

CELL BIOLOGY

Spatially resolved cell polarity proteomics of a human epiblast model

Sicong Wang^{1,2}, Chien-Wei Lin³, Amber E. Carleton², Chari L. Cortez⁴, Craig Johnson⁴, Linnea E. Taniguchi², Nikola Sekulovski², Ryan F. Townshend⁴, Venkatesha Basrur⁵, Alexey I. Nesvizhskii^{5,6}, Peng Zou⁷, Jianping Fu^{1,4,8*}, Deborah L. Gumucio^{4,9*}, Mara C. Duncan^{4*}, Kenichiro Taniguchi^{2,4,10*}

Critical early steps in human embryonic development include polarization of the inner cell mass, followed by formation of an expanded lumen that will become the epiblast cavity. Recently described three-dimensional (3D) human pluripotent stem cell–derived cyst (hPSC-cyst) structures can replicate these processes. To gain mechanistic insights into the poorly understood machinery involved in epiblast cavity formation, we interrogated the proteomes of apical and basolateral membrane territories in 3D human hPSC-cysts. APEX2-based proximity biotinylation, followed by quantitative mass spectrometry, revealed a variety of proteins without previous annotation to specific membrane subdomains. Functional experiments validated the requirement for several apically enriched proteins in cyst morphogenesis. In particular, we found a key role for the AP-1 clathrin adaptor complex in expanding the apical membrane domains during lumen establishment. These findings highlight the robust power of this proximity labeling approach for discovering novel regulators of epithelial morphogenesis in 3D stem cell–based models.

INTRODUCTION

As the human embryo implants into the uterine wall, the pluripotent cells of the inner cell mass become radially organized, undergo apicobasal polarization, and initiate lumen formation, forming a cyst with an expanded central lumen called the epiblast cavity (1–3). Although this polarized epiblast cyst is the essential substrate for subsequent critical developmental events (e.g., amniogenesis and gastrulation) (1, 2), we know little about the molecular machinery that is required for polarized cyst organization, because of the inaccessibility of peri-implantation human embryos. In vitro culture of intact human blastocysts is difficult, and only a small percentage of these embryos develop to the epiblast cavity stage (4–7). Recently, however, a three-dimensional (3D) culture system was defined, using primed human pluripotent stem cells (hPSC), that recapitulates epiblast-like luminal cavity formation, forming cysts of pluripotent cells (termed hPSC-cysts) (4, 8–10), enabling investigation of the essential polarization machinery at molecular and cellular levels.

Apicobasal organization of the epiblast cyst must be maintained by the polarized distribution of proteins (e.g., transmembrane proteins and intracellular cell polarity determinants) to the apical and basolateral surfaces, a process that is driven by intracellular

trafficking molecules. While several important studies have recently emerged (11–14), studies of such trafficking events are generally challenging at the global molecular level, because it is difficult to separate membrane fractions that are specific to apical or basolateral surfaces and such fractions are also contaminated by mitochondrial, endoplasmic reticulum, and Golgi membrane proteins (15). However, advances in spatial proteomics using domain-specific biotin labeling hold a promise to uncover novel players in polarization and trafficking. Labeling of the outer surface proteomes of apical and basal membranes in polarized epithelial monolayers of Madin-Darby canine kidney (MDCK) cells revealed novel aspects of polarity in this well-studied system (16). A recent major advance in spatial proteomics is the targeting of “promiscuous” tagging enzymes, such as the engineered ascorbate peroxidase APEX2, to subcellular regions of interest by fusing it to known regionally localized proteins. This strategy was recently used to examine the proteomes associated with Pals1 and Par3 in a 2D MDCK monolayer system (17). In the presence of H₂O₂ and biotin phenol (BP), APEX2 triggers the generation of free BP radicals, which quickly react with neighboring proteins. Because BP radicals display a short half-life and limited diffusion distance (18, 19), this APEX2-based approach enables robust proximity biotin labeling of endogenous proteins within targeted cellular compartments, a critical requirement for analysis of dynamic processes in the context of the epiblast-like cyst.

Here, we used APEX2-based spatial biotinylation, followed by quantitative mass spectrometry (MS) and bioinformatic analyses, to systematically profile apical and basolateral proteomes in live hPSC-derived 3D human epiblast-like cysts. Our goal was to find key molecules required for polarization and the critical trafficking machinery involved in this process. Analyses were done on 3-day (d3) cysts, which are well organized and contain an expanded central lumen. APEX2-tagged constructs were generated to enable the delivery of APEX2 to apical and basolateral territories. Ratio-metric analyses identified a cohort of apically and basolaterally enriched proteins, and microscopic analysis was used to validate the

¹Department of Mechanical Engineering, University of Michigan, Ann Arbor, MI 48109, USA. ²Department of Cell Biology, Neurobiology and Anatomy, Medical College of Wisconsin, Milwaukee, WI 53226, USA. ³Division of Biostatistics, Medical College of Wisconsin, Milwaukee, WI 53226, USA. ⁴Department of Cell and Developmental Biology, University of Michigan Medical School, Ann Arbor, MI 48109, USA. ⁵Department of Pathology, University of Michigan Medical School, Ann Arbor, MI 48109, USA. ⁶Department of Computational Medicine and Bioinformatics, University of Michigan Medical School, Ann Arbor, MI 48109, USA. ⁷College of Chemistry and Molecular Engineering, Peking University, Beijing 100871, China. ⁸Department of Biomedical Engineering, University of Michigan, Ann Arbor, MI 48109, USA. ⁹Department of Internal Medicine, University of Michigan Medical School, Ann Arbor, MI 48109, USA. ¹⁰Department of Pediatrics, Medical College of Wisconsin, Milwaukee, WI 53226, USA.

*Corresponding author. Email: ktaniguchi@mcw.edu (K.T.); mcduncan@med.umich.edu (M.C.D.); dgumucio@med.umich.edu (D.L.G.); jpfu@umich.edu (J.F.)

localization of selected proteins. Functional analysis of several apical proteins without prior known apical localization or function revealed novel aspects of the cell polarization machinery in hPSC epiblast-like cysts. In particular, we uncovered a previously uncharacterized role for the clathrin adaptor protein complex 1 (AP-1) in expansion of the luminal surface. Collectively, these analyses provide an extensive resource that illuminates the composition of apical and basolateral membrane territories in an *in vitro* 3D model of human epiblast morphogenesis, permitting further mechanistic examination of candidate proteins involved in cell polarization and epithelial remodeling during early human development.

RESULTS

Establishing an APEX2-based proximity labeling platform for proteomic mapping of apical and basolateral cell membrane territories in hPSC-cysts

To establish a global molecular view of the intracellular cell polarity machinery that governs polarized epiblast cyst organization, we used APEX2-based proximity labeling. Doxycycline (DOX)-inducible constructs of APEX2-tagged fusion proteins were generated using two known apical (PODXL and EZR) and two basolateral (ATP1B1 and SDC1) proteins (Fig. 1A). PODXL is a transmembrane protein that localizes to the apical surface of hPSC-cysts (20) and to the epiblast cavity of *in vitro* cultured human embryos (21). To capture proteins localized in close proximity to the intracellular apical surface, we fused APEX2 to the C terminus (facing the cell interior) of PODXL (PODXL-APEX2) (Fig. 1B). As an additional apical bait, APEX2 was fused to the C terminus of EZR (EZR-APEX2), an actin scaffold protein with subapical membrane localization (22). Previous studies showed that, in addition to the apical membrane, PODXL shows vesicular localization, while EZRIN is largely apical membrane specific, although nuclear and basolateral localization has also been documented (23–26). Furthermore, PODXL is known to bind to EZRIN (27). To label the basolateral domain, we fused APEX2 at N and C termini, respectively, to ATP1B1 and SDC1, transmembrane proteins that localize basolaterally (Fig. 1, A and B) (28, 29). Last, APEX2 was fused to a nuclear export sequence (APEX2-NES) (30) to allow biotin tagging of unpolarized cytoplasmic proteins for subsequent background subtraction during ratiometric analysis (Fig. 1, A and B).

To verify expression, correct localization, and labeling functionality, we established stable cell lines with each APEX2 construct. hPSC-cysts were then generated using a previously published protocol; mature luminal cysts were seen by 48 hours (d2) after the removal of the ROCK inhibitor Y-27632 (Fig. 1C) (20, 31, 32). To induce transgene expression, DOX (2 μ g/ml) was added at d2 and APEX2 labeling was performed 24 hours later (d3) (Fig. 1C). All five constructs (marked by FLAG) showed induced expression and expected enrichment at respective apical, basolateral, or cytoplasmic domains when expressed in hPSC-cysts (Fig. 1D and fig. S1, A and B). To verify the specificity of APEX2-mediated apical and basolateral membrane biotinylation, we first performed fluorescent staining with fluorophore-conjugated streptavidin that specifically recognizes biotin. d3 hPSC-cysts were incubated with BP for 1 hour, followed by 1.5-min H_2O_2 labeling to allow full penetration of H_2O_2 to the cyst structure (Fig. 1B). We observed extensive biotinylation in the expected membrane territories (Fig. 1E), but not in samples lacking BP or H_2O_2 (fig. S1C). Although labeling was concentrated

at the expected surface for each tag, all tags showed diffuse cytosolic localization of biotin signal, likely due to diffusion of the peroxidase activated label, or diffusion of cytosolic proteins labeled by the APEX2 enzyme. A streptavidin blot of whole-cell lysates revealed that each APEX2 fusion construct biotinylated a wide range of endogenous proteins compared to control samples without H_2O_2 labeling, and different constructs (apical versus basolateral) exhibited distinct biotinylation patterns (Fig. 1F and fig. S1D). Finally, successful pulldown of biotinylated proteins was confirmed by Coomassie staining of post-enrichment eluates (Fig. 1G). These analyses demonstrate that this selective APEX2-based apical and basolateral surface biotinylation can label and enrich cell surface proteins from specific plasma membrane subdomains in live, intact 3D tissues.

Proteomic mapping of apical and basolateral membrane territories in hPSC-cysts

Previous work has shown that a chemical tandem mass tag (TMT)-based quantitative ratiometric APEX2-tagging strategy can improve the specificity of protein identification (33, 34). In this study, our regions of interest (apical and basolateral territories) could also contain proteins that are freely diffusible cytosolic components, which become biotinylated by virtue of being close to the APEX2 enzyme at the time of its activation. In addition, BP radicals generated by apical or basolateral membrane targeted APEX2 could diffuse into the cytosol and biotinylate proteins that lack membrane proximity localization. Both possibilities would result in a pool of weakly biotinylated cytosolic proteins that are essentially “background.” In contrast, bona fide apically or basolaterally enriched proteins should be heavily biotinylated by the apical or basolateral APEX2 construct compared to the cytoplasmic APEX2 construct. Given that each territory should display a distinct protein abundance, ratiometric quantitative analysis can be performed for each protein by calculating the ratio of TMT values (e.g., apical/cytoplasmic and basolateral/cytoplasmic). To design ratiometric APEX2 tagging for quantitative proteomics experiments, we therefore used TMT tagging of 10 APEX2-labeled samples (Fig. 2A): three replicates biotinylated at the apical membrane territory (2 \times PODXL and 1 \times EZR), three replicates biotinylated at the basolateral membrane (2 \times ATP1B1 and 1 \times SDC1), three replicates biotinylated in the cytoplasm [nuclear export signal (NES)], and one nonbiotinylated negative control (NES without H_2O_2 labeling) (Fig. 2B). The use of two different baits (PODXL and EZR for apical and ATP1B1 and SDC1 for basolateral) for each membrane subdomain in this proteomic analysis allowed us to further enrich for relevant proteins based on identification in multiple replicates.

