGK1 and GPX1, which are known to be degraded in recipient cells, were included as negative controls. Whole cells are outlined with dashed lines. Cellular regions with orange signals are enlarged. Scale bar, 10 μm.

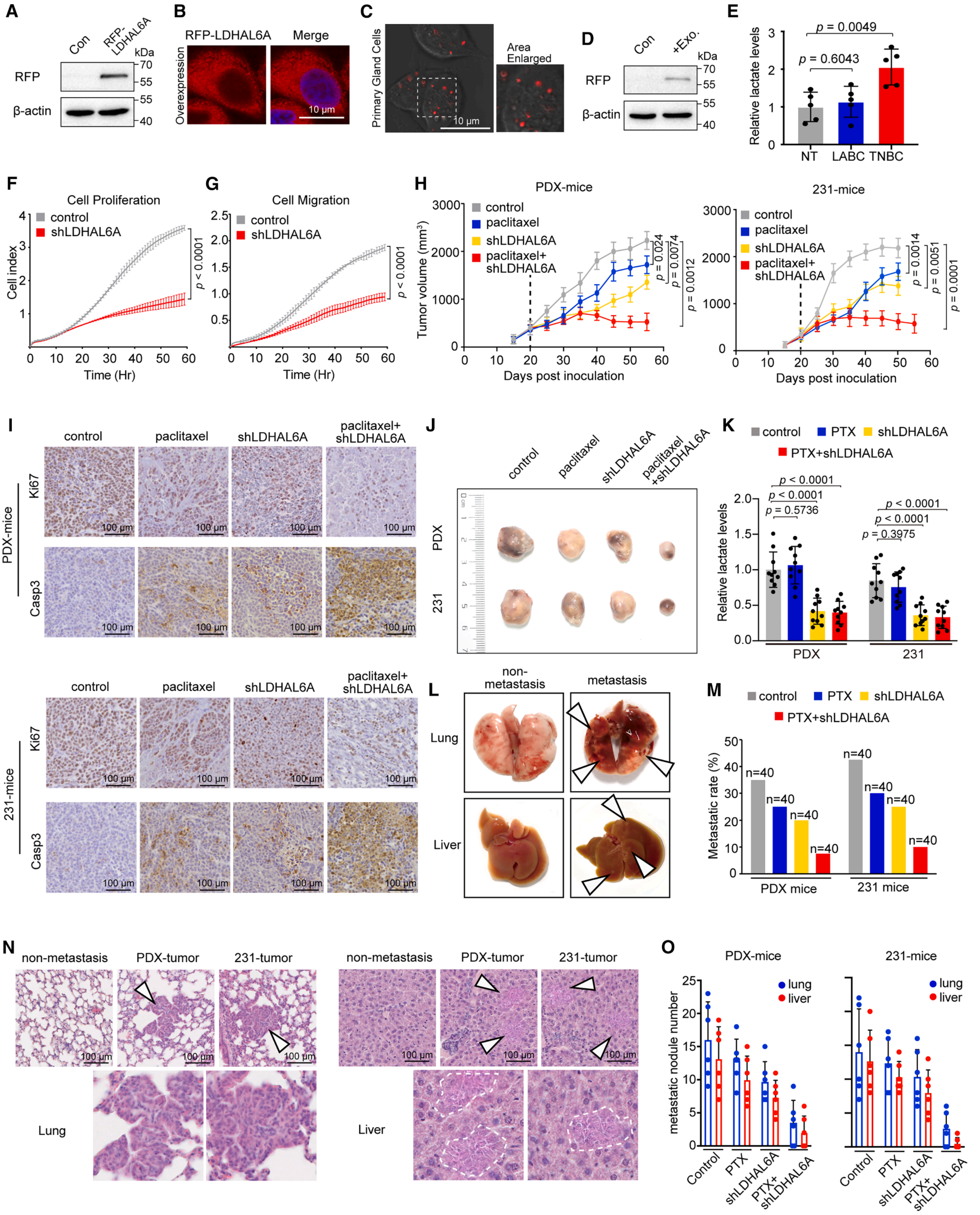
(G) Statistical analysis of the percentage of cells exhibiting orange fluorescence (blue column) and relative fluorescence intensity in these cells (red column) from (F). One hundred recipient cells per group from three independent biological replicates were examined. Data are presented as mean values ± SD.

(H) Ranking plot of proteins in the TNBC_{EXO}-uptake proteome compared with the NT_{EXO}-uptake proteome. Proteins that were at least 2-fold higher in abundance in the TNBC_{EXO}-uptake proteome than in the NT_{EXO}-uptake proteome are cut off by the red line. Representative proteins are labeled in red.

(I) Pathway enrichment analysis of the proteins upregulated in the TNBC_{EXO}-uptake proteome. The *x* axis indicates the enrichment ratio for specific terms. The color schema represents $-\log_{10}$ (*p* value) ranging from low (white) to high (red).

(J) *Z* score plot of proteins involved in the glycolysis/gluconeogenesis pathway. The *x* axis represents the number of standard deviations from each protein's average abundance.

See also Figure S5 and Table S6.



(legend on next page)

DISCUSSION

Proteomics has transformed biomedicine, but *in vitro* isolation erases crucial *in situ* and microenvironmental context.[34,35] Building on important advances, prior strategies based on non-canonical amino acid incorporation and Cre-dependent AAV-mediated APEX delivery enable tissue-specific labeling.[8,10] However, they tend to confine detection to nascent proteins or short labeling windows, with efficiency varying among tissues. By contrast, IVPL provides robust, heritable labeling across a wide range of tissues and cell types without metabolic incorporation or viral delivery, capturing both nascent and long-lived proteins, including low-abundance or exogenous species that persist in pathological niches. These strengths make IVPL a broadly applicable platform for *in vivo* proteomics, especially in long-term disease models like cancer progression and immunomodulation.

During tumor progression, T_{reg} cells infiltrate tumors and reshape the immune microenvironment,[36,37] yet conventional proteomics cannot profile their proteome with *in situ* context. Using IVPL, we decoded the proteome of intratumoral T_{reg} cells within intact solid tumors, identifying 2,787 proteins, of which 991 were absent from *in vitro* T_{reg} datasets,[22] underscoring context-dependent divergence. From this *in vivo* dataset, we extracted membrane receptors and predicted interacting ligands, revealing extensive receptor-ligand landscapes. These insights clarify T_{reg} dynamics in pathological niches (and extend to other infiltrative cells) and inform the design of tumor immunotherapies.

Deploying IVPL to trace exogenous proteomes *in vivo*, we mapped >900 "uptake exo-proteins" originating from LABC and TNBC patient samples, establishing an integrated translational pipeline that bridges clinical material with animal models. Among these, LDHAL6A, a lactate dehydrogenase sustaining glycolytic flux,[38] was the most differentially abundant in the $TNBC_{EXO}$-uptake proteome. It is known that TNBC cells are characterized by an increased rate of glycolysis due to their highly aggressive behaviors compared with other breast cancer subtypes.[39,40] Yet how TNBC cells meet this high level of glycolytic demand among heterogeneous tumor cells remains unclear. Exosomal delivery of the essential glycolysis-related enzyme LDHAL6A to recipient cells provides a plausible explanation: beyond serving as a fuel source, lactate promotes angiogenesis, regulates gene expression, and activates signaling pathways that promote cancer cell proliferation and metastasis.[41–43] Thus, direct introduction of the readily deployable protein LDHAL6A may be the most optimal way for metabolic reprogramming for tumors. Consistently, targeting LDHAL6A combined with standard chemotherapy suppressed TNBC growth and metastasis. Together, these findings establish IVPL as a translatable IVPL platform and reveal actionable protein effectors in recipient cells for targeted cancer therapy.

Limitations of the study

The current approach relies on gene-edited $APEX2^{f/f}$ mice and specific Cre driver lines, which are powerful for cell-type-specific studies but may limit rapid deployment or direct clinical application. Although control experiments showed minimal proteomic changes after short-term H_2O_2 exposure, subtle physiological perturbations cannot be fully excluded. In addition, APEX2 expression is presently confined to intracellular compartments, and developing complementary extracellular labeling variants could further expand the spatial scope of *in vivo* proteomic mapping.

RESOURCE AVAILABILITY

Lead contact

Requests for further information and resources should be directed to and will be fulfilled by the lead contact, Mo Li (limo@hsc.pku.edu.cn).

Figure 6. Targeting LDHAL6A suppresses TNBC growth and metastasis

(A) Overexpression of RFP-LDHAL6A in TNBC primary tumor cells examined by western blotting. β-actin was used as a loading control.

(B) Immunofluorescent detection of RFP-LDHAL6A in TNBC primary tumor cells. Scale bar, 10 μm.

(C) Detection of TNBC exosomes containing RFP-LDHAL6A that were engulfed by recipient cells harvested from mammary glands of mice 24 h post treatment of the exosomes. Scale bar, 10 μm.

(D) Western blot analysis of RFP-LDHAL6A in recipient cells harvested from mammary glands of mice 24 h post treatment of the exosomes.

(E) Intracellular lactate levels in recipient mammary gland cells treated with NT, LABC, or TNBC exosomes, measured from three independent biological replicates.

(F and G) Real-time cell proliferation and migration assays of control and shLDHAL6A TNBC cells. A total of 5,000–10,000 (20,000–40,000 for migration assay) cells were seeded into 16-well plates, and the proliferation and migration were monitored by the xCELLigence RTCA system over a course of 60 h. Data represent three independent biological replicates.

(H) Growth curve of tumors in the PDX model and 231 model upon treatment with paclitaxel (PTX) alone and combined with LDHAL6A inhibition using shLDHAL6A ($n = 9$ mice per group).

(I) Immunohistochemical staining of Ki67 and cleaved caspase3 in tumors harvested from PDX and 231 mice after single and combined treatments. Scale bar, 100 μm.

(J) Tumors harvested from PDX and 231 mice post treatment with paclitaxel alone or combined with LDHAL6A inhibition.

(K) Intracellular lactate levels in tumor tissues from PDX and 231 mice post treatment with paclitaxel alone or combined with LDHAL6A inhibition ($n = 10$ mice per group).