We prepared proteomic samples from d3 hPSC-cysts, which display a well-organized apicobasally polarized epithelial cyst structure with a central lumen (Figs. 1B and 2A). Each of the 10 samples was separately lysed and enriched for biotinylated proteins using streptavidin beads, followed by on-bead digestion and TMT labeling. All samples were then combined, fractionated, and analyzed as a pooled mixture by liquid chromatography (LC)-MS/MS (Fig. 2A). We identified 4454 unique proteins with two or more unique peptides and a false discovery rate (FDR) lower than 1% (Fig. 2C and table S1A). To assign proteins into subcellular territories, proteins were further filtered using a previously established ratiometric strategy (34–36) in which the TMT ratio of each protein reflects its differential enrichment in the experimental group versus negative control lacking H_2O_2 [see Materials and Methods for ratiometric

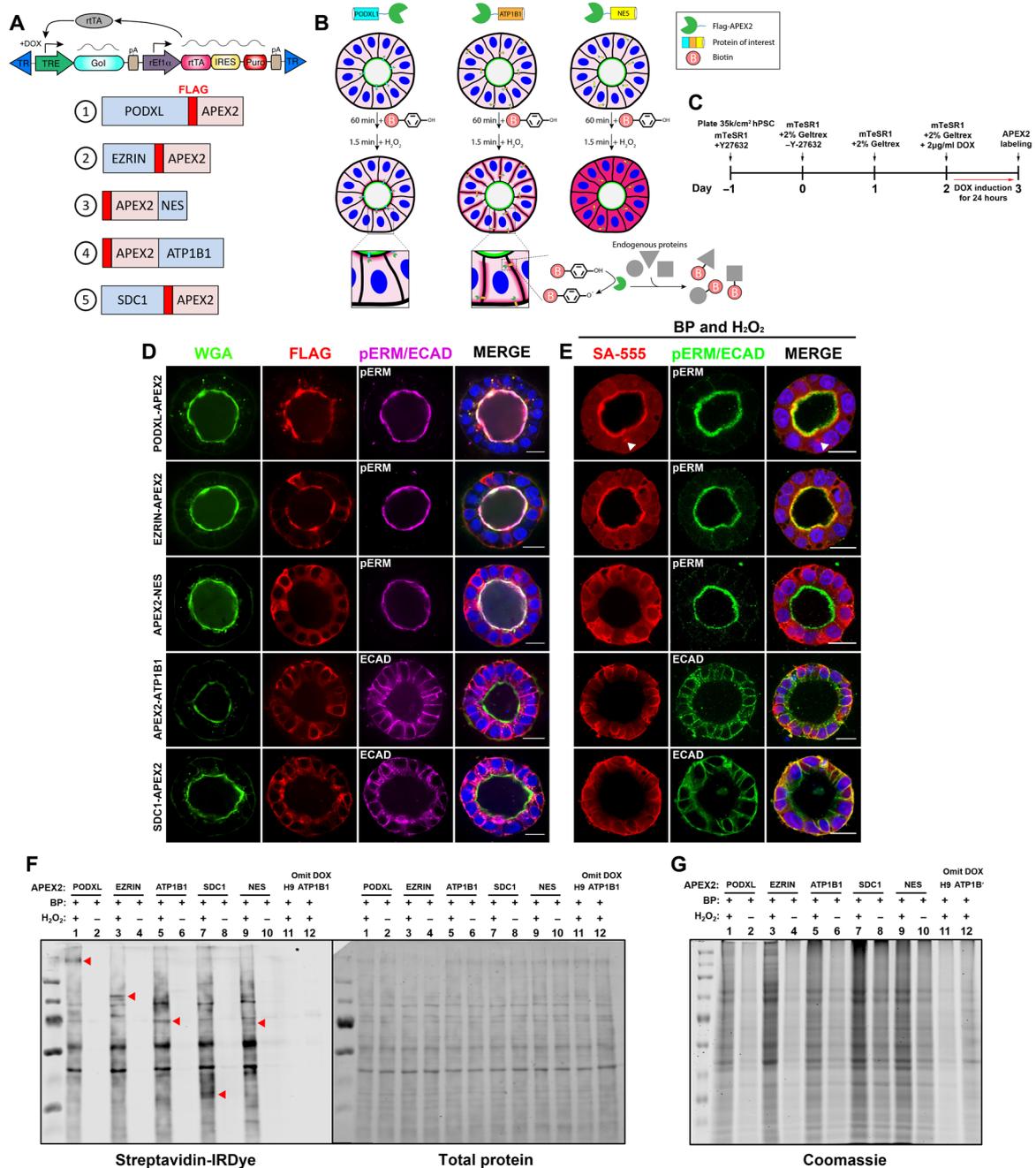


Fig. 1. APEX2-based spatial biotinylation of apical and basolateral membrane territories in an hPSC-cyst epiblast model. (A) Design of APEX2 fusion construct based on a DOX-inducible piggyBac transposon system. Gol, gene of interest; TR, terminal repeats; TRE, tetracycline-responsive element; pA, polyadenylation sequence; rE1 α , rat elongation factor 1 α promoter; rT α , reverse tetracycline transactivator; IRES, internal ribosome entry site; Puro, puromycin. (B) Flow diagram of APEX2-based proximity biotinylation of apical, basolateral, and cytoplasmic domains using APEX2 fusion proteins (e.g., PODXL, ATP1B1, or NES). hPSC-cysts expressing each construct were incubated with BP (biotin-phenol), followed by a 1.5-min H₂O₂ treatment to trigger biotinylation: APEX2 converts BP to biotin-phenoxy radicals, which form covalent bonds with proximal endogenous proteins. The red stripe indicates the labeling radius of APEX2 (~20 nm). (C) hPSC-cyst APEX2 labeling assay timeline. (D) Optical section of indicated d3 APEX2 hPSC-cysts stained with indicated markers. FLAG (red) indicates the APEX2 construct localization. (E) Streptavidin (SA-555) staining of d3 APEX2 hPSC-cysts after labeling, showing spatially distinct streptavidin signals. White arrowhead indicates biotinylated intracellular vesicles. (F) Western blot of whole-cell lysates from d3 hPSC-cysts using indicated conditions was probed for biotin, revealing distinct patterns of construct-specific biotinylation. Red arrowheads indicate unique bands. Right: Total protein staining. (G) Coomassie staining following enrichment of biotinylated proteins from indicated samples. Some background biotinylation signal is seen in unlabeled (2, 4, 6, 8, and 10) and negative control (11 and 12) samples. Scale bars, 20 μ m unless otherwise noted. Nuclei: Hoechst (blue).

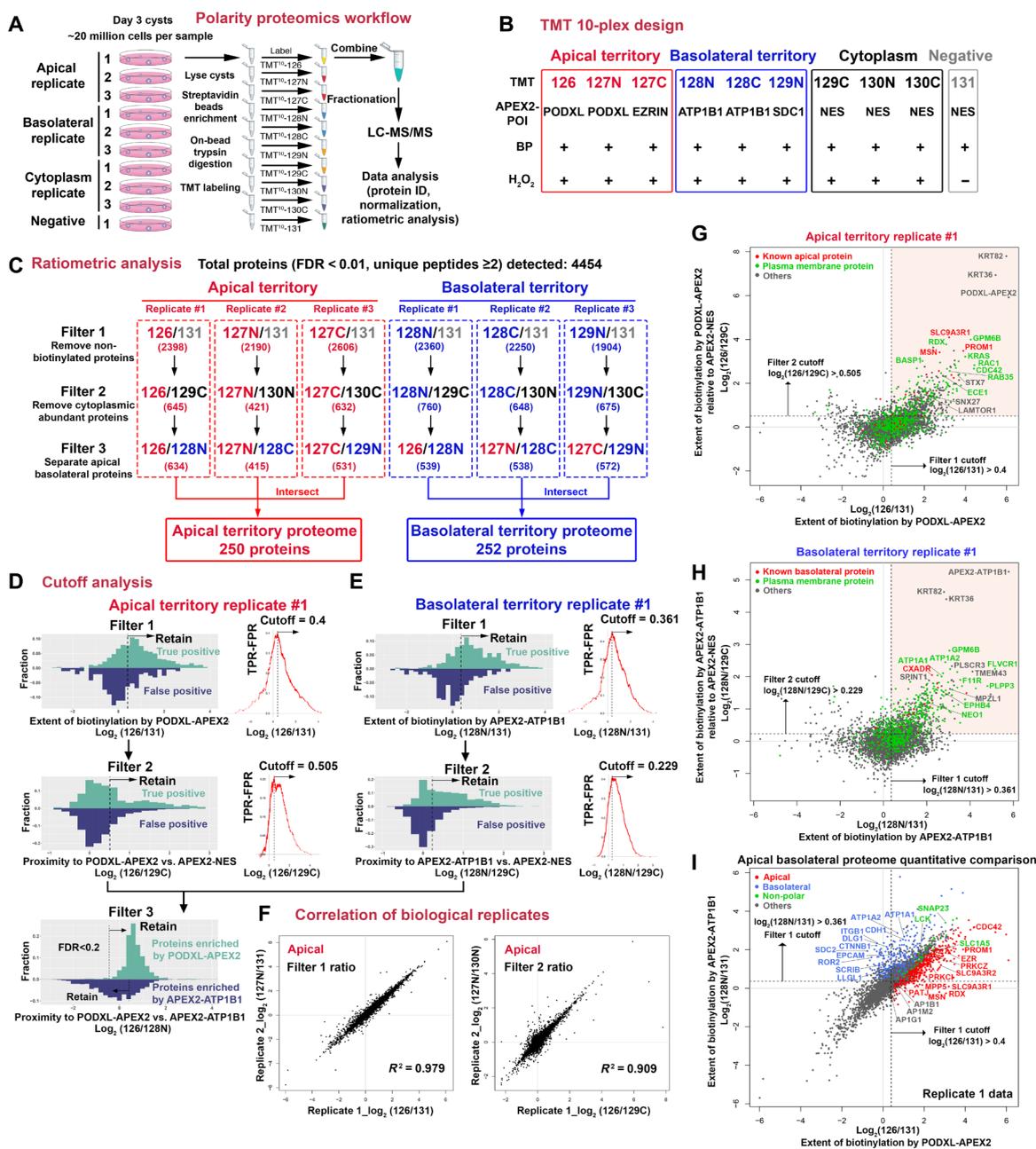


Fig. 2. Ratiometric profiling of cell polarity proteomes in hPSC-cysts. (A) Workflow of hPSC-cyst cell polarity proteome analysis, using three replicates of apical (2× PODXL and 1× EZR), basolateral (2× ATP1B1 and 1× SDC1), and 3× cytoplasmic (NES) samples and one negative control (NES without H₂O₂). (B) Experimental design of 10-plex TMT-based profiling. Row 1, specific TMT tags; row 2, APEX2 lines; rows 3 and 4, proximity labeling conditions. POI, protein of interest. (C) Flow diagram of ratiometric analyses to obtain proteins specific to apical and basolateral territories. Brackets: Number of proteins that passed each filter. (D and E) Representative histograms and receiver operating characteristic (ROC) curves of apical replicate 1 (D) and basolateral replicate 1 (E), illustrating how filters 1, 2, and 3 were applied. See fig. S2 (A to D) and Materials and Methods for additional cutoff analysis details. (F) Representative scatterplots (apical replicates 1 and 2) revealing strong correlation among biological replicates (additional plots in fig. S2E). (G and H) Representative ratiometric analysis scatter plots [(G), apical territory replicate #1 (PODXL-APEX2, TMT¹⁰-126); (H), basolateral replicate #1 (APEX2-ATP1B1, TMT¹⁰-128N)] showing the distribution of proteins based on TMT ratios [x axis, 126 or 128N/129C (cytoplasmic #1)]; y axis, 126 or 128N/129C (cytoplasmic #1)]. Orange shade (top right quadrant) indicates proteins that were analyzed for filter 3. Colored dots/fonts indicate proteins with existing apical (red, top)/basolateral (red, bottom), or plasma membrane and cell membrane (green) annotations; others in gray. See fig. S2H for additional plots. (I) Representative ratiometric scatter plots based on normalized TMT ratios of apical #1 (x axis, 126/131) and basolateral #1 (128N/131) of 4454 proteins revealing a clear enrichment of proteins from apical (red), basolateral (blue), and nonpolar (green) lists in distinct domains of the scatter plot. Proteins identified through filter 4 are indicated in gray fonts. See fig. S2I for additional plots.

analysis details including normalization, filtering, and derivation of cutoff values using receiver operating characteristic (ROC) analysis and FDR, based on UniProt annotations]. We first filtered out proteins that were nonspecifically bound to the streptavidin beads, by removing proteins that were enriched in the unlabeled samples compared to the apical or basolateral pools (filter 1; Fig. 2, C to E, and fig. S2, A to D). This removed, on average, $48.7\% \pm 5.3\%$ (SD) of the proteins from each membrane territory (see full lists in table S1, D to I). Filter 2 was used to remove membrane territory proteins that are abundant in the cytoplasm (displaying low apical/cytoplasmic or basolateral/cytoplasmic ratios), enriching for polarized membrane-specific proteins (Fig. 2, C to E, and fig. S2, A to D; full lists in table S1, J to O). Filter 3 was based on apical/basolateral ratio: We removed, from the apical territory proteome, a small number of proteins that were much more strongly biotinylated by basolateral APEX2 constructs than apical constructs, and vice versa, using FDR (cutoff set at $FDR < 0.2$). This enriched for proteins that are specifically localized to apical or basolateral membrane territories (see Fig. 2F and fig. S2E for correlation of TMT ratios among replicates for each filtering step). Intersection of the protein lists after filter 3 led to the lists of 250 apical and 252 basolateral membrane territory proteins (Fig. 2C and fig. S2F). Given its cytosolic vesicular localization (Fig. 1, D and E, top row), the PODXL-APEX2 construct biotinylated additional proteins that are not biotinylated by EZRIN-APEX2. Thus, 139 proteins that were shared among replicate #1 and replicate #2 (PODXL-APEX2 #1 and PODXL-APEX2 #2, respectively; see fig. S2F and table S2A), but were not included in the list of 250 apical proteins, were added to the apical list (389).