(L) Livers and lungs removed from xenograft mice with or without metastatic lesions. Arrowheads denote representative metastatic nodules in organs.

(M) Metastatic rate in livers and lungs of PDX and 231 mice after different treatments. Number of mice in each group is indicated above the column.

(N) Metastatic areas in livers and lungs are shown by H&E staining. Representative metastatic sites are marked by arrowheads, and their enlarged images are shown in the bottom row. Scale bar, 100 μm.

(O) Number of metastatic nodules in lungs and livers of PDX and 231 mice in different treatment groups ($n = 6$ mice per group). One dot denotes one mouse bearing lung or liver metastatic nodules. Data are presented as mean values ± SD (E–H and K). p values were calculated by unpaired two-tailed Student's t tests (E and K) or two-way ANOVA tests (F–H). $p < 0.05$ was considered statistically significant.

See also Figure S6.

Materials availability

All reagents and materials are listed in the key resources table. Additional materials are available from the corresponding author upon reasonable request.

Data and code availability

- All raw files and search results for mass spectrometry of the proteome have been deposited in ProteomeXchange via iProX (www.iprox.org) with the identification no. PXD053158 (for ProteomeXchange) and IPX0009025000 (for iProX). Original gel/blot images are available on Mendeley Data (https://doi.org/10.17632/8btrkzw4z4.1) as of the date of publication. Raw microscopy data reported in this paper will be shared by the lead contact upon request.
- The analysis codes have been deposited at Zenodo at https://doi.org/10.5281/zenodo.17493194.
- Any additional information required to reanalyze the data reported in this paper is available from the lead contact upon request.

ACKNOWLEDGMENTS

Mass spectrometry experiments were carried out by the State Key Laboratory of Natural and Biomimetic Drugs Large Instrument Technology Platform, Peking University Health Science Center. This work was supported by the National Natural Science Foundation of China (NSFC) (T2225006 to M.L., T2488301 to M.L., 82272948 to M.L., 82371640 to B.X., 82203103 to Y. Xiao, and 82201788 to N.Z.) and the Beijing Municipal Natural Science Foundation (key program Z220011 to M.L.).

AUTHOR CONTRIBUTIONS

M.L. conceived and designed the experiments. Q.W., Y.J., M.B., G.W., Y. Xiao, Y.G., X.L., W.Y., and N.Z. performed the experiments. Y.J. and B.X. operated mass spectrum equipment. Y.J., H.Z., and Y. Xue analyzed bioinformatics data. Y.J. and W.Y. harvested patients' samples. M.L., B.X., P.Z., and H.Y. wrote the manuscript.

DECLARATION OF INTERESTS

The authors declare the following patent related to this work: M.L., B.X., X.L., and W.Y. are listed as inventors on a patent (C.N. patent application no. ZL 2021 1 1358445.3) filed in November 2021 and granted in February 2022.

STAR★METHODS

Detailed methods are provided in the online version of this paper and include the following:

- KEY RESOURCES TABLE
- EXPERIMENTAL MODEL AND STUDY PARTICIPANT DETAILS
  - ○ Cell lines
  - ○ Mice
  - ○ Human samples
- METHOD DETAILS
  - ○ Plasmid construct
  - ○ IVPL substrate design and synthesis
  - ○ IVPL mice generation
  - ○ *In vivo* protein labeling
  - ○ Stability analysis of substrates
  - ○ Pharmacokinetic analysis of oral dosing *in vivo*
  - ○ Mass spectrometry
  - ○ Protein identification
  - ○ Bioinformatics analysis
  - ○ Volume staining and 3D reconstruction
  - ○ Isolation and culture of primary tumor cells
  - ○ Exosome isolation
  - ○ Exosome labeling and injection
  - ○ shRNA lentivirus generation and shRNA knockdown
  - ○ PDX and CDX mouse models
  - ○ Southern blot
  - ○ Western blot
  - ○ Electron microscopy
  - ○ Histologic staining
  - ○ Immunofluorescence
  - ○ Lactate measurements
  - ○ Cell proliferation and migration assays
- QUANTIFICATION AND STATISTICAL ANALYSIS

SUPPLEMENTAL INFORMATION

Supplemental information can be found online at https://doi.org/10.1016/j.molcel.2025.11.008.

Received: June 18, 2024
Revised: September 2, 2025
Accepted: November 5, 2025
Published: December 1, 2025

REFERENCES

1. Larance, M., and Lamond, A.I. (2015). Multidimensional proteomics for cell biology. Nat. Rev. Mol. Cell Biol. *16*, 269–280. https://doi.org/10.1038/nrm3970.
2. Toby, T.K., Fornelli, L., and Kelleher, N.L. (2016). Progress in Top-Down Proteomics and the Analysis of Proteoforms. Annu. Rev. Anal. Chem. (Palo Alto. Calif) *9*, 499–519. https://doi.org/10.1146/annurev-anchem-071015-041550.
3. Lyssiotis, C.A., and Kimmelman, A.C. (2017). Metabolic Interactions in the Tumor Microenvironment. Trends Cell Biol. *27*, 863–875. https://doi.org/10.1016/j.tcb.2017.06.003.
4. Elia, I., and Haigis, M.C. (2021). Metabolites and the tumour microenvironment: from cellular mechanisms to systemic metabolism. Nat. Metab. *3*, 21–32. https://doi.org/10.1038/s42255-020-00317-z.
5. Rhee, H.W., Zou, P., Udeshi, N.D., Martell, J.D., Mootha, V.K., Carr, S.A., and Ting, A.Y. (2013). Proteomic mapping of mitochondria in living cells via spatially restricted enzymatic tagging. Science *339*, 1328–1331. https://doi.org/10.1126/science.1230593.
6. Hung, V., Zou, P., Rhee, H.W., Udeshi, N.D., Cracan, V., Svinkina, T., Carr, S.A., Mootha, V.K., and Ting, A.Y. (2014). Proteomic mapping of the human mitochondrial intermembrane space in live cells via ratiometric APEX tagging. Mol. Cell *55*, 332–341. https://doi.org/10.1016/j.molcel.2014.06.003.
7. Hung, V., Udeshi, N.D., Lam, S.S., Loh, K.H., Cox, K.J., Pedram, K., Carr, S.A., and Ting, A.Y. (2016). Spatially resolved proteomic mapping in living cells with the engineered peroxidase APEX2. Nat. Protoc. *11*, 456–475. https://doi.org/10.1038/nprot.2016.018.
8. Dumrongprechachan, V., Salisbury, R.B., Soto, G., Kumar, M., MacDonald, M.L., and Kozorovitskiy, Y. (2021). Cell-type and subcellular compartment-specific APEX2 proximity labeling reveals activity-dependent nuclear proteome dynamics in the striatum. Nat. Commun. *12*, 4855. https://doi.org/10.1038/s41467-021-25144-y.
9. Soto, J.S., Jami-Alahmadi, Y., Chacon, J., Moye, S.L., Diaz-Castro, B., Wohlschlegel, J.A., and Khakh, B.S. (2023). Astrocyte-neuron subproteomes and obsessive-compulsive disorder mechanisms. Nature *616*, 764–773. https://doi.org/10.1038/s41586-023-05927-7.
10. Alvarez-Castelao, B., Schanzenbächer, C.T., Hanus, C., Glock, C., Tom Dieck, S., Dörrbaum, A.R., Bartnik, I., Nassim-Assir, B., Ciirdaeva, E., Mueller, A., et al. (2017). Cell-type-specific metabolic labeling of nascent proteomes in vivo. Nat. Biotechnol. *35*, 1196–1201. https://doi.org/10.1038/nbt.4016.
11. Krogager, T.P., Ernst, R.J., Elliott, T.S., Calo, L., Beránek, V., Ciabatti, E., Spillantini, M.G., Tripodi, M., Hastings, M.H., and Chin, J.W. (2018).

Labeling and identifying cell-specific proteomes in the mouse brain. Nat. Biotechnol. *36*, 156–159. https://doi.org/10.1038/nbt.4056.