This analysis accurately identified many known interacting partners for each of the APEX2 labeled baits. Each interactome was highly enriched in known binding partners for each bait as identified in BioGrid (table S2B) (37). This indicates that the APEX2-based labeling and TMT-based approach are effective for identification of interactions in the hPSC-cyst system.

The final lists [apical, 389 (table S2C); basolateral, 252 (table S2D)] were then intersected (identifying proteins that are shared among apical/basolateral replicates; Fig. 2C and fig. S2G), to reveal proteins specific to apical (359), basolateral (222), or general membrane (30, nonpolar; table S2E) territories. Figure 2 (G and H) and fig. S2H show distributions of TMT ratios for each protein and provide graphical representations of how the filtering of 4454 proteins was performed, based on the polarized enrichment of each protein [low or high TMT ratio: apical or basolateral/negative (x axis) versus cytoplasmic (y axis)], to establish a list of apical or basolateral territory proteins for each replicate (top right quadrant). In Fig. 2I and fig. S2I, we examined the normalized TMT ratiometric distribution of the 4454 proteins in apical (x axis, e.g., PODXL-APEX2 #1) versus basolateral (y axis, e.g., APEX2-ATP1B1 #1) samples and highlighted proteins found in apical (359, red), basolateral (222, blue), and nonpolar (30, green) territories. Proteins from each territory are well segregated: Proteins in the apical proteome appear below the diagonally distributed plots (including many established apical proteins, e.g., PRKCZ, CDC42, SLC9A3R1/NHE-RF1, and RDX), while basolateral (e.g., CDH1, ITGB1, and CTNNB1) or nonpolar proteins are distributed above or in the middle of the diagonal, respectively (Fig. 2I and fig. S2I). Notably, our analysis failed to identify some proteins with known polar enrichment, such as PTEN and CRB3 (31, 38–40); this may be due to protein or

membrane interactions that shield these proteins from BP radical labeling in intact cells, a limitation of this approach.

Our analysis resulted in a membrane proteome list that was highly enriched in proteins previously annotated as having plasma membrane/cell membrane annotations in UniProt (31.6% of apical proteins and 50.4% of basolateral proteins) (Fig. 3A, top). In addition, 59 of 174 (33.9%) of the plasma membrane proteins identified in the SubCell BarCode database [a resource for spatial localization of subcellular molecules (11)] were found in these lists. Among the captured proteins with previous polarized annotations, most (88% for apical and 90% for basolateral) were found in the expected territory (Fig. 3A, bottom, see adjacent expanded bar graph), indicating the highly specific nature of these cell polarity proteome datasets. However, proteins with previous apical/basolateral membrane annotations make up only 13% (apical) and 36% (basolateral) of these proteome lists. The remaining proteins (labeled as “no apical/basolateral annotation” in Fig. 3A) are candidates for previously unrecognized apical or basolateral membrane territory proteins [apical/basolateral annotation based on UniProt and Gene Ontology (GO); see table S2F for the complete list] with potential functions in cell polarity.

As expected, each proteomic list contains numerous known components of polarized membrane trafficking, adhesion, and transporter machineries (Fig. 3B, full list in table S2G). The apical and basolateral proteomes identified here show distinct and non-overlapping signatures regarding their cellular localization and biological functions based on GO (Fig. 3, C to E, and fig. S3, A to F): The apical membrane proteome highlights the processes of vesicular trafficking, actin cytoskeletal organization, and guanosine triphosphatase activity, whereas the basolateral proteome features categories covering junction formation, morphogenesis, and activities related to transmembrane proteins. Furthermore, clustering of proteins based on their reported protein-protein interactions reveals a wide spectrum of functional relationships (Fig. 3, F and G). Given its known nuclear localization (23), 182 proteins (table S2H) found in the EZRIN-APEX2 list, but not found in the PODXL-APEX2 lists after filter 3, included a set of cell cycle proteins (e.g., ATM, CDC6, CDK2, CDK7, PCNA, and WEE1) and showed a marked enrichment of cell cycle and nucleotide binding GO terms (fig. S3, D to F).

We next compared these hPSC lists with two recently published MDCK monolayer proteomes of the apical-lateral border (17) and apical and basolateral surface (16). In the apical-lateral border study, the spatial proteomes surrounding Pals1 (which marks the apical vertebrate marginal zone, a restricted membrane compartment apical to tight junctions) and Par3 (a tight junction marker, located more basally than Pals1) territories were examined using an APEX2 strategy (17). While 37.3% (41 of 110) of the Pals1 territory proteins were shared with our hPSC-cyst apical proteome list, only 8.1% (7 of 86) of the Par3 territory proteins were shared (fig. S3G). Thus, the hPSC-cyst apical membrane territory proteome is highly similar to an apical membrane compartment previously defined in MDCK cells.

For the MDCK surface proteome dataset defined by surface biotinylation (16), 32.7% (52 of 159) of basolateral proteins were in common with the hPSC basolateral list. However, only 7.4% (9 proteins of 122) and 15.2% (5 of 33) of apical and nonpolar proteins, respectively, were shared with the respective hPSC lists. This disparity may be due to differences in cell type/state (pluripotent versus kidney; stem cells versus more differentiated cells), species (human versus dog), experimental setup (labeling of internal

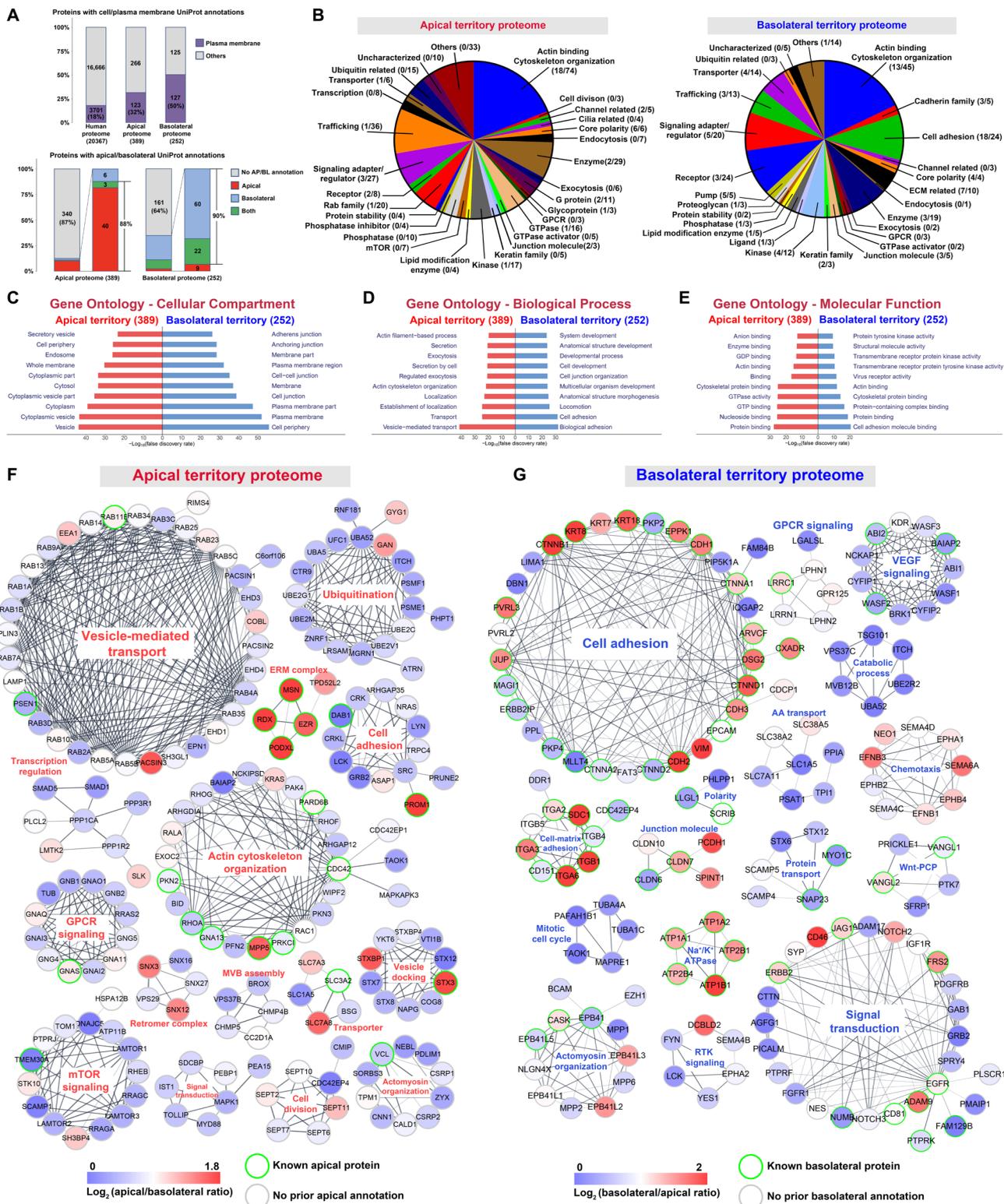


Fig. 3. Characterization of the apical and basolateral proteomes. (A) Bar plots showing the fraction of proteins with UniProt membrane annotations in the entire human proteome and apical and basolateral territories (top). Previously annotated apical or basolateral proteins are enriched in respective proteomic lists (bottom). (B) Functional classification of the proteins in apical and basolateral territory proteomes. Values in brackets indicate the number of proteins with existing membrane subdomain annotation (apical or basolateral). (C to E) Gene Ontology (GO) analysis (top 10 shown) of Cellular Compartment (C), Biological Process (D), and Molecular Function (E) terms in the apical and basolateral proteome lists. (F and G) STRING protein network analysis of apical (F) and basolateral (G) proteomes. Gray lines indicate established interactions. Clusters with more than four proteins are shown.

versus external membrane surfaces), or culture conditions (2D versus 3D).

To identify proteins that are enriched in both cytoplasm and apical or both cytoplasm and basolateral territories (e.g., trafficking molecules, enzymes, and scaffolds), proteins that passed filter 1 were directly filtered by filter 3 (cutoff, FDR < 0.2), skipping filter 2. These proteins (filter 4) are also candidates with potential functions in cell polarity (table S3, A to C).

Cell polarity proteomes reveal new apically and basolaterally polarized proteins

To validate the spatial assignments from our proteomics approach, we selected proteins from each membrane territory (apical, basolateral, and unpolarized) for further study. We focused on proteins important for vesicular trafficking as well as those with unexpected polarized localizations. Our selections were also guided by the availability of commercial antibodies and transgenes for recombinant expression. We examined seven apically enriched proteins (BASP1, ECE1, LAMTOR1, RAB35, SNX27, AP1G1, and STX7, all without UniProt apical annotation), six basolaterally enriched proteins (EPHB4, SPINT1, NEO1, MPZL1, BMPR1A, and ACVR1B, all without UniProt basolateral annotation), and three unpolarized membrane territory proteins (SNAP23, SLC1A5, and LCK) in d2 hPSC-cysts (enrichment and presence/absence of apical/basolateral UniProt annotation found in Fig. 2, G and H, fig. S2H, and table S2, F and G). Most of the selected proteins are involved in signaling pathways, cytoskeleton remodeling, or secretory trafficking systems (fig. S4A; see table S2G for functional annotation). Given the critical role of trafficking in cell polarity (41) and in pluripotency exit (42), the polarized membrane proteins involved in membrane traffic (e.g., SNX27 and AP1G1) in 3D hPSC systems are of particular interest for further investigation.

Of note, AP1B1 and AP1M2, the β 1 and μ 1B subunits, respectively, of the clathrin-associated adaptor protein complex 1 (AP-1), were found in the apical filter 4 list; in addition, AP1G1 was just below the filter 4 cutoff. AP-1 is an important complex that is known for its role in clathrin-dependent traffic at the trans-Golgi network and endosomes (43, 44). To better understand the role of the AP-1 complex in hPSC-cyst polarity, we focused our efforts on the AP1G1 subunit for several reasons. First, the β -subunit of the AP-1 complex is highly similar to the β -subunit of the endocytic AP-2 complex, and these two proteins may be able to functionally substitute for one another (45). Second, the μ 1B subunit is one of two alternate AP-1 μ subunits that perform distinct cargo selective functions (43, 44); thus, its functional analysis might not reveal the entire role of AP-1 in apical trafficking. In contrast, the γ subunit plays an integral role in maintaining AP-1 complex structure and it is therefore likely to substantially affect AP-1 complex function (46). Although there are two alternate γ -subunit proteins, γ 1 (encoded by *APIG1*) is detected at higher levels than γ 2 (encoded by *APIG2*) in hPSC and is more highly conserved evolutionarily (47). In addition, the *ApIgl*-KO (knockout) mouse displays preimplantation lethality (48), but the role of AP1G1 during peri-implantation has not been examined.

We found that LAMTOR1, SNX27, and STX7 are localized to vesicles proximal to the PODXL-marked apical surface [demarcated by wheat germ agglutinin (WGA) staining] in d2 hPSC-cysts using immunofluorescence (IF; fig. S4B): d2 cysts display luminal and polarized epithelial characteristics largely identical to d3. To examine the localization of the other proteins, hPSC lines stably

expressing tagged versions of the respective recombinant proteins were generated and analyzed in d2 cysts: BASP1 (BASP1-HA), ECE1 (ECE1-HA), AP1G1 (AP1G1-mCherry), and RAB35 (EGFP-RAB35). BASP1 and ECE1 show a clear colocalization with PODXL at the apical surface, while AP1G1-mCherry shows a subapical localization and diffuse cytosolic staining (fig. S4B), findings consistent with our quantitative ratiometric data. EGFP-RAB35 was localized throughout the plasma membrane, potentially due to overexpression (fig. S4B). Consistent with their basolateral enrichment, EPHB4, SPINT1, NEO1, and MPZL1 show a clear colocalization with E-CADHERIN by IF (fig. S4C). Furthermore, in accord with previous findings (49), BMPR1A-Myc or ACVR1B-Myc exhibits a predominantly basolateral localization in cysts expressing DOX-inducible constructs (fig. S4C). Three proteins detected in the nonpolar proteome (SNAP23, SLC1A5, and LCK) exhibited the expected pattern of uniform membrane staining (fig. S4D). These results demonstrate that the APEX2 proximity biotinylation approach used in this study captures the proteome of respective membrane territories with high accuracy.

Functional tests of selected proteins identified via proteome-driven analysis

We next examined the functional roles of novel apical territory proteins using loss- and gain-of function analyses. We focused on BASP1 (50–53), a novel apical territory protein, and several apical territory proteins implicated in endolysosomal membrane traffic (LAMTOR1, SNX27, and AP1G1) (43, 44, 54–56). Two known apical proteins (EZR and RAB35) were also studied. CRISPR-Cas9-based genome editing was used to generate hPSC lines carrying loss-of-function mutations in targeted genes. Although we confirmed the successful deletion of *LAMTOR1* using IF 3 days after introducing the CRISPR-Cas9 construct, these cells could not be expanded beyond 5 days, presumably due to the critical role of the mammalian target of rapamycin (mTOR) signaling pathway in hPSC (57). For the other five genes, loss-of-function lines were successfully generated and confirmed using Western blot and Sanger sequencing (fig. S5, A to E). All edited cell lines maintained pluripotency over at least 10 passages. hPSC-cysts were generated from each of the KO cell lines and were stained with apical markers [PODXL and pERM (phosphorylated EZRIN/RADIXIN/MOESIN)], E-CADHERIN (basolateral), WGA (cell membrane), and F-ACTIN (actin cytoskeleton) to examine overall luminal cyst morphogenesis.

A spectrum of defects in epithelial morphogenesis were seen in all five KOs. Loss of EZR, an actin scaffolding protein involved in apical morphogenesis (22), results in hPSC-cysts with a small and off-centered lumen with ectopic PODXL⁺ internal compartments in d2 cysts (Fig. 4, A and B). Overexpression of EZR in hPSC-cysts does not lead to obvious defects in most cysts but promotes the formation of a normal central lumen (Fig. 4, C and D). Most RAB35-KO hPSC-cysts display luminal cyst organization similar to controls (Fig. 4, A and B), although a minority of the cysts exhibit inverted polarity, with EZR localized to the cyst periphery (fig. S5F). This contrasts with the strongly penetrant inverted polarity phenotype seen in MDCK cysts with *Rab35* deficiencies (58, 59), perhaps indicating that other RAB proteins substitute for RAB35 in this system but not in MDCK cells. RAB35 overexpression leads to cysts with single or multiple small PODXL⁺ lumens (Fig. 4, C and D, and fig. S5G).

SNX27, BASP1, and AP1G1 have not been previously implicated in luminal morphogenesis. SNX27 has been shown to cooperate

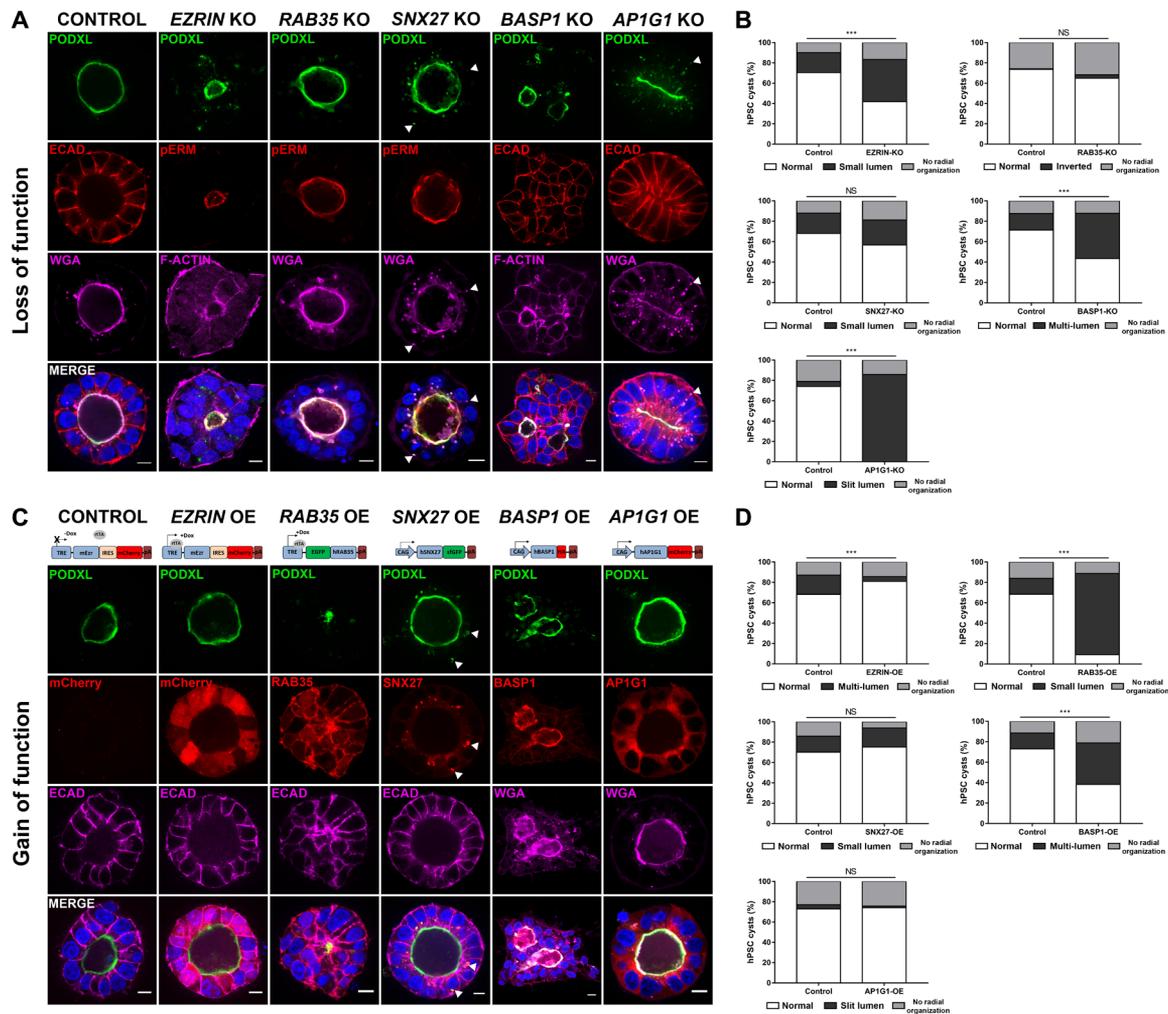


Fig. 4. Loss- and gain-of-function analysis of apical membrane territory proteins during hPSC-cyst formation. (A) Loss-of-function tests. d2 cysts generated from H9 hESC lines carrying null alleles of *EZR*, *RAB35*, *SNX27*, *BASP1*, and *AP1G1* were stained with indicated markers. See fig. S5 (F and H) for additional data and Materials and Methods for details. (B) Quantitation of luminal morphology from samples in (A) (EZRIN, 264 control and 241 KO cysts; RAB35, 159 control and 154 KO cysts; SNX27, 141 control and 245 KO cysts; BASP1, 192 control and 147 KO cysts; AP1G1, 158 control and 241 KO cysts). (C) Gain-of-function tests (OE, overexpression). Schematics of constructs used to generate transgenic H9 hESC lines are shown. Controls are from the DOX-inducible EZR line (EZRIN-IRES-mCherry) without DOX treatment. d2 cysts were stained as indicated (also see fig. S5G). (D) Quantitation of luminal morphology from samples in (C) (EZRIN, 202 control and 243 OE cysts; RAB35, 168 control and 152 OE cysts; SNX27, 178 control and 149 OE cysts; BASP1, 200 control and 193 OE cysts; AP1G1, 122 control and 185 OE cysts). Nuclei: Hoechst (blue) for all images. Scale bars, 10 μm (A) and (C). Small lumen: Luminal area smaller than 200 μm^2 . χ^2 test; *** $P < 0.0001$; NS, not significant.

with the retromer complex to promote recycling of transmembrane proteins from endosomes (55, 56). The majority of *SNX27*-KO cysts form luminal cysts comparable to controls, while the rest form cysts with a smaller or disorganized lumen (Fig. 4, A and B, and fig. S5H). In all *SNX27*-KO cysts (regardless of lumen size), PODXL shows ectopic vesicular accumulation (Fig. 4A, arrowheads; also see fig. S5H), consistent with a role for *SNX27* in apical protein recycling (60–62). Overexpression of *SNX27* does not appear to perturb cyst formation (Fig. 4, C and D). *BASP1* is a cytoskeleton-associated protein that promotes nerve sprouting (52, 53). Notably, the loss of *BASP1* results in disorganized cyst structures with a multi-lumen phenotype, in which two to three PODXL apical domains are seen per cyst; overexpression of *BASP1* causes a similar phenotype (Fig. 4, A to D).

Last, *AP1G1*-KO hPSC lines exhibit a novel phenotype: cysts contain a central disc-like lumen (appearing as a slit in a single

optical section), surrounded by tall, thin, radially organized cells (Fig. 4, A and B). While PODXL is enriched apically, demarcating the disc-shaped lumen, ectopic PODXL localization is also observed on intracellular vesicles, a portion of which localize proximal to the basolateral membrane (arrowhead in Fig. 4A), suggesting that AP-1 may play a role in apical sorting of PODXL vesicles. Notably, cells in a 2D *AP1G1*-KO monolayer display only a few PODXL vesicles (fig. S6A). These findings are highly intriguing given that, in MDCK monolayers, PODXL localization is also largely intact after AP-1 knockdown (16), and suggest that apical surface generation/maintenance can be affected by topology. E-CADHERIN, a known AP-1 cargo in MDCK cells (63), shows normal basolateral membrane localization in *AP1G1*-KO cysts (Fig. 4A). Overexpression of *AP1G1* results in cysts similar to controls (Fig. 4, C and D).

Together, these genetic loss- and gain-of-function studies suggest previously undescribed roles for RAB35, EZR, SNX27, BASP1, and AP1G1 in luminal morphogenesis in the setting of a 3D epiblast-like model. Moreover, the distinct differences in phenotypes for RAB35 and AP-1 in hPSC-cysts versus in MDCK cells emphasize that context-dependent features characterize the process of polarization in distinct cell types and that topology matters. Our list of polarized membrane proteins generated based on the hPSC-cyst cell polarity proteomics is a rich source that can be further mined to identify additional protein candidates that control epithelial morphogenesis.