12. van Niel, G., D'Angelo, G., and Raposo, G. (2018). Shedding light on the cell biology of extracellular vesicles. Nat. Rev. Mol. Cell Biol. *19*, 213–228. https://doi.org/10.1038/nrm.2017.125.
13. Tian, T., Wang, Y., Wang, H., Zhu, Z., and Xiao, Z. (2010). Visualizing of the cellular uptake and intracellular trafficking of exosomes by live-cell microscopy. J. Cell. Biochem. *111*, 488–496. https://doi.org/10.1002/jcb.22733.
14. Kalluri, R., and LeBleu, V.S. (2020). The biology, function, and biomedical applications of exosomes. Science *367*, eaau6977. https://doi.org/10.1126/science.aau6977.
15. Zhou, Y., Wang, G., Wang, P., Li, Z., Yue, T., Wang, J., and Zou, P. (2019). Expanding APEX2 Substrates for Proximity-Dependent Labeling of Nucleic Acids and Proteins in Living Cells. Angew. Chem. Int. Ed. Engl. *58*, 11763–11767. https://doi.org/10.1002/anie.201905949.
16. O'Hagan, D. (2008). Understanding organofluorine chemistry. An introduction to the C-F bond. Chem. Soc. Rev. *37*, 308–319. https://doi.org/10.1039/b711844a.
17. Shah, P., and Westwell, A.D. (2007). The role of fluorine in medicinal chemistry. J. Enzyme Inhib. Med. Chem. *22*, 527–540. https://doi.org/10.1080/14756360701425014.
18. Thelen, K., and Dressman, J.B. (2009). Cytochrome P450-mediated metabolism in the human gut wall. J. Pharm. Pharmacol. *61*, 541–558. https://doi.org/10.1211/jpp/61.05.0002.
19. El-Rayes, B.F., Ali, S., Heilbrun, L.K., Lababidi, S., Bouwman, D., Visscher, D., and Philip, P.A. (2003). Cytochrome p450 and glutathione transferase expression in human breast cancer. Clin. Cancer Res. *9*, 1705–1709.
20. Minchinton, A.I., and Tannock, I.F. (2006). Drug penetration in solid tumours. Nat. Rev. Cancer *6*, 583–592. https://doi.org/10.1038/nrc1893.
21. Franzen, O., Gan, L.M., and Bjorkegren, J.L.M. (2019). PanglaoDB: a web server for exploration of mouse and human single-cell RNA sequencing data. Database (Oxford) *2019*, baz046. https://doi.org/10.1093/database/baz046.
22. Cuadrado, E., van den Biggelaar, M., de Kivit, S., Chen, Y.Y., Slot, M., Doubal, I., Meijer, A., van Lier, R.A.W., Borst, J., and Amsen, D. (2018). Proteomic Analyses of Human Regulatory T Cells Reveal Adaptations in Signaling Pathways that Protect Cellular Identity. Immunity *48*, 1046–1059.e6. https://doi.org/10.1016/j.immuni.2018.04.008.
23. Jeppesen, D.K., Fenix, A.M., Franklin, J.L., Higginbotham, J.N., Zhang, Q., Zimmerman, L.J., Liebler, D.C., Ping, J., Liu, Q., Evans, R., et al. (2019). Reassessment of Exosome Composition. Cell *177*, 428–445.e18. https://doi.org/10.1016/j.cell.2019.02.029.
24. Jensen, H.M. (1986). On the origin and progression of human breast cancer. Am. J. Obstet. Gynecol. *154*, 1280–1284. https://doi.org/10.1016/0002-9378(86)90713-1.
25. Figueroa, J.D., Pfeiffer, R.M., Patel, D.A., Linville, L., Brinton, L.A., Gierach, G.L., Yang, X.R., Papathomas, D., Visscher, D., Mies, C., et al. (2014). Terminal duct lobular unit involution of the normal breast: implications for breast cancer etiology. J. Natl. Cancer Inst. *106*, dju286. https://doi.org/10.1093/jnci/dju286.
26. Coleman, O., Henry, M., O'Neill, F., Roche, S., Swan, N., Boyle, L., Murphy, J., Meiller, J., Conlon, N.T., Geoghegan, J., et al. (2018). A Comparative Quantitative LC-MS/MS Profiling Analysis of Human Pancreatic Adenocarcinoma, Adjacent-Normal Tissue, and Patient-Derived Tumour Xenografts. Proteomes *6*, 45. https://doi.org/10.3390/proteomes6040045.
27. Mirhadi, S., Tam, S., Li, Q., Moghal, N., Pham, N.A., Tong, J., Golbourn, B.J., Krieger, J.R., Taylor, P., Li, M., et al. (2022). Integrative analysis of non-small cell lung cancer patient-derived xenografts identifies distinct proteotypes associated with patient outcomes. Nat. Commun. *13*, 1811. https://doi.org/10.1038/s41467-022-29444-9.
28. Rodrigues, G., Hoshino, A., Kenific, C.M., Matei, I.R., Steiner, L., Freitas, D., Kim, H.S., Oxley, P.R., Scandariato, I., Casanova-Salas, I., et al. (2019). Tumour exosomal CEMIP protein promotes cancer cell colonization in brain metastasis. Nat. Cell Biol. *21*, 1403–1412. https://doi.org/10.1038/s41556-019-0404-4.
29. Kugeratski, F.G., Hodge, K., Lilla, S., McAndrews, K.M., Zhou, X., Hwang, R.F., Zanivan, S., and Kalluri, R. (2021). Quantitative proteomics identifies the core proteome of exosomes with syntenin-1 as the highest abundant protein and a putative universal biomarker. Nat. Cell Biol. *23*, 631–641. https://doi.org/10.1038/s41556-021-00693-y.
30. Certo, M., Llibre, A., Lee, W., and Mauro, C. (2022). Understanding lactate sensing and signalling. Trends Endocrinol. Metab. *33*, 722–735. https://doi.org/10.1016/j.tem.2022.07.004.
31. Zong, Z., Xie, F., Wang, S., Wu, X., Zhang, Z., Yang, B., and Zhou, F. (2024). Alanyl-tRNA synthetase, AARS1, is a lactate sensor and lactyltransferase that lactylates p53 and contributes to tumorigenesis. Cell *187*, 2375–2392.e33. https://doi.org/10.1016/j.cell.2024.04.002.
32. Liu, W., Wang, Y., Bozi, L.H.M., Fischer, P.D., Jedrychowski, M.P., Xiao, H., Wu, T., Darabedian, N., He, X., Mills, E.L., et al. (2023). Lactate regulates cell cycle by remodelling the anaphase promoting complex. Nature *616*, 790–797. https://doi.org/10.1038/s41586-023-05939-3.
33. Chen, X., Gu, X., Shan, Y., Tang, W., Yuan, J., Zhong, Z., Wang, Y., Huang, W., Wan, B., and Yu, L. (2009). Identification of a novel human lactate dehydrogenase gene LDHAL6A, which activates transcriptional activities of AP1(PMA). Mol. Biol. Rep. *36*, 669–676. https://doi.org/10.1007/s11033-008-9227-2.
34. Timp, W., and Timp, G. (2020). Beyond mass spectrometry, the next step in proteomics. Sci. Adv. *6*, eaax8978. https://doi.org/10.1126/sciadv.aax8978.
35. Xu, J.Y., Zhang, C., Wang, X., Zhai, L., Ma, Y., Mao, Y., Qian, K., Sun, C., Liu, Z., Jiang, S., et al. (2020). Integrative Proteomic Characterization of Human Lung Adenocarcinoma. Cell *182*, 245–261.e17. https://doi.org/10.1016/j.cell.2020.05.043.
36. Tay, C., Tanaka, A., and Sakaguchi, S. (2023). Tumor-infiltrating regulatory T cells as targets of cancer immunotherapy. Cancer Cell *41*, 450–465. https://doi.org/10.1016/j.ccell.2023.02.014.
37. Togashi, Y., Shitara, K., and Nishikawa, H. (2019). Regulatory T cells in cancer immunosuppression - implications for anticancer therapy. Nat. Rev. Clin. Oncol. *16*, 356–371. https://doi.org/10.1038/s41571-019-0175-7.
38. Bian, X., Jiang, H., Meng, Y., Li, Y.P., Fang, J., and Lu, Z. (2022). Regulation of gene expression by glycolytic and gluconeogenic enzymes. Trends Cell Biol. *32*, 786–799. https://doi.org/10.1016/j.tcb.2022.02.003.
39. Gong, Y., Ji, P., Yang, Y.S., Xie, S., Yu, T.J., Xiao, Y., Jin, M.L., Ma, D., Guo, L.W., Pei, Y.C., et al. (2021). Metabolic-Pathway-Based Subtyping of Triple-Negative Breast Cancer Reveals Potential Therapeutic Targets. Cell Metab. *33*, 51–64.e9. https://doi.org/10.1016/j.cmet.2020.10.012.
40. Wu, Q., Ba-Alawi, W., Deblois, G., Cruickshank, J., Duan, S., Lima-Fernandes, E., Haight, J., Tonekaboni, S.A.M., Fortier, A.M., Kuasne, H., et al. (2020). GLUT1 inhibition blocks growth of RB1-positive triple negative breast cancer. Nat. Commun. *11*, 4205. https://doi.org/10.1038/s41467-020-18020-8.
41. Végran, F., Boidot, R., Michiels, C., Sonveaux, P., and Feron, O. (2011). Lactate influx through the endothelial cell monocarboxylate transporter MCT1 supports an NF-κB/IL-8 pathway that drives tumor angiogenesis. Cancer Res. *71*, 2550–2560. https://doi.org/10.1158/0008-5472.CAN-10-2828.
42. Roland, C.L., Arumugam, T., Deng, D., Liu, S.H., Philip, B., Gomez, S., Burns, W.R., Ramachandran, V., Wang, H., Cruz-Monserrate, Z., et al. (2014). Cell surface lactate receptor GPR81 is crucial for cancer cell survival. Cancer Res. *74*, 5301–5310. https://doi.org/10.1158/0008-5472.CAN-14-0319.

43. Harmon, C., Robinson, M.W., Hand, F., Almuaili, D., Mentor, K., Houlihan, D.D., Hoti, E., Lynch, L., Geoghegan, J., and O'Farrelly, C. (2019). Lactate-Mediated Acidification of Tumor Microenvironment Induces Apoptosis of Liver-Resident NK Cells in Colorectal Liver Metastasis. Cancer Immunol. Res. *7*, 335–346. https://doi.org/10.1158/2326-6066.CIR-18-0481.

44. Shen, B., Zhang, W., Zhang, J., Zhou, J., Wang, J., Chen, L., Wang, L., Hodgkins, A., Iyer, V., Huang, X., et al. (2014). Efficient genome modification by CRISPR-Cas9 nickase with minimal off-target effects. Nat. Methods *11*, 399–402. https://doi.org/10.1038/nmeth.2857.

45. Hwang, W.Y., Fu, Y., Reyon, D., Maeder, M.L., Tsai, S.Q., Sander, J.D., Peterson, R.T., Yeh, J.R.J., and Joung, J.K. (2013). Efficient genome editing in zebrafish using a CRISPR-Cas system. Nat. Biotechnol. *31*, 227–229. https://doi.org/10.1038/nbt.2501.

46. Norrander, J., Kempe, T., and Messing, J. (1983). Construction of improved M13 vectors using oligodeoxynucleotide-directed mutagenesis. Gene *26*, 101–106. https://doi.org/10.1016/0378-1119(83)90040-9.

47. Stewart, S.A., Dykxhoorn, D.M., Palliser, D., Mizuno, H., Yu, E.Y., An, D.S., Sabatini, D.M., Chen, I.S.Y., Hahn, W.C., Sharp, P.A., et al. (2003). Lentivirus-delivered stable gene silencing by RNAi in primary cells. Rna *9*, 493–501. https://doi.org/10.1261/rna.2192803.

48. Xie, B., Pu, Y., Yang, F., Chen, W., Yue, W., Ma, J., Zhang, N., Jiang, Y., Wu, J., Lin, Y., et al. (2022). Proteomic Mapping and Targeting of Mitotic Pericentriolar Material in Tumors Bearing Centrosome Amplification. Cancer Res. *82*, 2576–2592. https://doi.org/10.1158/0008-5472.CAN-22-0225.

49. Schwanhäusser, B., Busse, D., Li, N., Dittmar, G., Schuchhardt, J., Wolf, J., Chen, W., and Selbach, M. (2011). Global quantification of mammalian gene expression control. Nature *473*, 337–342. https://doi.org/10.1038/nature10098.

50. Giansanti, P., Samaras, P., Bian, Y., Meng, C., Coluccio, A., Frejno, M., Jakubowsky, H., Dobiasch, S., Hazarika, R.R., Rechenberger, J., et al. (2022). Mass spectrometry-based draft of the mouse proteome. Nat. Methods *19*, 803–811. https://doi.org/10.1038/s41592-022-01526-y.

51. Shao, X., Liao, J., Li, C., Lu, X., Cheng, J., and Fan, X. (2021). CellTalkDB: a manually curated database of ligand-receptor interactions in humans and mice. Brief. Bioinform. *22*, bbaa269. https://doi.org/10.1093/bib/bbaa269.

52. Li, W., Germain, R.N., and Gerner, M.Y. (2019). High-dimensional cell-level analysis of tissues with Ce3D multiplex volume imaging. Nat. Protoc. *14*, 1708–1733. https://doi.org/10.1038/s41596-019-0156-4.

53. Xie, B.T., Liang, X.L., Yue, W., Ma, J.H., Li, X.Y., Zhang, N., Wang, P., Liu, C., Shi, X.M., Qiao, J., et al. (2021). Targeting cytokinesis bridge proteins to kill high-CIN type tumors. Fundam. Res. *1*, 752–766. https://doi.org/10.1016/j.fmre.2021.08.015.

54. Théry, C., Witwer, K.W., Aikawa, E., Alcaraz, M.J., Anderson, J.D., Andriantsitohaina, R., Antoniou, A., Arab, T., Archer, F., Atkin-Smith, G.K., et al. (2018). Minimal information for studies of extracellular vesicles 2018 (MISEV2018): a position statement of the International Society for Extracellular Vesicles and update of the MISEV2014 guidelines. J. Extracell. Vesicles *7*, 1535750. https://doi.org/10.1080/20013078.2018.1535750.

STAR★METHODS

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Antibodies		
Rabbit polyclonal anti-GFP	Abcam	Cat#ab6556; RRID: AB_305564
Mouse recombinant monoclonal Anti-CD63	Abcam	Cat#ab193349; RRID: AB_3095976
Rabbit monoclonal anti-TSG101	Abcam	Cat#ab125011; RRID: AB_10974262
Mouse monoclonal anti- beta Actin	Abcam	Cat#ab8226; RRID: AB_306371
Rabbit polyclonal anti-Ki67	Novus	Cat#NB500-170; RRID: AB_10001977
Rabbit polyclonal anti-CUL4B	Proteintech	Cat#12916-1-AP; RRID: AB_2086699
Mouse monoclonal anti-RFP (clone 6G6)	Proteintech	Cat#6G6; RRID: AB_2631395
Rabbit monoclonal anti-cleaved caspase-3 (Asp175) (clone 5A1E)	Cell Signaling Technology	Cat#9664; RRID: AB_2070042
Rabbit monoclonal anti-CD9 (clone D8O1A)	Cell Signaling Technology	Cat#13174; RRID: AB_2798139
Alexa Fluor® 647 anti-mouse CD31	BioLegend	Cat#102416; RRID: AB_493410
Rabbit polyclonal anti-UBAP2L	Thermo Fisher Scientific	Cat#PA5-36998; RRID: AB_2553856
Rabbit polyclonal Anti-APLP2	Thermo Fisher Scientific	Cat#PA5-117016; RRID: AB_2901646
Rabbit polyclonal Anti- RPL7	Thermo Fisher Scientific	Cat#PA5-36571; RRID: AB_2553591
Rabbit polyclonal Anti- CDV3	Thermo Fisher Scientific	Cat#PA5-56251; RRID: AB_2639631
Rabbit polyclonal Anti- DYNC1I2	Thermo Fisher Scientific	Cat#PA5-75613; RRID: AB_2719341
HRP goat anti-mouse IgG (H+L) secondary antibody	Thermo Fisher Scientific	Cat#32430; RRID: AB_1185566
HRP goat anti-rabbit IgG (H+L) secondary antibody	Thermo Fisher Scientific	Cat#31466; RRID: AB_10960844
Biological samples		
Human breast tissue samples (tumor and normal)	This paper	Table S5
Chemicals, peptides, and recombinant proteins		
Hoechst 33342	Beyotime	Cat#C1027
HRP Streptavidin Protein	Thermo Fisher Scientific	Cat#21127
Alexa Fluor® 555 streptavidin	Thermo Fisher Scientific	Cat#S21381
Paclitaxel	Selleck	Cat#S1150
Critical commercial assays		
L-Lactate Assay Kit (Colorimetric/Fluorometric)	Abcam	Cat#ab65330
PKH26 Red Fluorescent Cell Linker Kit	Sigma Aldrich	Cat#PKH26GL
Pierce™ BCA Protein Assay Kits	Thermo Fisher Scientific	Cat#23227
mMESSAGE mMACHINE™ T7 ULTRA Transcription Kit	Thermo Fisher Scientific	Cat# AM1345
MEGAclear™ Transcription Clean-Up Kit	Thermo Fisher Scientific	Cat#AM1908
MEGAshortscript™ T7 Transcription Kit	Thermo Fisher Scientific	Cat#AM1354
Deposited data		
Proteomics data	This paper	ProteomeXchange: PXD053158; iProX: IPX0009025000
Original gels/blots images	This paper	Mendeley Data: https://doi.org/10.17632/8btrkzw4z4.1
In vitro T_{reg} proteome	Cuadrado et al.[22]	PubMed: 29752063;DOI: 10.1016/j.immuni.2018.04.008
Experimental models: Cell lines		
MDA-MB-231 cells	ATCC	CRM-HTB-26

(*Continued on next page*)

Continued

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Experimental models: Organisms/strains		
Mouse: APEX2-EGFP$^{\text{f/f or f/+}}$	This paper	N/A
Mouse: C57BL/6JGpt-H11$^{\text{em1Cin(Vil1-icre)}}$/Gpt (Villin-Cre)	GemPharmatech	Cat#T004714
Mouse: C57BL/6JGpt-H11$^{\text{em1Cin(MMTV-LTR-iCre)}}$/Gpt (MMTV-Cre)	GemPharmatech	Cat#T005635
Mouse: C57BL/6JGpt-Tg(Foxp3-iCre)97/Gpt (Foxp3-Cre)	GemPharmatech	Cat#T010853
Mouse: C57BL/6JGpt-H11$^{\text{em1Cin(Camk2a-iCre-polyA)}}$/Gpt (Camk2a-Cre)	GemPharmatech	Cat#T054722
Mouse: C57BL/6 J	Peking University Health Science Center	N/A
Mouse: ICR	Peking University Health Science Center	N/A
Mouse: NOD-SCID	Peking University Health Science Center	N/A
Mouse: BALB/c Nude	Peking University Health Science Center	N/A
Oligonucleotides		
guide RNA for Rosa26 locus	This paper	method details
Primers for genotyping assay	This paper	method details
Primers for Southern Blot	This paper	method details
Oligonucleotides of knock-down	This paper	method details
Recombinant DNA		
pUC57-sgRNA expression vector	Shen et al.[44]	Addgene #51132
Cas9 expression vector	Hwang et al.[45]	Addgene #42251
pUC19 plasmid	Norrander et al.[46]	Addgene #50005
pcDNA3-mRFP plasmid	Doug Golenbock	Addgene #13032
pLKO.1 plasmid	Stewart et al.[47]	Addgene #8453
Software and algorithms		
IVPL analysis code	This paper	GitHub: https://github.com/rebeccajiang611/IVPL_Analysis; Zenodo: https://doi.org/10.5281/zenodo.17493194
Proteome Discoverer v2.4	Thermo Fisher Scientific	https://www.thermofisher.cn/cn/zh/home/industrial/mass-spectrometry/liquid-chromatography-mass-spectrometry-lc-ms/lc-ms-software/multi-omics-data-analysis/proteome-discoverer-software.html
GraphPad Prism v9.0	GraphPad Software, Inc.	RRID:SCR_002798
SPSS Statistics 27	IBM Corporation	RRID:SCR_016479
RTCA Software v2.0	Agilent	RRID:SCR_014821
Imaris	Bitplane	RRID:SCR_007370
Other		
xCELLigence RTCA-DP system	Agilent	https://www.agilent.com.cn/zh-cn/product/cell-analysis/real-time-cell-analysis/rtca-analyzers/xcelligence-rtca-dp-cell-invasion-migration-741226
BioTek Synergy H1 Multimode Reader	Agilent	https://www.agilent.com.cn/en/product/microplate-instrumentation/microplate-readers/multimode-microplate-readers/biotek-synergy-h1-multimode-reader-1623193

(*Continued on next page*)

Continued

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Zeiss LSM880 microscope	ZEISS	https://www.zeiss.com/microscopy/en/products/light-microscopes/confocal-microscopes.html
Olympus BX51 microscope	OLYMPUS	https://evidentscientific.com/en/microscope-resource/knowledge-hub/anatomy/bx51cutaway
JEOL JEM-1400 transmission electron microscopy	JEOL	RRID:SCR_020179
NanoSight NS300	Malvern Panalytical	https://www.malvernpanalytical.com/en/support/product-support/nanosight-range/nanosight-ns300

EXPERIMENTAL MODEL AND STUDY PARTICIPANT DETAILS

Cell lines

MDA-MB-231 cells were cultured in RPMI-1640 medium, supplemented with 10% FBS, at 37°C in a humidified 5% CO_2 incubator. The cell line was obtained from ATCC and was routinely tested for mycoplasma. Cell line identity was confirmed by short tandem repeat (STR) profiling.