Detailed analysis of AP1G1 in epithelial morphogenesis

AP1G1-KO hPSC-cysts display an unusual phenotype. The AP1G1-KO cells that form the cyst are consistently taller and thinner than control cells, which are uniformly cuboidal/columnar. The lumen in these KO cysts is flat and disc shaped, in contrast to the balloon-like large lumen of controls, suggesting that while apical polarization initiates in AP1G1-KO hPSC-cysts, expansion of the apical surface/lumen fails. Transmission electron microscopy (TEM) confirms the presence of a thin open lumen in KO cysts, with an apical surface containing microvilli and tight junctions (Fig. 5, A and B). While a growing body of evidence implicates the role of the AP-1 complex in apical traffic in diverse metazoans (64–69), this type of disc-like luminal phenotype has not previously been reported, to our knowledge.

Given this unique phenotype, we examined these cysts using additional apical markers, including pERM (actin scaffolding proteins), phosphorylated myosin light chain (pMLC; a marker of intracellular actin cytoskeletal contractility), phalloidin (marking F-ACTIN), and CDC42. In contrast to PODXL, which is partially mislocalized in AP1G1-KO cells, pERM, pMLC, F-ACTIN, and CDC42 localize normally in KO cysts (Fig. 5, C and D, and fig. S6B). In addition, ZO1 (a component of the tight junction) and RAB11 (a recycling endosome marker) show that ZO1⁺ or RAB11⁺ puncta are enriched apically as expected (Fig. 5, C and D). However, the distance between adjacent ZO1 puncta is greatly diminished in KO cysts, consistent with a decrease in apical surface diameter/cell (Fig. 5E). Notably, the basolateral domain appears largely intact in the absence of AP1G1, as the localization of basolateral cell adhesion molecules (CDH1 and CTNBN1) and ion transporters (ATP1A1 and ATP1B1) is comparable to controls (Fig. 5, C and D, and fig. S6C). Thus, in hPSC cysts, the AP-1 complex controls trafficking of PODXL vesicles and apical/luminal surface expansion.

To verify these findings, we performed a rescue experiment by introducing a transgenic construct expressing mCherry fused to a CRISPR-resistant form (construct details in Materials and Methods) of human AP1G1 into the AP1G1-KO background. For all three clonal lines of AP1G1-KO cells expressing the AP1G1-mCherry rescue construct, pluripotency (OCT4⁺/NANOG⁺) was maintained and hPSC-cysts were comparable to controls, with a round central lumen and cuboidally shaped cells (Fig. 5, E to G). Furthermore, the AP1G1-mCherry fusion protein exhibited an apically oriented localization in both monolayers and cysts (Fig. 5, F and G, and fig. S6A), overlapping extensively with RAB11 (Fig. 5H), a marker for recycling endosomes. This localization is consistent with a role in transmembrane protein recycling suggested by the ectopic PODXL foci found in the cytoplasm of AP1G1-KO cysts (Fig. 4A).

To monitor the kinetics of lumen formation, we used live-cell imaging of control or AP1G1-KO hPSC cysts stably expressing mTnG [membrane tdTomato (mT) and H2B (nuclear)-EGFP (nG)]. The balloon-like lumens of control cysts formed within 2 days and continued to expand between 1 and 4 days; cyst cells remained cuboidal to columnar in shape throughout this process (Fig. 5I). While some mT⁺ membrane vesicles remained in the sub-apical domain of control d1 to d2 cysts, few such vesicles could be seen in d3 to d4 cysts. In contrast, KO cysts initiated radial organization and formed a disc-like lumen but never displayed hollow central lumens (Fig. 5J; see 3D reconstructions in movies S1 and S2). Cells remained tall and thin, with reduced apical surface area throughout the process of lumen maturation (also see d5 cysts in fig. S6D). Consistent with the finding that apically charged membrane vesicles (PODXL⁺/WGA⁺) displayed pronounced cytoplasmic localization (a phenomenon rarely seen in control cysts; Fig. 4A), numerous subapical mT⁺ vesicles were obvious at all stages (Fig. 5J). Together, observation of these subapical PODXL vesicles, in combination with the restricted apical surfaces of AP1G1-KO cells, as documented by ZO1 staining, suggests a role for AP-1-dependent trafficking of apically charged vesicles during expansion of apical cell membranes during 3D cyst formation. We propose that, in the absence of sufficient apical cell membrane, each cell's apical surface area is reduced and the luminal cyst surface itself remains small and disc like.

Impaired apicosome formation and trafficking in cells lacking AP1G1

In hPSC-cysts, a central lumen forms after the coalescence of apicosomes, specialized intracellular structures enriched with apical membrane proteins (20). Apicosome-like structures have been observed in several lumen-forming tissues (20, 70–72), and a recent study showed that apicosomes are critical for mouse blastocoel cavity formation (73). In isolated hPSC, PODXL vesicles fuse in the peri-nuclear region of each cell and recruit F-ACTIN and actin scaffolding proteins (e.g., EZRIN) to give rise to a small intracellular PODXL/EZRIN vesicle that grows in size to form a highly organized apically charged membrane structure studded with microvilli and a primary cilium (the apicosome). After mitosis, apicosomes fuse with the newly established cytokinetic plane to establish the nascent apical (luminal) surface of a two-cell cyst; with further cell division, such cysts continue to expand (20). Because of its important role in lumen initiation, we compared apicosome behavior in control, AP1G1-KO, and rescue cells.

In individually plated control cells, this previously described apicosome formation process was recapitulated (Fig. 6A). However, this process was impaired in KO cells, as PODXL vesicles failed to merge into mature apicosomes (Fig. 6, A and B). In AP1G1-KO rescue cells, apicosome formation was similar to controls and AP1G1-mCherry vesicles appeared directly adjacent to apicosomes, but not in the apicosomal membrane proper (Fig. 6, A and B).

When control hPSC are plated as small cell clusters, apicosomes form in a peri-nuclear location in most cells of the cluster. As cells begin to radially organize, apicosomes are transported to the center of the radial structure, where they fuse, creating the apical surface of each cell and forming a central lumen for the cyst (20). To examine this process in AP1G1-KO cysts, we used 3D high-resolution time-lapse (20-min interval) confocal microscopy to follow control and KO cells from freshly plated aggregates (before lumen formation:

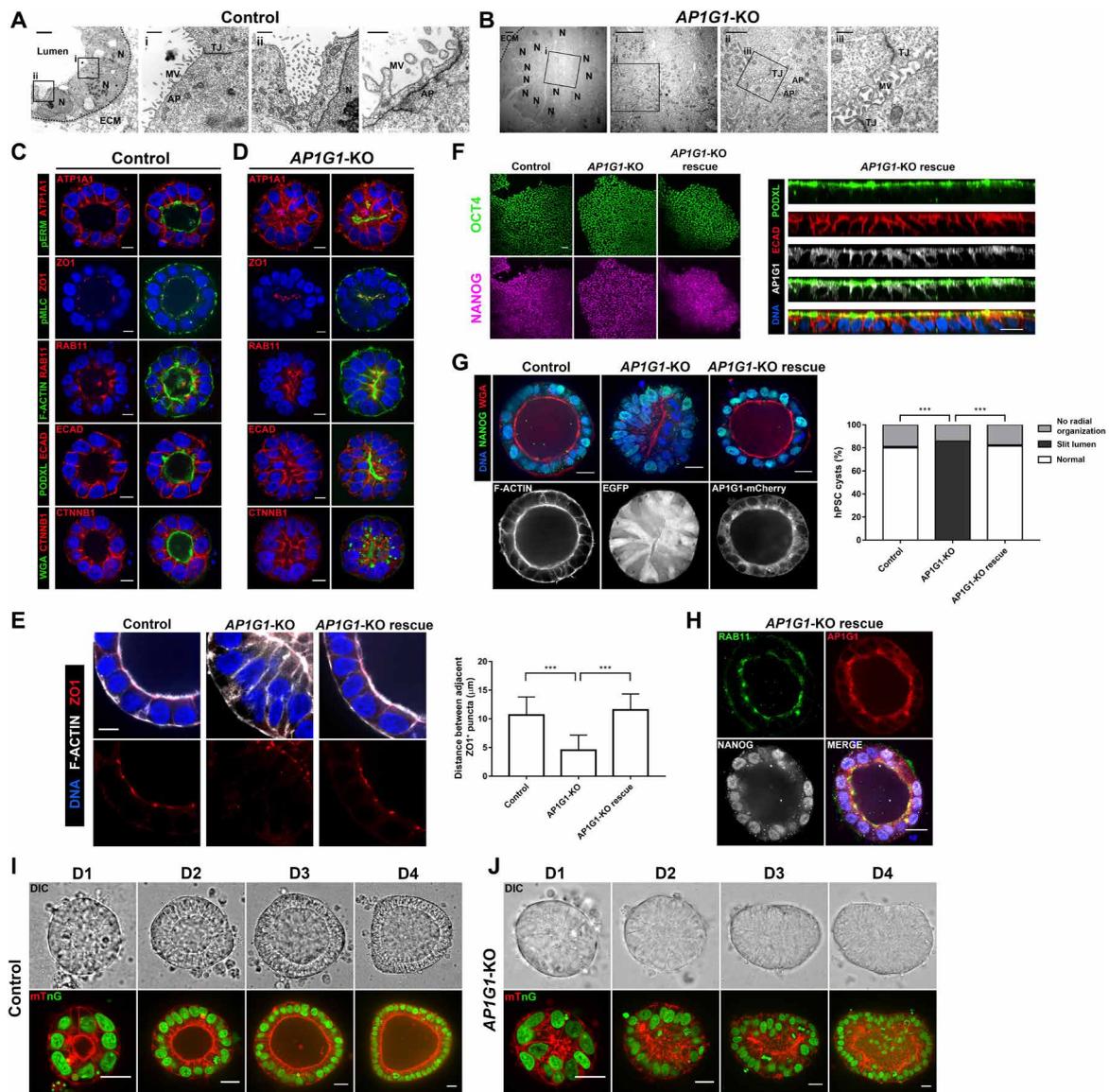


Fig. 5. Detailed analysis of lumen expansion defects in *AP1G1*-KO hPSC-cysts. (A and B) TEM analysis of d3 control (A) and *AP1G1*-KO (B) cysts. The disc-shaped lumen in *AP1G1*-KO displays a slit-like shape in the TEM micrograph. N, nucleus; ECM, extracellular matrix; TJ, tight junction; MV, microvillus; AP, apical surface. Scale bars, 5 μ m, 1 μ m, 1 μ m, and 500 nm (A, left to right) and 5 μ m, 5 μ m, 2 μ m, and 500 nm (B, left to right). (C and D) Confocal images of d2 control (C) and *AP1G1*-KO (D) cysts stained with indicated markers [apical: pERM, F-ACTIN, and PODXL; basolateral: ATP1A1, E-CADHERIN, and CTNNB1; recycling endosome: RAB11; tight junction: ZO1; phosphorylated myosin light chain (pMLC); membrane marker (WGA, shows apical enrichment)]. Scale bars, 10 μ m (C and D). (E) Magnified images of the apical surface of d3 control, *AP1G1*-KO, and *AP1G1*-KO rescue (KO + CAG-driven expression of *AP1G1-mCherry*) hPSC-cysts stained for ZO1 and F-ACTIN. Right: Quantitation of the distance between adjacent ZO1⁺ foci (apical surface width, mean \pm SD, one-way ANOVA with Tukey's multiple comparison test; ****P* < 0.0001). Scale bar, 10 μ m. (F) Left: Confocal images of control, *AP1G1*-KO, and *AP1G1*-KO rescue hPSC monolayers stained for pluripotency markers, OCT4 and NANOG. Right: *AP1G1*-mCherry is subapical in the *AP1G1*-KO rescue cells. Scale bars, 50 μ m and 20 μ m (for monolayer images). (G) d3 control, *AP1G1*-KO, and *AP1G1*-KO rescue hPSC-cysts stained with indicated markers. *AP1G1*-mCherry is localized proximal to the apical surface. Quantitation of luminal morphology is shown (right). χ^2 test; ****P* < 0.0001. Scale bars, 20 μ m. (H) d3 *AP1G1*-KO rescue cyst, stained for RAB11 and NANOG. The *AP1G1*-mCherry signal largely colocalizes with RAB11⁺ recycling endosomes. Scale bar, 20 μ m. (I and J) Time-lapse imaging (24-hour intervals) of cyst formation in control (I) and *AP1G1*-KO (J) cysts stably expressing H2B-EGFP and membrane-tdTomato (mTnG). Top, DIC (differential interference contrast); bottom, confocal. Scale bars, 20 μ m.

imaging started 5 hours after lumenogenesis is triggered, 00:00) to 24 hours (19:00), when expanded lumens are established in controls (dotted circles in Fig. 6C, top; fig. S7, A and B, and movie S3) but not in KO cysts (Fig. 6C, bottom; fig. S7, A and B, and movie S4). Six optical sections were captured in the *z* axis with 3- μ m steps,

starting in close proximity to the coverslips (0 to 15 μ m; to accommodate the increase in cyst height over time, only 3-, 6-, and 9- μ m steps are shown in Fig. 6D). In controls, apicosomes were readily observed in newly plated individual cells, and over time, apicosomes from several cells fused to form a central lumen (movie S3).