Mice

Mice handling and experiments were conducted in accordance with policies promulgated by the Ethics Committee of the Peking University Health Science Center (SA2022065).

For the generation of APEX2-EGFP$^{f/f}$ (IVPL) mice, male and female C57BL/6J mice aged 6–8 weeks were used for CRISPR/Cas9-mediated gene targeting in zygotes. 6–8-week-old pseudopregnant ICR female mice were used as foster mothers. C57BL/6J and ICR mice were obtained from the Peking University Health Science Center.

To establish tissue-specific IVPL mouse models, 6–8-week-old APEX2-EGFP$^{f/f}$ mice were crossed with the following 6–8-week-old Cre-driver strains purchased from GemPharmatech: (i) C57BL/6JGpt-H11$^{em1Cin(Vil1\text{-}icre)}$/Gpt (Cat#T004714) to generate IVPL^{INTESTINE} mice, (ii) C57BL/6JGpt-H11$^{em1Cin(MMTV\text{-}LTR\text{-}iCre)}$/Gpt (Cat#T005635) to generate IVPL^{BREAST} mice, (iii) C57BL/6JGpt-Tg (Foxp3-iCre)97/Gpt (Cat#T010853) to generate IVPL^{Treg} mice, and (iv) C57BL/6JGpt-H11$^{em1Cin(Camk2a\text{-}iCre\text{-}polyA)}$/Gpt (Cat#T054722) to generate IVPL^{BRAIN} mice.

For xenograft tumor modeling, 5–6-week-old female NOD-SCID and BALB/c nude mice were used to establish patient-derived xenograft (PDX) tumor model, while 4-week-old female BALB/c nude mice were used to establish cell-derived xenograft (CDX) tumor model. NOD-SCID and BALB/c nude mice were obtained from the Peking University Health Science Center.

All animals were housed in a specific pathogen-free barrier facility under controlled environmental conditions (temperature, 20–26 °C; relative humidity, 30–70%) with a 12-hour light/dark cycle (lights on at 6 a.m., off at 6 p.m.), and provided ad libitum access to food and water.

Human samples

Tumor samples obtained from breast cancer patients and normal breast tissues obtained from reduction mammoplasty procedures were collected from adult Asian female patients in the Department of General Surgery, Peking University Third Hospital. All procedures were approved by the Medical Science Research Ethics Committee of Peking University Third Hospital (IRB00006761-M2019471). Written informed consent was obtained from all participants prior to sample collection. Detailed clinical information for the human samples is provided in Table S5.

METHOD DETAILS

Plasmid construct

For the generation of Cre-dependent APEX2-EGFP mouse, the guide RNA for the *Rosa26* locus was designed at https://crispor.gi.ucsc.edu/. The guide RNA (5’-CTCCAGTCTTTCTAGAAGAT-3’) was then cloned into the pUC57-sgRNA expression vector (Addgene #51132). The Cas9 expression vector was purchased from Addgene (#42251). The donor plasmid, which contained the loxP-Stop-loxP-APEX2-EGFP sequence flanked by $\sim$2000-bp homology arms of the *Rosa26* locus, was constructed using the pUC19 plasmid (Addgene #50005). To overexpress proteins (APLP2, CUL4B, RPL7, CDV3, DYNC1I2, UBAP2L, PGK1, and GPX1) with both N-terminal EGFP and C-terminal RFP in LABC primary tumor cells, the coding sequences of the eight human genes were subcloned into pEGFP-C1 plasmid containing an N-terminal EGFP tag. Subsequently, each EGFP-gene tandem sequence was subcloned into pcDNA3 with a C-terminal RFP tag (Addgene #13032). To knockdown the expression of *LDHAL6A*, oligo encoding *LDHAL6A* shRNA was cloned into the pLKO.1 plasmid (Addgene #8453). The sequences used were as follows: shLDHAL6A top strand, 5’-CCGGGCCTGGAAGTTGAGTGGATTTCTCGAGAAATCCACTCAACTTCCAGGCTTTTTG-3’; shLDHAL6A bottom strand,

5’-AATTCAAAAAGCCTGGAAGTTGAGTGGATTTCTCGAGAAATCCACTCAACTTCCAGGC-3’; scramble (non-targeting) control shRNA top strand, 5’-CCGGCCTAAGGTTAAGTCGCCCTCGCTCGAGCGAGGGCGACTTAACCTTAGGTTTTTG-3’; scramble (non-targeting) control shRNA bottom strand, 5’-AATTCAAAAACCTAAGGTTAAGTCGCCCTCGCTCGAGCGAGGGCGACTTAACCT TAGG-3’.

IVPL substrate design and synthesis

The synthesis strategies for the two newly designed small molecules, Btn-Ph-3F and Btn-O-Ph, are illustrated in Figures S2A and S2B, respectively. The synthesis of Btn-Ph-3F commenced with the preparation of Intermediate A1. Initially, 3-Fluoro-L-tyrosine (200 mg, 0.94 mmol) was suspended in methanol (5 mL), and the mixture was stirred in an ice bath. Thionyl chloride (366 μL, 5.04 mmol) was added dropwise to the mixture. The mixture was incubated on ice for 30 minutes and then warmed to room temperature. After reacting overnight, the solvent was removed under reduced pressure to yield Intermediate A1 as a solid, which was used directly in the next step without further purification. Next, Intermediate A1 (142 mg, 0.67 mmol) was dissolved in 10 mL DMF. Triethylamine (556 μL, 4.02 mmol) was slowly added to the solution, followed by biotin-NHS (228 mg, 0.67 mmol). The mixture was stirred at room temperature overnight. After solvent removal via rotary evaporation, the residue was purified using a C18 reverse phase column on a semi-preparative HPLC to afford Btn-Ph-3F as a white solid (165 mg, yield 56%). ^{1}H-NMR (400 MHz, d_6-DMSO): 9.66 (s, 1H), 8.22 (d, 1H), 6.97 (m, 1H), 6.82 (m, 2H), 6.42 (s, 1H), 6.39 (s, 1H), 4.41 (m, 1H), 4.31 (m, 1H), 4.11 (m, 1H), 3.60 (s, 3H), 3.05 (m, 1H), 2.85 (m, 2H), 2.77 (m, 1H), 2.57 (d, 1H), 2.06 (m, 2H), 1.44 (m, 4H), 1.22 (m, 2H). ^{13}C-NMR (100 MHz, d_6-DMSO): 25.61, 28.41, 35.19, 16.11, 40.33, 52.27, 53.88, 55.85, 59.67, 61.47, 117.01, 117.89, 125.54, 128.96, 143.73, 149.86, 152.25, 163.21, 172.65, 172.70. Calculated molecular weight for $C_{20}H_{27}FN_3O_5S$: $[M+H]^+$: 440.17; found: 440.42.

The synthesis of Btn-O-Ph involved three major steps, starting with the formation of two intermediates, B1 and B2. Initially, 2-bromoethanamine hydrobromide (2.25 g, 11 mmol) and triethylamine (2.07 mL, 15 mmol) were stirred in 1,4-dioxane (20 mL) within an ice bath, and di-tert-butyl dicarbonate (2.62 g, 12 mmol) was added dropwise. After 16 hours, the solvent was removed under reduced pressure. The resulting residue was dissolved in 50 mL of ethyl acetate and sequentially washed with saturated solutions of NH_4Cl, $NaHCO_3$, and NaCl. After drying over Na_2SO_4, the organic layer was concentrated to yield Intermediate B1 as a yellow oil, which was then used directly in the next synthesis step without further purification. For the synthesis of Intermediate B2, hydroquinone (1.0 g, 9.1 mmol) and KOH (0.51 g, 9.1 mmol) were stirred in methanol (50 mL) at room temperature for 30 minutes before Intermediate B1 (2.03 g, 9.1 mmol) was introduced dropwise. The solution was refluxed for 24 hours, followed by evaporation of the solvent under reduced pressure to produce a dark brown solid. This solid was resuspended in 50 mL of DCM, washed with a saturated NaCl solution, dried over Na_2SO_4, filtered, and then concentrated. Purification was carried out using silica column chromatography (EtOAc = 1:20 to 1:5), producing Intermediate B2 as a yellow oil (550 mg, 24% yield). Finally, Intermediate B2 (150 mg, 0.59 mmol) was dissolved in a 1:1 v/v mixture of TFA and DCM and stirred at room temperature for 2 hours. After vacuum evaporation, the residue was directly used in the subsequent reaction. The residue and triethylamine (489 μL, 3.54 mmol) were dissolved in DMF (10 mL), and biotin-NHS (201 mg, 0.59 mmol) was added. The mixture was stirred overnight at room temperature and purified using a C18 reverse phase column on semi-preparative HPLC to yield Btn-O-Ph as a light yellow solid (156 mg, 70% yield). ^{1}H-NMR (400 MHz, d_6-DMSO): 8.92 (s, 1H), 8.02 (t, 1H), 6.65-6.76 (m, 4H), 6.45 (s, 1H), 6.38 (s, 1H), 4.29 (m, 1H), 4.10 (m, 1H), 3.86 (t, 2H), 3.35 (m, 2H), 3.04 (m, 1H), 2.79 (dd, 1H), 2.58 (d, 1H), 2.09 (t, 2H), 1.24-1.64 (m, 6H). ^{13}C-NMR (100 MHz, d_6-DMSO): 25.73, 28.47, 28.60, 35.53, 38.76, 40.32, 55.88, 59.66, 61.48, 67.35, 115.95, 116.17, 151.69, 151.75, 163.21, 172.83. Calculated molecular weight for $C_{18}H_{26}N_3O_4S$: $[M+H]^+$: 380.16; found: 380.42.

IVPL mice generation

The Cre-dependent APEX2-EGFP mouse was generated by CRISPR/Cas9-assisted gene targeting in zygotes as follows. Briefly, pronuclear stage zygotes were obtained by mating C57BL/6 J males with superovulated C57BL/6 J females. A 1-2 pL mixture containing 100 ng/μL Cas9 mRNA, 50 ng/μL sgRNA, and 200 ng/μL targeting vector was microinjected into the male pronucleus of the zygote using a FemtoJet 4i and Injectman 4 microinjection system (Eppendorf) fitted with Femtotips II (Eppendorf, 5242957000). Cas9 mRNA was prepared from the AgeI-linearized Cas9 expression vector (Addgene #42251) by *in vitro* transcription using the mMESSAGE mMACHINE™ T7 ULTRA Transcription Kit (Thermo Fisher Scientific, Cat#AM1345) and purified using the MEGAclear™ Transcription Clean-Up Kit (Thermo Fisher Scientific, Cat#AM1908). Rosa26-specific sgRNA was prepared from the ApaI-linearized sgRNA expression vector (Addgene #51132) using the MEGAshortscript™ T7 Transcription Kit (Thermo Fisher Scientific, Cat#AM1354) and purified with the MEGAclear™ Transcription Clean-Up Kit. After overnight culture in KSOM medium at 37 °C under 5% CO_2, the injected embryos that had reached a 2-cell stage were implanted into oviducts of pseudopregnant foster ICR female mice. For germline transmission of targeted alleles, founder mice (F0) were crossed to WT C57BL/6 mice for the generation of F1 offspring. Positively heterozygous mice were identified by PCR targeting 5’ and 3’ inserting sites in the gene. PCR primers for 5’ inserting sites are: F1, 5’- TACGCCACAGGGAGTCCAAGAATG-3’; R1, 5’- GATGGGGAGAGTGAAGCAGAACG-3’. PCR primers for 3’ inserting sites are F2, 5’-CTGCTGTCCATTCCTTATTCCATAG-3’; R2, 5’-CTGGAAATCAGGCTGCAAATCTC-3’, resulting in a 2.7 kb and a 2.5 kb PCR product, respectively. The PCR products were then purified and sequenced by the Sanger method. After genotyping and Southern blot analysis against F1 mice, the heterozygous F1 mice were intercrossed for the generation of homozygous (LSL-APEX2-EGFP$^{f/f}$) mice (F2, i.e. IVPL mice). Primers used for genotyping of F2 IVPL mice were as follows: F3, 5’-CACTTGCTCTCC CAAAGTCGCTC-3’; R3, 5’-ATACTCCGAGGCGGATCACAA-3’; F4, 5’-GCATCTGACTTCTGGCTAATAAAG-3’. PCR with F3/R3 yields a 453 bp band from the WT allele, while F4/R3 yields a 644 bp band from the targeted allele. Heterozygous mice produce both bands.

In vivo protein labeling

For the IVPL^{INTESTINE} model, after anesthetizing the mice, we administered five different reaction substrates (Btn-Ph-3F, Btn-O-Ph, Btn-An, Btn-Nap, and BP), each at a concentration of 500 μM dissolved in phosphate-buffered saline (PBS), to five separate groups by oral gavage. Each substrate was allowed to incubate *in vivo* for 60 minutes to ensure optimal tissue penetration and substrate absorption. A PBS-only control was also included. To trigger the *in vivo* labeling, 1 mM H_2O_2 was injected through the same gavage route. After a 90-second reaction period, the intestines were promptly excised from the anesthetized mice, immediately subjected to two rapid washes with a quencher solution (30 s × 2), and then processed either for lysis and protein harvesting or for fixation for subsequent staining. Lysis buffer: 50 mM Tris-HCl [pH 7.4], 150 mM NaCl, 0.5% SDS, 0.5% sodium deoxycholate, 1% Triton X-100, 1× protease inhibitor cocktail, 10 mM sodium azide, 10 mM sodium ascorbate, 5 mM Trolox and 20 mM DTT. After gently sonicated, the lysates were cleared by centrifugation at 12,000 rpm for 5 min at 4°C, and then ten times volume of ice-cold methanol was added followed by precipitation at -80°C for 6 h. Proteins were harvested by centrifugation at 4,000 rpm for 30 min at 4 °C, and then were re-dissolved in 1% SDS RIPA buffer. 4 mg protein of each sample (in 0.2% SDS RIPA buffer) was incubated with 150 μL of streptavidin-coated magnetic beads slurry under rotation for 2 h at room temperature. The beads were subsequently washed twice with 1 mL of RIPA lysis buffer, once with 1 mL of 1 M KCl, once with 1 mL of 0.1 M Na_2CO_3, once with 1 mL of 2 M urea in 10 mM Tris-HCl (pH 8.0), and twice with 1 mL RIPA lysis buffer. Biotinylated proteins were then eluted by boiling the beads in 60 μL of protein loading buffer supplemented with 20 mM DTT and 2 mM biotin. The enriched protein sample was prepared for mass spectrometry.

For the IVPL^{BREAST} model, we administered the same set of five substrates and a PBS control directly into the nipple of anesthetized mice using a microinjection needle. Each substrate, at a concentration of 500 μM, was preloaded for 45 minutes *in vivo*. The *in vivo* labeling was then initiated by injecting 1 mM H_2O_2 into the same site and allowing a 90-second reaction. The mammary glands were harvested post-reaction, immediately subjected to two rapid washes with a quencher solution (30 s × 2), and then lysed for protein extraction and enrichment, following the same protocol used for intestinal epithelial cells. In addition, an independent validation group was processed using an extended quenching step (20 minutes) prior to lysis to assess the robustness of the standard protocol. Comparative proteomic analysis revealed minimal differences between the standard and extended quenching conditions, indicating that the original quenching strategy was sufficient (Figures S3N and S3O).

For the IVPL^{Treg} model, each of the five reaction substrates and a PBS control was directly injected into the tumor mass at a concentration of 500 μM using a fine gauge needle. After a 90-minute incubation period to accommodate the dense tumor matrix, 1 mM H_2O_2 was injected into the tumor for a 90-second *in vivo* labeling. The tumors were subsequently harvested, immediately subjected to two rapid washes with a quencher solution (30 s × 2), and then processed using a similar lysis and protein extraction protocol as described for other tissues.

Stability analysis of substrates

Stability analysis of five substrates (Btn-Ph-3F, BP, Btn-Nap, Btn-An, and Btn-O-Ph) was conducted on tissue lysates from mouse intestine and liver. To prepare the lysates, tissues were first harvested, washed with PBS, and homogenized in RIPA lysis buffer containing protease inhibitors to prevent protein degradation. After sonication, the homogenates were then centrifuged at 12,000 rpm for 10 minutes at 4°C to remove the cellular debris, and the supernatants were collected. Each substrate was added to the lysates to achieve a final concentration of 1 ng/μL and incubated at 37 °C for specified durations. Sample extraction and compound quantification were performed in accordance with the methods described in the pharmacokinetic analysis.

Pharmacokinetic analysis of oral dosing *in vivo*

Pharmacokinetic analyses were conducted based on time-course concentration data (0.5h, 1h, 1.5h, 2h, 4h. 6h) collected from plasma and multiple tissues, including liver, spleen, lung, kidney, small intestine, colon, heart, and brain, following intravenous (IV, 20 mg/kg), intraperitoneal (IP, 50 mg/kg), and oral (PO, 50 mg/kg) administration of the tested compounds. Each concentration at a given time point represents the mean of three biological replicates. Samples from untreated mice were harvested to serve as the 0 h control. For organ samples, tissues were homogenized at low temperatures in an 80% methanol solution (methanol/water, v/v, 4/1) at a ratio of 1:40 (20 mg tissue to 800 μL solution), and then centrifuged at 15,000 rpm for 30 minutes. After resting for 10 minutes, the supernatant was collected and concentrated at low temperatures followed by reconstitution in 200 μL of 50% methanol solution (methanol/water, v/v, 1/1). The reconstituted solution was centrifuged for another two rounds under the same conditions, with the supernatant collected for Liquid Chromatography-Tandem Mass Spectrometry (LC-MS/MS) analysis. Blood samples were obtained into heparinized tubes, followed by centrifugation at 3,000 rpm and 4 °C for 15 minutes. The plasma supernatant was then precipitated with five volumes of methanol and centrifuged at 15,000 rpm and 4 °C for 30 minutes. Then the supernatant was collected and subjected to a second low-temperature centrifugation for measurements. LC-MS/MS analysis was performed using Waters ACQUITY UPLC i-Class Plus for chromatographic separation and Waters Xevo TQ-S for mass spectrometry, with data captured and processed using Waters MassLynx V4.2 software and Waters TargetLynx XS V.2 software, respectively. For absolute quantification, a standard curve was prepared by spiking various concentrations of the standard (0, 0.05, 0.1, 0.2, 1, 2, 5, 10, 20, 50, 100, 200, 500 ng/mL) into homogenates of blank tissues or blood. The maximum concentration (C_{max}) and time to reach maximum concentration (T_{max}) were directly obtained from the observed data points. The area under the concentration–time curve from 0 to 6 hours ($AUC_{0-6}h$) was calculated using the linear trapezoidal rule. The terminal half-life ($t_{1/2}$) was estimated by linear regression of the log-transformed concentration values over time points within the terminal phase (typically 2–6 hours).

Mass spectrometry

An intestinal segment (~ 2 cm) from one mouse, two mammary glands from one mouse, and one tumor each from two mice were used to generate *in situ* proteomes of intestinal epithelial cells, mammary epithelial cells, and tumor-infiltrating T_{reg} cells, respectively. The sample for mass spectrometry was prepared according to previous studies.[5,48] Briefly, biotinylated proteins eluted from streptavidin beads were separated on SDS-PAGE. After staining with Coomassie G-250 and destaining with water, each lane of each sample was manually cut into 6 fractions. Each fraction was dried with a vacuum concentrator after in-gel digestion into peptides.