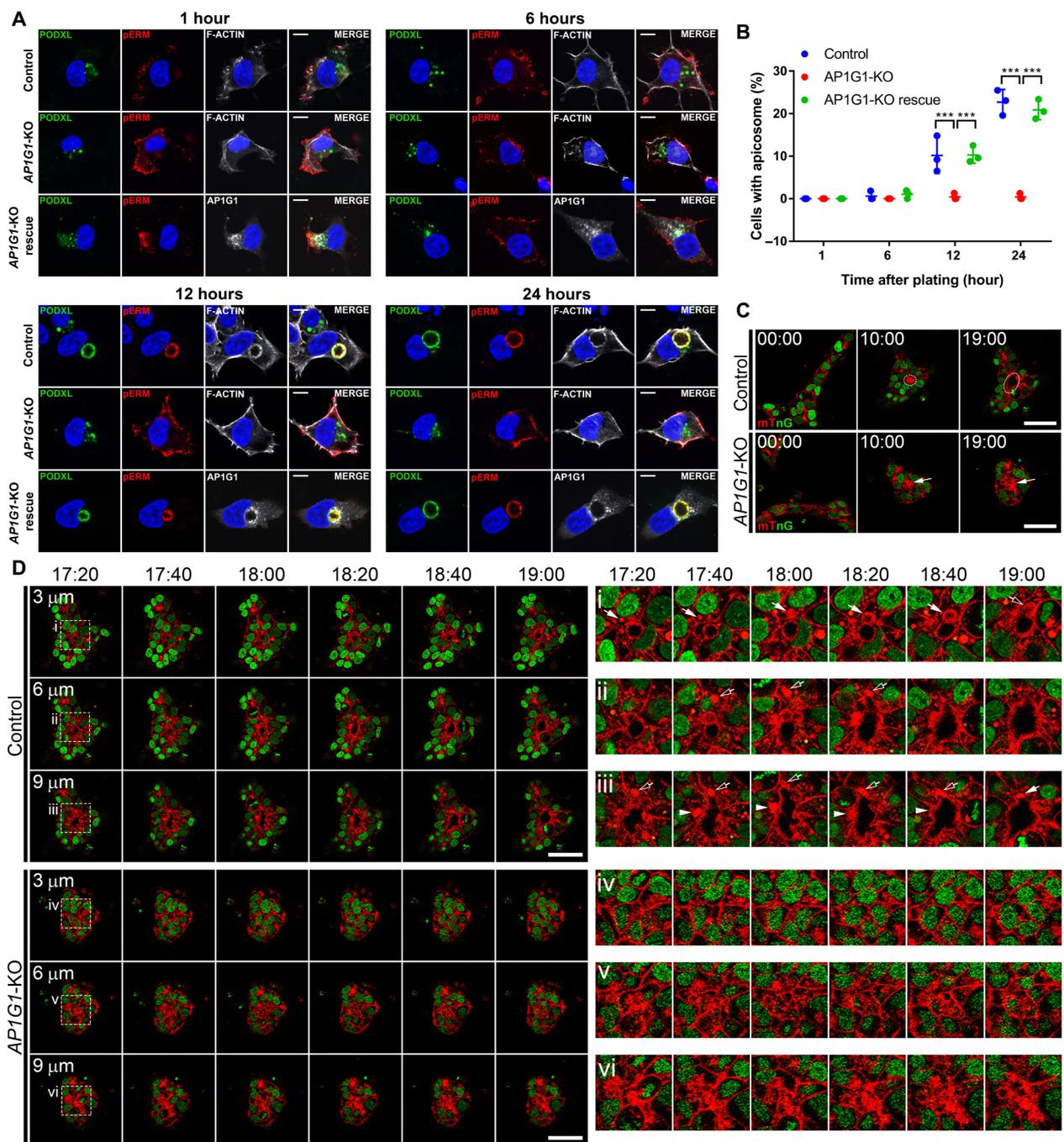


Fig. 6. Apicosome formation is impaired in cells lacking AP1G1. (A and B) Apicosome formation time-course assay (1, 6, 12, and 24 hours after plating). Singly dissociated control and *AP1G1*-KO and *AP1G1*-KO rescue hESC were stained with indicated apicosome markers. Apicosome formation is defective in *AP1G1*-KO cells and rescued by expression of AP1G1-mCherry [quantitated in (B), mean \pm SD shown, two-way ANOVA with Tukey's multiple comparison tests; $***P < 0.0001$]. Scale bars, 10 μ m. (C and D) Snapshots of magnified images from high-resolution confocal 3D time-lapse analysis (movies S3 and S4, 20-min intervals, six of 3- μ m Z-steps) of mTnG [membrane-Tomato/nuclear (H2B)-GFP] control and *AP1G1*-KO hPSC-cyst formation from aggregates: Imaging started 5 hours after inducing cyst morphogenesis (00:00) and ended at 24 hours (19:00, 58 frames). (C) Dotted circles indicate an expanding central lumen. Arrows: Central accumulation of membrane materials. Original images shown in fig. S7. (D) Magnified images from 3, 6, and 9 μ m at indicated time points reveal that apicosome fusion events result in central lumen expansion in controls. In KO, while a central accumulation of mT⁺ membrane materials is seen (but not organized apicosome structures), a hollow lumen structure does not form. Insets (i to vi) indicate magnified regions (forming lumens) shown in adjacent panels. Arrow, open arrow, and arrowhead in (i), (ii), and (iii) indicate three separate apicosome formation and fusion events. Scale bars, 50 μ m.

Lumen size was increased and apical surface was expanded as more apicosomes were integrated into the cortical membrane (Fig. 6D, i to iii). In the *AP1G1*-null background, however, organized apicosome structures were rarely seen. Although by 14 hours a limited mT^+ strip of central apical membrane could be discerned, surrounded by radially organized cells, this apical surface remained disc like (Fig. 6D, iv to vi, and movie S4) and cells were uniformly thin and tall.

Given the mislocalization of PODXL vesicles in *AP1G1*-KO cysts and the documented defective apicosome formation (which is also marked by apparent defects in trafficking of PODXL vesicles), we speculated that the odd disc-shaped lumen seen in these cysts might be due, at least in part, to mistrafficking of apically charged PODXL vesicles, leading to reduced apical membrane at the apicosome and at the luminal surface. To test this, we generated hPSC-cysts overexpressing a PODXL-mCherry construct and found a modest increase in apicosome formation (fig. S8A); furthermore, lumens of d2 *AP1G1*-KO-PODXL-mCherry cysts were somewhat expanded (fig. S8B). In addition to mistrafficking of PODXL-charged membrane, impaired osmotic swelling caused by mislocalized ion transporters might contribute to the disc-like lumens seen in *AP1G1*-KO cysts. Indeed, NHERF1/SLC9A3R1, a modulator of the NHE3/SLC9A3 sodium-hydrogen transporter, a major regulator of luminal electrolyte balance, activity, and localization (74, 75), was identified in the apical membrane territory proteome list (Fig. 2I and table S2C). Using IF, we found that NHERF1 was mislocalized in KO cysts (found in ectopic PODXL vesicles; fig. S8, C and D) while maintaining NHERF1 levels comparable to controls (fig. S8E). Before cyst formation, NHERF1 was found in apicosomes and surrounding PODXL vesicles (fig. S8F). Together, these results are consistent with a novel role for the AP-1 complex in proper trafficking of essential apicosomal and vesicular membrane components (e.g., PODXL and ion transporters) that contribute apical membrane and function to expand the apical lumen during hPSC-cyst formation.

DISCUSSION

In this study, we used APEX2-based spatial proteomics to examine the cell polarity proteome of a 3D hPSC-derived system that models the early human epiblast cavity. This unbiased, systematic characterization of apical and basolateral membrane proteomes has allowed us to identify important players in the molecular machinery that is responsible for the polarized organization of the early human epiblast cyst. We demonstrate that *SXN27*, *BASP1*, and *AP1G1*, proteins not previously implicated in luminal morphogenesis, as well as *EZR* and *RAB35*, play critical roles in lumenogenesis and luminal cyst organization in hPSC. Furthermore, we establish an essential role for the AP-1 clathrin adapter complex in apical membrane trafficking in the context of hPSC-cysts. These findings and the spatially specific proteome lists presented here provide an important resource for further exploration of peri-implantation human development.

Our data highlight clear functional differences in the regulation of cell polarization and lumenogenesis between the hPSC-cyst model compared to other established 3D models of epithelial cyst morphogenesis. For example, in MDCK cysts, *Rab35* deficiency results in a highly penetrant inverted polarity phenotype (58, 59). In contrast, most hPSC-cysts lacking *RAB35* form normally, similar to

controls (Fig. 4A), suggesting that other RAB proteins may substitute for *RAB35* in hPSC-cysts but not in MDCK models. Furthermore, *AP1G1*-KO hPSC-cysts show no apparent changes in basolateral protein targeting, based on markers used in this study (*ATP1A1*, *ATP1B1*, *ECAD*, and *CTNNB1*), although more work is needed to definitively rule out such changes. These results underline emerging evidence that molecular mechanisms governing trafficking and polarization vary among epithelial cell types (43, 63, 65, 76, 77). Application of the APEX2 proteomics approach described here to other cell types or tissues/organoids will help decipher which aspects of polarity and trafficking are generalizable and which are tissue specific (e.g., differences in molecular requirements between species, tissue types, topology, and developmental timing).

The phenotype of *AP1G1*-KO cysts is, to our knowledge, unique, in that it affects not only apical membrane delivery but also cell shape and lumen shape (Figs. 4A and 5). These KO cysts exhibit a flat, disc-shaped lumen surrounded by tall, thin cells with a small apical surface area, in stark contrast to control cysts with their round lumens and cuboidal/columnar cells with large apical domains. Several molecular processes could contribute to this *AP1G1* phenotype (e.g., changes in osmotic swelling, tension, and apical constriction). Together, our data suggest that at least one important component is *AP1G1*-dependent apical membrane trafficking (Fig. 7). *AP1G1* is localized apically in control hPSC-cysts, where it colocalizes with *RAB11* vesicles (Fig. 5H). However, the subluminal

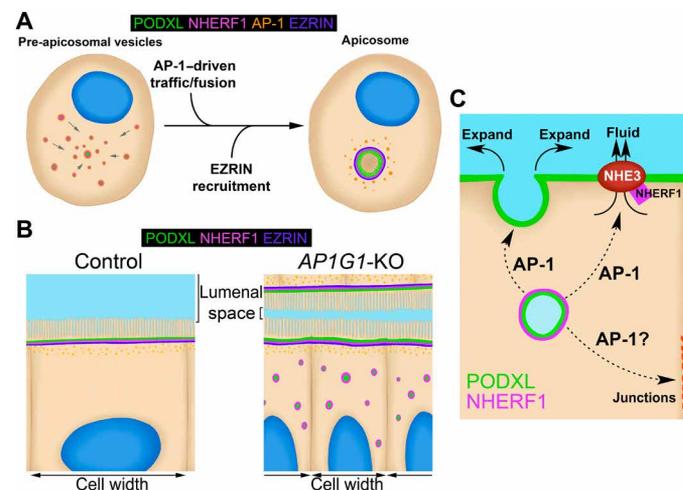


Fig. 7. A proposed model of AP-1 function during hPSC-cyst morphogenesis.