The extracted peptides were dissolved with 25 μL of Solvent A (0.1% formic acid in water) and separated by a 120 min gradient elution at a flow rate 0.300 μL/min with the Thermo Ultimate 3000 nano-UPLC system which was directly interfaced with the Thermo Fusion LUMOS mass spectrometer. The analytical column was an Acclaim PepMap RSLC column (75 μm ID, 250 mm length, C18) coupled with a trap column (Acclaim PepMap™ 100, 75 μm ID, 20mm length, C18). Mobile phase A consisted of 0.1% formic acid, and mobile phase B consisted of 100% acetonitrile and 0.1% formic acid. The gradient elution conditions (120 min) were: 3% B for 5min; 3% to 22% for 75min; 22% to 35% for 12 min; 35% to 90% for 11 min; 90% for 6 min; 90% to 3% for 1 min; 3% for 10min. The Fusion LUMOS mass spectrometer was operated in the data-dependent acquisition mode using Xcalibur 4.1.50 software and there is a single full-scan mass spectrum in the Orbitrap (375-1500 m/z, 60,000 resolution) followed by data-dependent MS/MS scans. The most intense ions selected under top-speed mode were isolated in Quadrupole with a 1.6 m/z window and activated by high energy collision dissociation (HCD) with a normalized collision energy of 30%, then detected in the Orbitrap at resolution of 15,000. The Maximum Injection Time was set to dynamic mode and the AGC Target was set to Standard mode.

Protein identification

Raw mass spectrometry data were processed using Proteome Discoverer (Thermo Fisher Scientific, version 2.4). Proteomic data from each sample were queried against the Homo sapiens database (for uptake exosomal proteome) and the Mus musculus database in the protein sequence database UniProt (https://www.uniprot.org). For uptake exosomal proteomes, only PSMs matching to one unambiguous protein was allowed to ensure the retention of only species-specific peptides. If evidence for both human and mouse peptides from an orthologous protein were detected, then peptides that cannot distinguish between the two species were ignored.[26,27] Precursor mass tolerance: 20 ppm; fragment mass tolerance: 0.05 Da; variable modifications: oxidation of methionine, carbamidomethylation of cysteine, and acetylation of protein N-terminal; maximum two missed cleavage sites were allowed. Proteins with at least one unique peptide with <1% FDR and sum PEP score > 120 were kept for uptake exo-proteomes. For protein quantification, the intensity-based absolute quantification (iBAQ) intensities were used to represent the abundances of proteins,[49] and were further normalized to a fraction of total (FOT) based on the total iBAQ intensities of all proteins within each sample.[50] The FOT values were multiplied by 10^5 for the ease of visualization. Replicates were performed in each group and proteins that were present in at least two samples from one group were selected for further analysis.

Bioinformatics analysis

Pearson correlation coefficients were calculated to estimate the similarities of samples within and between groups using ggpairs() function in R package *GGally* (version 2.2.1) or ggcorrplot() function in R package *ggcorrplot* (version 0.1.4.1). The distribution of $\log_2$-transformed protein abundances in each sample was visualized as violin plots generated using R package *ggplot2* (version 3.5.1). For Gene Set Enrichment Analysis (GSEA), we carried out our T_{reg} proteomics datasets on Java based GSEA software (version 4.3.2) with the reference gene sets from Molecular signatures database (MSigDB, https://www.gsea-msigdb.org/gsea/msigdb/index.jsp) and the gene set identified from the literature.[22] To conduct ligand-receptor (LR) interaction analysis, we obtained mouse LR pairs from CellTalkDB (https://xomics.com.cn/celltalkdb/), one of the most extensive collection of LR pairs, featuring 2033 mouse LR interactions identified through extensive data mining and manual verification.[51] To display the interactions between twenty receptors with the highest abundances identified in our tumor-infiltrating T_{reg} proteomes and other potential ligands from cells in the tumor microenvironment, we used R package *igraph* (version 2.0.3) and *ggraph* (version 2.2.1) to create a bipartite LR communication network plot. Heatmaps were visualized by R package *pheatmap* (version 1.0.12). Functional enrichment cluster of uniquely identified *in situ* 991 proteins in T_{reg} proteomes was visualized using GOCluster() function in R package *GOplot* (version 1.0.2). Enrichment network was plotted using ClueGo plugin of Cytoscape software. Circos plot and bubble plots were created using R package *circlize* (version 0.4.16) and *ggplot2* (version 3.5.1). Proteins with log2 (fold change) ≥ 1 and p value < 0.05 (two-tailed t-test) were considered significantly differential abundant proteins. To further ensure robustness, we performed downstream enrichment analysis and multiple layers of annotation. Gene Ontology (GO) and KEGG pathway analyses were conducted using WebGestalt (http://www.webgestalt.org/) and clusterProfiler (R package version 4.10.1), with enrichment results subjected to redundancy reduction (e.g., by grouping semantically similar GO terms and applying leading-edge clustering). Enriched terms were visualized using network-based layouts to highlight representative biological processes. Additionally, exosomal protein candidates were cross-referenced with the ExoCarta and Vesiclepedia databases to validate their exosome relevance. For glycolysis/gluconeogenesis–associated proteins, standardized z-scores were calculated as (individual abundance − global mean abundance)/standard deviation to allow comparison across samples.

Volume staining and 3D reconstruction

The immunostaining of solid tumors and mammary glands was performed according to the protocol described previously.[52] Briefly, the tumors and mammary glands were stripped out from the mouse and fixed overnight at 4°C using BD Cytofix on a shaker. The fixed mammary tissue was washed for 1 h with washing buffer (1× PBS supplemented with 0.3% Triton X-100 and 0.5% 1-thioglycerol) three times. Subsequently, the tissue was blocked and permeabilized at 37°C for 24 h with blocking buffer (1× PBS plus 0.3% Triton X-100, 1% BSA, and 1% normal mouse serum). Following incubation with primary antibody of CD31 (1:200, BioLegend, Cat#102416; RRID: AB_493410) at 37°C for 4 d, the tissues were washed with washing buffer at 37°C for 12 h, then immersed in fresh washing buffer and incubated at room temperature for 3 d. During the incubation period, the buffer was changed every 12 h. After that, the tissues were incubated with fluorescent secondary antibody (1:400) and Hoechst 33342 (10 μg/mL, Beyotime, Cat#C1027) at 37°C for 4 d. After three more washes, the tissue was cleared by immersion in clearing solution (2.75 mL of 40% N-methyl acetamide, 4 g of Histodenz, and 5 μL Triton X-100) with a refractive index of 1.495-1.505, for 24 h at room temperature before being imaged. The fluorescence images were captured using Lightsheet Z.1 microscope with a 5× detection objective (NA = 1.0). The images were treated with dual side fusion and converted to Imaris files using Imaris File Converter. The 3D surface-rendered images and Spots-rendered images were analyzed and assembled using the Imaris software (Bitplane, RRID: SCR_007370) to reconstruct the tissues.

Isolation and culture of primary tumor cells

Fresh breast cancer tissue samples were collected from patients who underwent tumor resection surgery. Isolation and culture of primary tumor cells were performed according to our previous study.[53] In specific, the tumor specimen was minced into pieces smaller than 1 mm^3 followed by digestion in 1× collagenase/hyaluronidase buffer (STEMCELL Technologies) at 37 °C with agitation for 6 h. Further digestion was performed by gentle pipetting in trypsin (0.25%) and then in a solution of dispase (5 units/ml) and DNase I (0.05 mg/ml; STEMCELL Technologies) for 5-10 min. By filtration through a 40-μm filter, single-cell suspensions were obtained. These cells were then seeded at a density of $0.5\text{-}1\times10^5$/well onto 6-well plates coated with collagen I and cultured for 36-72 h. The culture medium composes of DMEM:F12 supplemented with 5% FBS, penicillin/streptomycin (1%), gentamycin (0.2%), EGF (10 ng/ml), adenine (20 μg/ml), cholera toxin (10 ng/ml), HEPES (15 mM), insulin (5 μg/ml), hydrocortisone (0.32 μg/ml) and ROCK inhibitor (5 μM). After stable adhesion, cells could be further cultured or frozen in FBS/DMSO (9:1) with 5 μM ROCK inhibitor in liquid nitrogen.

Exosome isolation

Exosome isolation was performed by standard differential ultracentrifugation.[54] Cells at 80% confluence were cultured in media supplemented with 10% exosome-depleted FBS, obtained by removing bovine exosomes through overnight centrifugation at 100,000 × g (Beckman 32Ti rotor). After incubation for 48 h, the conditioned media were collected and sequentially centrifuged at 3,000 × g for 20 min and 10,000 × g for 40 min at 4 °C to remove cell debris and microvesicles. The resulting supernatants were filtered through a 0.22 μm filter (Millipore), and exosomes were harvested by ultracentrifugation at 100,000 × g for 70 min at 4 °C. The exosome pellets were then washed with cold PBS and re-ultracentrifuged at 100,000 × g for 70 min at 4 °C. Finally, the pelleted exosomes were resuspended in PBS and either used immediately or stored at −80 °C. The exosomal protein concentration was measured by BCA assay (Pierce, Thermo Fisher Scientific, Cat#23227). The size and concentration of isolated exosomes were assessed with nanoparticle tracking analysis using the NS300 nanoparticle characterization system (NanoSight, Malvern Instruments). Exosomes were further verified by western blot analysis and electron microscopy.

Exosome labeling and injection

To visualize exosomal distribution *in vivo*, purified exosomes derived from tumors of breast cancer patients were labeled with the fluorescent PKH26 (Sigma Aldrich, Cat#PKH26GL) according to the manufacturer's protocol. Briefly, exosomes in PBS were diluted to 1ml using Diluent C and mixed with 1ml Diluent C containing 6μl of PKH26 dye. After 5 min of incubation at room temperature, the mixture was ultracentrifuged at 130,000×g at 4 °C for 2 h, and the obtained pellet (PKH26-labeled exosomes) was dissolved in PBS. Finally, the labeled exosomes were intravenously injected into $\text{IVPL}^{\text{BREAST}}$ mice via tail vein (10 μg of exosomes in 100μl of PBS/mouse) for subsequent experiments.

shRNA lentivirus generation and shRNA knockdown

For shRNA lentivirus generation, the pLKO.1 plasmid comprising shRNA was co-transfected with the packaging plasmids (psPAX2 and pMD2.G) into HEK293T cells using Lipofectamine 3000™ according to the manufacturer's protocol. 6 h after transfection, cells were washed and further cultured in fresh growth culture media for another 48 h. Then the culture media containing viral particles were harvested and centrifuged at 3,000 ×*g* for 5 min to remove the cell debris and filtered by a 0.45-μm filter. The viral supernatant was further concentrated with a Centricon Plus-20 Centrifugal Filter at 4,000 ×*g*. The concentrated lentivirus supernatant was aliquoted and kept at -80 °C before use. To knockdown interested genes in primary tumor cells, 10^5 cells were seeded onto 6-well plates and incubated at 37 °C with 5% CO_2 until reaching 30-40% confluence. The concentrated viral supernatant was added into the culture medium at a multiplicity of infection (MOI) of 20 for 16 h.