(A) During apicosome formation, the AP-1 complex aids in traffic and fusion of apicosome precursor vesicles (left, PODXL/NHERF1/AP-1) to form an apicosome (PODXL/NHERF1/EZRIN) surrounded by AP-1 vesicles. (B) Once hPSC-cysts have formed, AP-1-dependent trafficking of PODXL/NHERF1 vesicles adds additional membrane and potentially fluid to drive luminal expansion. While the apical lumen is demarcated by PODXL, NHERF1, and EZRIN in both control and *AP1G1*-KO cysts, ectopic PODXL/NHERF1 vesicles (demarcated by green and magenta) and significant reductions in luminal space and cell width are observed in KO cysts. *RAB11*⁺ recycling endosomal vesicles (orange foci) are observed directly adjacent to the apical lumen in control and KO cysts (these vesicles are AP-1⁺ in controls). (C) Three proposed AP-1-dependent mechanisms regulating hPSC-cyst morphogenesis. AP-1-driven trafficking of PODXL/NHERF1 vesicles may (i) physically expand the apical (luminal) membrane (and potentially add fluid), (ii) act to apically localize NHE3 (and potentially other pumps) involved in luminal expansion, and (iii) permit proper localization of junctional proteins.

domains of mutant cysts contain numerous mislocalized PODXL⁺ large vesicular puncta, and live imaging reveals a marked failure of apicosome formation in KO cells (Figs. 5 and 6). It is interesting that the ectopic PODXL⁺ vesicular puncta seen in KO cysts are much less prominent in the context of 2D monolayers of KO cells (fig. S6A). Similarly, it has been previously demonstrated that impaired AP-1 activity has little or no effect on PODXL localization in MDCK monolayers (16). These findings suggest that different topologies (3D versus 2D) may place different demands on the polarity trafficking machinery. Notably, the ectopic expression of PODXL-mCherry in KO cysts increases apicosome formation (though apicosomes are generally smaller than in controls; fig. S8A) and the resulting cysts exhibit slightly expanded lumens (fig. S8B). Together, these data suggest that apicosome-associated membrane trafficking, as well as trafficking of apically loaded membrane vesicles, is perturbed in AP1G1 mutant cells. We speculate that, with reduced apical surfaces, cells must be taller and thinner to maintain volume, while the lumen remains disc like, with a morphology that is highly reminiscent of other epithelial structures composed of thin pseudostratified cells, such as the embryonic small intestine (78, 79). While an AP1G1-null mouse line exists, its only known phenotype is preimplantation lethality (48). Recent evidence suggests that apicosomes participate in mouse blastocoel cavity formation (73), although the possibility that lethality of the mouse AP1G1 mutant is due to failure of apicosome trafficking during blastocyst cavity formation has not been specifically investigated.

The APEX2-based spatial proteomics approach developed in this study will be a powerful tool to examine the cell polarity proteomes of other 3D systems that remain unexplored because of limitations of previous methods. Because APEX2 can be specifically delivered to desired subcellular compartments by fusion to known localized proteins, and because chemicals used for APEX2-based biotinylation (biotin-phenol and H₂O₂) are cell permeant, APEX2-based cell polarity proteomics can be extended to a wide range of 3D systems (e.g., embryoid, organoid, and tissue). Moreover, tissue-specific proximity biotinylation can be used in vivo, using tissue-specific promoters, a technique recently used in proteomic mapping of fly olfactory projection neurons (35). Spatial quantitative proteomics profiling could also be used to decipher proteomic changes in mutants under pathological conditions. Because cell surface proteins are major targets for drug development (80), probing cell surface proteome changes under pathological conditions or in response to drug application could help identify therapeutic targets or monitor treatment efficacy.

Overall, this study opens the door for spatial proteomics to be used in a greater diversity of stem cell-derived embryoid/organoid models to elucidate mechanisms associated with development and disease, generating rich resources for cell, stem cell, and developmental biologists. In addition to in vivo validation of findings in stem cell models, future exploration of the new cadre of polarized proteins identified in this study will expand our understanding of the machinery underlying epithelial cell polarization and lumen expansion during critical morphogenetic events of peri-implantation human embryogenesis.

MATERIALS AND METHODS

hESC lines

Human embryonic stem cell (hESC) line H9 was used in this study [WA09, P48, WiCell; National Institutes of Health (NIH) registration

number: 0062]. All protocols for the use of hPSC lines were approved by the Human Pluripotent Stem Cell Research Oversight Committee at the University of Michigan and Human Stem Cell Research Oversight Committee at the Medical College of Wisconsin. H9 hESC were maintained in a feeder-free system for at least 20 passages and authenticated as karyotypically normal at the indicated passage number. Karyotype analysis was performed at Cell Line Genetics. All hPSC lines tested negative for mycoplasma contamination (LookOut Mycoplasma PCR Detection Kit, Sigma-Aldrich). All transgenic and KO hPSC lines in this study used H9 as the parental line.

hESC were maintained in a feeder-free culture system with mTeSR1 medium (STEMCELL Technologies). hPSC were cultured on 1% (v/v) Geltrex (Thermo Fisher Scientific)-coated six-well plates (Nunc). Cells were passaged as small clumps every 4 to 5 days with dispase (Gibco). All cells were cultured at 37°C with 5% CO₂. Medium was changed every day. hESC were visually checked every day to ensure the absence of spontaneously differentiated, mesenchymal-like cells in the culture. Minor differentiated cells were scratched off the plate under a dissecting scope once identified. The quality of all hESC lines was periodically examined by immunostaining for pluripotency markers and successful differentiation to three germ layer cells. All hESC were used before reaching the 70th passage.

hPSC-cyst and apicosome formation assays

Methods for these assays are previously as described (1, 8, 31). In short, singly dissociated cells were prepared using Accutase (Sigma-Aldrich) and were plated on coverslips coated with 1% Geltrex at 10,000 cells/cm² (apicosome) or 35,000 cells/cm² (hPSC-cyst, densely plated to form aggregates). For apicosome formation assays, cells were plated in mTeSR1 containing Y-27632 (STEMCELL Technologies) and 2% Geltrex. Apicosome formation is initiated soon after plating. For hPSC-cyst assays, cells were plated in mTeSR1 + Y27632 without Geltrex. After 24 hours, cells were then incubated in mTeSR containing 2% Geltrex overlay without Y-27632 with daily medium changes: The removal of Y-27632 triggers the apicosome-dependent hPSC-cyst morphogenesis.

Confocal microscopy of fixed samples

Confocal images of fixed samples were acquired using a Nikon-TiE fluorescence microscope equipped with a CSU-X1 spinning-disc unit (Yokogawa), or a Nikon-A1 and a Leica SP8 laser scanning confocal microscope. NIS Elements (Nikon) was used to generate 3D reconstruction images. Non-3D images were generated using FIJI (NIH) and Photoshop (Adobe).

Live-cell imaging

Zeiss LSM980 and Nikon A1R confocal microscopes configured with an environmental chamber (37°C and 5% CO₂) were used for high-resolution 3D hPSC-cyst formation time-lapse imaging using a 35-mm glass bottom culture dish (MatTek). Movies were generated using FIJI and Photoshop.

Transmission electron microscopy

d3 hPSC-cysts plated on a tissue culture plate were gently scraped and pelleted before processing. These samples were fixed with 2.5% glutaraldehyde in Sorenson's phosphate buffer for 1 hour, postfixed in 1% osmium tetroxide solution for 1 hour, and embedded in HistoGel (Thermo Fisher Scientific). Samples were dehydrated in a series of EtOH solutions (30, 50, 90, 95, and 100%) for 5 min each and

infiltrated with epoxy resin. These samples were then sectioned at 70 nm (Ultracut E; Reichert-Jung), placed on carbon slotted grids, stained using uranyl acetate, and imaged using Hitachi H600 equipped with a Hamamatsu digital camera and AMT image processing software.

Immunostaining

hPSC monolayers, hPSC aggregates, or hPSC-cysts on the coverslip were rinsed with phosphate-buffered saline (PBS; Gibco) twice, fixed with 4% paraformaldehyde (Sigma-Aldrich) for 40 to 60 min, then rinsed with PBS three times, and permeabilized with 0.1% SDS (Sigma-Aldrich) solution for 40 min. The samples were blocked in 4% heat-inactivated goat serum (Gibco) or 4% normal donkey serum (Gibco) in PBS 1 hour to overnight at 4°C. The samples were incubated with primary antibody solution prepared in blocking solution at 4°C overnight, washed three times with PBS (10 min each), and incubated in blocking solution with goat or donkey raised Alexa Fluor-conjugated secondary antibodies (Thermo Fisher Scientific) at room temperature for 2 hours. Counterstaining was performed using Hoechst 33342 (nucleus, Thermo Fisher Scientific), Alexa Fluor-conjugated WGA (membrane, Thermo Fisher Scientific), and phalloidin (F-ACTIN, Thermo Fisher Scientific). All samples were mounted on slides using 90% glycerol (in 1× PBS). When mounting hPSC-cyst samples, modeling clays were used as spacers between coverslips and slides to preserve luminal cyst morphology. Antibodies for IF staining are found in table S4A.

DNA constructs

APEX2 constructs

Detailed information regarding primers and DNA constructs can be found in table S4 (B to D). A DOX-inducible piggyBac transposon system (PB-TA-ERP2; gift of K. Woltjen; Addgene, #80477) (81) was used to generate five stable transgenic H9 hESC lines expressing APEX2 fused to human PODXL (OriGene; NM_00108111), EZR (gift of D. Louvard, Curie Institute, Paris), ATP1B1 (GenScript; NM_001677.3), SDC1 (GenScript; NM_002997.4), or NES (gift of A. Ting, Stanford; Addgene, #49386) (Fig. 1A) (82). We first generated a constitutively active piggyBac (gift of J. LoTurco)-based APEX2 universal vector [pPBCAG-Flag-APEX2-C1 and pPBCAG-Flag-APEX2-N1, containing multiple cloning sites (MCSs) at the linker region (83): APEX2-C1: Xho I, Nhe I, and Not I; APEX2-N1: Eco RI, Xho I, and Nhe I] by PCR amplification of Flag-APEX2 (Addgene, #49386) using primers (APEX2-C1: Clo-5EcoRI-Flag-APEX2-Fw and Clo-3NotI-Linker-APEX2-Rv; APEX2-N1: Clo-5EcoRI-Linker-Flag-APEX2-Fw and Clo-3NotI-Stop-APEX2-Rv). For APEX2-N1, a stop codon (TAA) was added at the end of the APEX2 sequence. Amplified products were cloned into pPBCAG digested with Eco RI and Not I. Next, we performed polymerase chain reaction (PCR) amplification of respective APEX2 tags from each donor construct (hPODXL: Clo-EcoRI-5-PODXL-Fw and Clo-NheI-3-PODXL-Rv; hEZR: Clo-EcoRI-5-hEZR-Fw and Clo-NheI-2xGS-3-hEZR-Rv; hATP1B1: Clo-NheI-5-ATP1B1-Fw and Clo-NotI-3-ATP1B1-Rv; hSDC1: Clo-XhoI-5-hSDC1-Fw and Clo-NheI-2xGS-3-hSDC1-Rv), which were then subcloned into the APEX2 universal vector (digested with Eco RI and Nhe I for hPODXL and hEZR; Nhe I and Not I for hATP1B1; Xho I and Nhe I for SDC1) to generate pPBCAG-Flag-APEX2-N1-hPODXL (hPODXL-APEX2), Flag-APEX2-N1-hEZR (hEZR-APEX2), Flag-APEX2-C1-hATP1B1 (APEX2-hATP1B1), and Flag-APEX2-N1-hSDC1 (hSDC1-APEX2). pPBCAG-Flag-APEX2-NES was generated

by PCR-amplifying Flag-APEX2-NES sequence (Addgene, #49386; using primers Clo-5EcoRI-Flag-APEX2-Fw and Clo-NotI-3-NES-Rv) and by subcloning the amplified product into the pPBCAG backbone digested with Eco RI and Not I. These APEX2 constructs were PCR-amplified (primers: PODXL, TOPO-5-PODXL-Fw and TOPO-3-APEX-Rv; EZR, TOPO-CACC-5-hEZR-Fw and TOPO-3-APEX-Rv; ATP1B1, TOPO-5-Flag-Fw and TOPO-3-ATP1B1-Rv; SDC1, TOPO-5-hSDC1-Fw and TOPO-3-APEX-Rv; NES, TOPO-5-Flag-Fw and TOPO-3-NES-Rv) and subcloned into pENTR/D-TOPO (Life Technologies), which were then cloned into PB-TA-ERP2 (Addgene, #80477) destination vector using the Gateway cloning system (Life Technologies).

mTnG constructs

The CAG promoter (Addgene, #48753; gift of A. Smith) (84) was excised using Xho I and Hind III and was subcloned into ePB-hUBC-Puro and ePB-hUBC-Neo (gift of A. Brivanlou) (49, 85) to generate ePB-CAG-hUBC-Puro and ePB-CAG-hUBC-Neo. Tcf-Lef-H2B-GFP (green fluorescent protein) [Addgene, #32610; gift of A.-K. Hadjantonakis (86) and pQC membrane TdTomato IX (Addgene, #37351; gift of C. Cepko (87))] were PCR-amplified (primers: H2B-GFP, Clo-HindIII-5-H2B-Fw and Clo-NotI-3-EGFP-Rv; mTdTtomato, Clo-HindIII-5-Palm-Fw and Clo-NotI-3-tdTomato-Rv) and cloned into ePBCAG-hUBC-Puro (H2B-GFP: Hind III and Not I) or ePBCAG-hUBC-Neo (mTdTtomato: Hind III and Not I).