PDX and CDX mouse models

The PDX-mouse model was established with our previous report with minor modifications.[53] In brief, 5–8 mm^3 of tumor tissue harvested from TNBC patients was subcutaneously implanted into the flanks of 5∼6 weeks old NOD-SCID mice for the first generation of tumor growth, followed by weekly measurement of tumor volume with calipers. When tumor volume reached 1000–1,500 mm^3, the mice were sacrificed and the tumor was removed. The fresh tumor was cut into small pieces (4-8 mm^3) and re-implanted into the flanks of 5–6 weeks old BALB/c Nude mice for the second generation of tumor growth. For the LDHAL6A knockdown *in vivo*, 4×10^7 copies of gene targeting shRNA lentivirus or scramble shRNA lentivirus was injected into the tumor. For CDX mouse (231-mouse) model established, 1×10^6 MDA-MB-231 breast cancer cells in 100 μL mixture (Matrigel : PBS = 1 : 1) were injected orthotopically into the 4th mammary fat pad of each 4-week-old female BALB/c nude mouse. The tumor growth was monitored for 40–60 days when the volume reached an ethical limitation of around 2000 mm^3. Tumor volume was calculated as the formula: $(\text{shortest diameter})^2\times\text{longest diameter}\times0.5$. Paclitaxel (Selleck Cat#S1150) was used as a standard chemotherapeutic agent.

Southern blot

Southern blotting was performed with DIG-High prime DNA labeling and detection starter kit II (Roche cat. no. 11585614910) as follows. The external 5’ probe and external 3’ probe were amplified with primers 5’ probe F/R (F: 5’-AAACGTGGAGTAGGCAATACCCAGG-3’, R: 5’-AAAGAAGGGTCACCTCAGTCTCCCT-3’) and 3’ probe F/R (F: 5’-TTCTGGGCAGGCTTAAAGGCTAAC-3’, R: 5’-AGGAGCGGGAGAAATGGATATGAAG-3’), respectively. All the probes were purified and labeled with DIG-dUTP. Mouse genomic DNA was isolated, purified, and digested by BamHI and BstEII, respectively. DNA fragments were separated on a 0.7% agarose gel. The gel was washed two times for 5 min in ddH_2O, incubated two times for 15 min in 0.125 M HCl for depurination, washed two times for 5 min in H_2O, and was finally denatured by being incubated two times for 15 min in 0.5 M NaOH/1.5 M NaCl. DNA was transferred and fixed to positively-charged nylon membranes. DNA hybridization and detection were performed as described in the kit instructions. The hybridized probes were immunodetected with anti-digoxigenin-AP and then visualized with the chemiluminescence substrate CSPD. Hybridization of BamHI-digested genomic DNA with the 5′ probe resulted in a WT 5.83 kb-band and a knockin 4.80 kb-band. Hybridization of BstEII-digested genomic DNA with the 3′ probe resulted in a WT 4.77 kb-band and a knockin 9.67 kb-band.

Western blot

Total protein was extracted from cell or tissue lysate by RIPA buffer. Protein samples were electrically transferred to polyvinylidene fluoride membranes after being separated using sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE). Following transfer, the membranes were blocked in Tris buffered saline with Tween 20 (TBST) containing 5% skim milk for 2 h before being incubated with indicated primary antibodies (1:500–1:1000 dilution) overnight at 4°C. After washing in TBST three times, the membranes were then incubated with HRP-conjugated secondary antibodies (1:1000 dilution) for 1 h at 37 °C. After another round of washing in TBST three times, protein bands were visualized using an enhanced chemiluminescence detection system (Amersham Biosciences). The following primary antibodies and reagents were used for Western blotting: anti-CD63 (Abcam, Cat#ab193349, RRID: AB_3095976), anti-TSG101 (Abcam, Cat#ab125011, RRID: AB_10974262), anti-CD9 (Cell Signaling Technology, Cat#13174, RRID: AB_2798139), anti-β-actin (Abcam, Cat#ab8226, RRID: AB_306371), anti-CUL4B (Proteintech, Cat#12916-1-AP, RRID: AB_2086699), anti-UBAP2L (Thermo Fisher, Cat# PA5-36998, RRID: AB_2553856), anti-APLP2 (Thermo Fisher, Cat# PA5-117016, RRID: AB_2901646), anti-RPL7 (Thermo Fisher, Cat#PA5-36571, RRID: AB_2553591), anti-CDV3 (Thermo Fisher, Cat#PA5-56251, RRID: AB_2639631), anti-DYNC1I2 (Thermo Fisher, Cat#PA5-75613, RRID: AB_2719341), and anti-RFP (ChromoTek, Cat#6G6, RRID: AB_2631395). HRP-conjugated reagents included HRP–Streptavidin protein (Thermo Fisher, Cat#21127), HRP–goat anti-mouse IgG (H+L) secondary antibody (Thermo Fisher, Cat#32430, RRID: AB_1185566), and HRP–goat anti-rabbit IgG (H+L) secondary antibody (Thermo Fisher, Cat#31466, RRID: AB_10960844).

Electron microscopy

For morphological characterization, fresh exosome samples in PBS at a concentration of 0.1 μg/μL were placed on formvar-coated copper grids for 1 min. 1% osmium containing 1.5% potassium ferrocyanide in 0.1 M PB buffer was then loaded onto the samples and settled for 30 min at room temperature. After three washes in ddH_2O, samples were performed with negative staining with 2% uranyl acetate. Following air drying, the grids carrying samples were observed under JEOL JEM-1400 transmission electron microscopy (JEOL, RRID:SCR_020179) at 120 kV.

Histologic staining

Tissues were fixed in 10% neutral buffered formalin for 24 h followed by dehydration under gradual serial of ethanol (75%, 85%, 90%, 100%), clearing and paraffin-embedding. Then the tissues were cut into 5 μm sections on polylysine-coated slides used for hematoxylin and eosin (HE) staining and immunohistochemistry. The dilutions of anti-Ki-67 (Novus, Cat#NB500-170, RRID: AB_10001977) and cleaved-caspase3 (Cell Signaling Technology, Cat#9664, RRID: AB_2070042) antibodies were 1:1000, and 1:500, respectively. Staining was imaged and evaluated with an Olympus BX51 microscope and DP73 CCD photographic system.

Immunofluorescence

Cells were fixed for 30 min using 4% paraformaldehyde in PBS, then permeabilized in 0.5% Triton-X-100 for 25 min at room temperature. After blocking with 1% bovine serum albumin-supplemented PBS for 1 h, cells were incubated with the indicated primary antibodies (1:100–1:500) at 4 °C overnight. After washing three times in PBS containing 0.1% Tween 20 and 0.01% Triton-X 100, cells were incubated with an appropriate fluorescent secondary antibody for 1 h at room temperature. After several washes, cells were stained with Hoechst 33342 (10 μg/mL) for 15 min. Samples were eventually put on glass slides and imaged under a confocal laser scanning microscope at ×63/1.40 (Carl Zeiss 880). The following antibodies and reagents were used for immunofluorescence staining: Alexa Fluor® 555–conjugated streptavidin (Thermo Fisher Scientific, Cat#S21381) and anti-GFP antibody (Abcam, Cat#ab6556, RRID: AB_305564).

Lactate measurements

Cell culture medium first underwent deproteinization using a 10kD Spin Column (Abcam, Cat#ab93349). Following this, the collected supernatant was diluted 5 to 10 fold. Tumor samples were harvested in ice-cold PBS and lysed in assay buffer before measuring lactate concentration. Lactate levels were then measured colorimetrically according to the manufacturer's instructions (Abcam, Cat#ab65330). Specifically, 10 μL of the supernatant from each sample was transferred into a 96-well plate, which was combined with OxiRed Probe, enzyme mix, and assay buffer followed by incubation for 30 minutes at room temperature in the dark. The absorbance was measured at 570nm using a spectrophotometer (BioTek Synergy H1, Agilent). The lactate quantification was normalized to cell number or tissue weight.

Cell proliferation and migration assays

The proliferation and migration abilities of control and shLDHAL6A cells were assessed using the xCELLigence RTCA-DP system (Agilent). For the proliferation assay, E-16-well plates were added with 20 μL complete medium for baseline measurements. Meanwhile, cells were digested with trypsin, suspended in serum-free medium, and 5,000-10,000 cells in 80 μL complete medium were seeded in each well and monitored for 60 h. For the migration assay, the lower chambers of CIM plate 16 were filled with complete medium, and the upper chambers were pre-equilibrated with 30 μL serum-free medium for 1 h. Then, 20,000-40,000 cells in 100 μL serum-free medium were seeded in the upper chambers and allowed to culture for 60 h. Cell index values were detected every 30 min, and the data were collected and analyzed with RTCA software (version 2.0).

QUANTIFICATION AND STATISTICAL ANALYSIS

Statistical analyses were performed using GraphPad Prism (version 9.0, GraphPad Software, USA) and SPSS Statistics 27 (IBM Corporation, USA). All experiments were carried out in three independent biological replicates unless indicated otherwise. Means and standard deviations (SD) were plotted in Figures. Two-tailed unpaired Student's *t*-test and Welch's t-test were used for two group comparisons, whereas multiple-comparison two-way analysis of variance (ANOVA) was used for between-group comparisons over time. Kaplan-Meier survival curves were generated using Survival section of GraphPad Prism, and *p* values were calculated using log-rank test. p value < 0.05 was regarded as statistically significant. Statistical analysis details were indicated in Figures or Figure legends.