Validation and gain-of-function constructs

MCSs of ePBCAG_hUBC vectors (containing unique Cla I, Hind III, and Not I sites 3' of CAG promoter) were expanded (ExMCS) by inserting synthetic DNA sequence (primers: Anneal-MCS expand-Fw and Anneal-MCS expand Rv) at Hind III and Not I sites to generate ePBCAG-ExMCS_hUBC-Puro/Neo (Cla I, Hind III, Bam HI, Pst I, Nhe I, Age I, Pac I, and Not I).

PCR-amplified hBASP1-HA (pRK7_CMV-hBASP1-HA, gift of K. Inoki) or hECE1-HA (GenScript, NM_001397.3) sequence (primers: BASP1, Clo-HindIII-5-hBASP1-Fw and Clo-NheI-stop-3-HA-Rv; ECE1, Clo-NheI-5-ECE1-Fw and Clo-NotI-HA-3-ECE1-Rv) was cloned into ePBCAG_hUBC-Puro (digested with: PODXL, Hind III and Nhe I; ECE1, Nhe I and Not I).

Universal ePBCAG-sfGFP-N1 and ePBCAG-mCherry-N1_hUBC-Puro vectors were generated through PCR amplification of sfGFP (Addgene, #54737; gift of G. Waldo and M. Davidson) and mCherry using shared primers [Clo-AgeI-3xGS-5-sfGFP-Fw and Clo-NotI-stop-3-sfGFP-Rv incorporating a stop (TAA) at the 3' end] and subcloned into ePBCAG_hUBC-Puro digested with Age I and Not I.

PCR-amplified hSNX27 (GenScript, NM_030918.6; Clo-HindIII-5-hSNX27_Fw and Clo-NheI-GS-3-hSNX27_Rv) or hAP1G1 [IDT, synthesized gBlocks fragment, NM_001030007.1 with CRISPR-resistant sequence modification in guide RNA (gRNA) target site and PAM sequence in its coding sequence 1—CCATCCGGACAGC—CCGAACCCAA changed to CGATCAGAACGGC~~C~~CGGACACAA; primers: Clo-BamHI-5-hAP1G1-Fw and Clo-NheI-3-hAP1G1-Rv] sequence was subcloned into ePBCAG-sfGFP-N1_hUBC-Puro (SNX27, Hind III and Nhe I) or ePBCAG-mCherry-N1_hUBC-Puro (AP1G1, Bam HI and Nhe I).

PCR-amplified mEzr (IMAGE: 6826190), EGFP-hCDC42 [Addgene, #12975; gift of G. Bokoch (88)], or hRAB35 [Addgene, #47424; gift of P. McPherson (89); primers: Clo-dTOPO-EGFP-fw and Clo_dTOPO_hRAB35_rv] product was subcloned into pENTR/dTOPO, which was then cloned into PB-TAC-ERP2 (mEzr; Addgene,

#80478) or into the PB-TA-ERP2 (hRAB35) destination vector using the Gateway cloning system.

PiggyBac-CRISPR-Cas9 (pBACON) constructs

A piggyBac-CRISPR-Cas9 (pBACON) vector that contains SpCas9-T2A-GFP and hU6-gRNA expression cassettes flanked by piggyBac transposon terminal repeat elements (pBACON-GFP), which allows subcloning of annealed oligos containing gRNA sequence at Bbs I site, has been previously described (32). In addition, a pBACON system expressing SpCas9-T2A-puro (pBACON-puro) was generated by PCR amplification of SpCas9-T2A-Puro [Addgene, #62988; gift of Feng Zhang (90); primers: Clo_5NheI_#115fw and Clo_3NotI#116rv]. gRNA targeting genomic sites and oligo sequences to generate pBACON-GFP-hEZR (CRISPR_hEz_E3#1_s and CRISPR_hEz_E3#1_as), pBACON-GFP-hAPIG1 (CRISPR_hAPIG1_E1#1_s and CRISPR_hAPIG1_E1#1_as), and pBACON-GFP-hLAMTOR1 (gRNA_hLAMTOR1_E2_s and gRNA_hLAMTOR1_E2_as), as well as pBACON-puro-hRAB35 (gRNA_hRAB35_E3_s and gRNA_hRAB35_E3_as), pBACON-puro-hSNX27 (gRNA_hSNX27_E2_s and gRNA_hSNX27_E2_as), and pBACON-puro-hBASP1 (gRNA_hBASP1_CDS1_s and gRNA_hBASP1_CDS1_as) are found in fig. S5 and table S4B; these were designed using publicly available tools (<http://crispr.mit.edu> or https://idtdna.com/site/order/designtool/index/CRISPR_CUSTOM).

PiggyBac-based transgenic and genome-edited hESC lines

PiggyBac constructs (3 μ g) and pCAG-ePBase (1 μ g; gift from A. Brivanlou) were cotransfected into H9 hESC (70,000 cells cm^{-2}) using GeneJammer transfection reagent (Agilent Technologies). To enrich for successfully transfected cells, fluorescence-activated cell sorting (FACS) or drug selection (puromycin, 2 μ g/ml for 4 days; neomycin, 250 μ g/ml for 10 days) was performed 48 to 72 hours after transfection. hESC stably expressing each construct maintained the expression of pluripotency markers and formed hPSC-cysts, unless otherwise noted. For inducible constructs, DOX (500 ng/ml) treatment was performed for 24 hours before harvesting for all experiments, unless otherwise noted.

During pBACON-based genome editing, GFP⁺ FACS-sorted or puro-selected cells were cultured at low density (300 cells cm^{-2}) for clonal selection. Established colonies were manually picked and expanded for screening indel mutations using PCR amplification of a region spanning the targeted gRNA region (primers: hEZR, PCR_EZRIN_RI_fw and PCR_EZRIN_NI_rv; hRAB35, Seq_hRAB35_fw and Seq_hRAB35_rv; hSNX27, Genotype_hSNX27-Exon2_EcoRI_Fw and Genotype_hSNX27-Exon2_NotI_Rv; hBASP1, Genotype_hBASP1-Exon2_EcoRI_Fw and Genotype_hBASP1-Exon2_NotI_Rv; APIG1, PCR_hAPIG1-CDS1_RI_fw and PCR_hAPIG1-CDS1_NI_rv), which were subcloned into pPBCAG-GFP (83) at Eco RI and Not I sites, and sequenced (Seq-3'TR-pPB-Fw). Genomic DNA was isolated from individual clones using DirectPCR Lysis Reagent (Tail) (VIAGEN). At least 12 to 15 bacterial colonies were sequenced to confirm genotypic clonality. Western blots for each protein were also performed to further validate the KO lines. Note that the RAB35-KO line was not validated by Western blot because of the lack of a specific antibody. KO analyses were performed using at least two distinct clonal lines per targeted gene. Control cells are H9 hESC in all loss-of-function experiments.

APEX2 labeling and sample preparation

d3 hPSC-cysts were generated from five different APEX2 hESC lines described in Fig. 1A. DOX (2 μ g/ml) was added at d2 to induce APEX2 transgene expression for 24 hours. On d3, mTeSR1 containing

500 μ M BP (BP/biotinyl tyramide, AdipoGen) and 1% Geltrex was added for 1 hour at 37°C, 5% CO₂. Hydrogen peroxide (H₂O₂) was then added directly into the medium to a final concentration of 1 mM for 90 s at room temperature to initiate biotinylation, immediately followed by 3 \times washes using quencher solution [10 mM sodium ascorbate (Spectrum Chemical), 10 mM sodium azide (Sigma-Aldrich), and 5 mM Trolox (Sigma-Aldrich) in Dulbecco's PBS (Gibco)]. For microscopy, hPSC-cysts were fixed for immunostaining using streptavidin conjugated to Alexa Fluor dye (Thermo Fisher Scientific).

To prepare proteomic samples, 10 hPSC-cyst samples were prepared individually (shown in Fig. 2, A and B) in two tissue culture-treated 100-mm dishes (Thermo Fisher Scientific, precoated with 1% Geltrex). For each plate, 2.2 \times 10⁶ hPSC (DOX-inducible stable lines) were plated (totaling 4.4 \times 10⁶ hPSC per sample) to obtain approximately 20 \times 10⁶ cells at d3, sufficient for at least 5.0 mg of total protein. At d3 (DOX added at d2), APEX2 labeling and quenching steps were performed as described above. After three times of quenching, APEX2-labeled hPSC-cysts from two dishes were resuspended as a pool in quencher buffer and centrifuged at 500g for 5 min to collect the cell pellet for lysis in radioimmunoprecipitation assay (RIPA) lysis buffer (Pierce) containing 1 \times Halt protease inhibitor cocktail (Thermo Fisher Scientific), 1 mM phenylmethylsulfonyl fluoride (Sigma-Aldrich), 10 mM sodium azide, 10 mM sodium ascorbate, and 5 mM Trolox. Cell lysates were centrifuged at 15,000g for 15 min at 4°C, and supernatant was collected for enriching biotinylated proteins using streptavidin beads.

The Pierce 660-nm assay (Pierce) was used to quantify protein concentrations in sample supernatants. Streptavidin-coated magnetic beads (Pierce) were first washed twice with RIPA lysis buffer. For each sample, 5.0 mg of total protein was incubated with 500 μ l of streptavidin beads overnight at 4°C with gentle rotation. Beads were subsequently pelleted using MagnaRack (Thermo Fisher Scientific) and washed twice with RIPA lysis buffer, once with 1 M KCl, once with 0.1 M Na₂CO₃, once with 2 M urea in 10 mM Tris-HCl (pH 8.0), twice with RIPA lysis buffer, and three times with PBS. Last, PBS was removed as much as possible and beads were frozen in -80°C before performing on-bead digestion, TMT labeling, peptide pooling, fractionation, and LC-MS/MS. All these steps were performed at 4°C unless otherwise noted. Alternatively, biotinylated proteins bound to streptavidin beads were eluted by boiling in homemade 1 \times protein loading buffer (34) supplemented with 2 mM biotin (Sigma-Aldrich, B4501) and 20 mM dithiothreitol (DTT; Sigma-Aldrich) for 15 min. Beads were pelleted by magnetic rack, and the supernatant was collected. These samples were loaded and separated on 10% SDS-polyacrylamide gel electrophoresis (PAGE) gels and processed by Coomassie staining, following the manufacturer's protocol (QC Colloidal Coomassie, Bio-Rad) to validate successful enrichment of biotinylated proteins. As shown in Fig. 1G, some background biotinylation signal was seen in samples lacking H₂O₂ (lanes 2, 4, 6, 8, and 10), untransfected H9 (lane 11), and ATP1B1-APEX2 without DOX induction (lane 12), likely due to background peroxidase activity and 16-hour incubation with streptavidin beads.

Western blot

SDS-PAGE gels (10% or gradient gel, 4 to 20%, Bio-Rad) and polyvinylidene difluoride membranes were used. Membranes were blocked using Intercept (TBS) Blocking Buffer (LI-COR), total protein quantification was performed by using Revert 700 Total Protein Stain (LI-COR), and primary antibody overnight incubation was performed at 4°C,

followed by 2-hour IRDye (LI-COR) secondary antibody incubation. Biotinylated proteins were detected and quantified by streptavidin-IRDye conjugate (LI-COR). Blots were imaged using LI-COR Odyssey Infrared Imaging system. Alternatively, for figs. S1A and S8E, after SDS-PAGE, samples were processed as described by Zysnarski *et al.* (91), with the exception that before blocking, total protein quantification was performed with Ponceau S staining imaged on an Azure c600 imaging system (Azure Biosciences) with blue light settings.

Quantitative MS

Proteins bound to streptavidin beads were digested by trypsin following the standard on-bead trypsin digestion workflow (92). Samples were proteolysed and labeled with TMT 10-plex by following the manufacturer's protocol (Thermo Fisher Scientific) with minor modifications. Ten was the highest number of available isobaric mass tags when our spatial proteomics assay was performed in 2018. Briefly, upon reduction [10 mM DTT in 0.1 M Triethylammonium bicarbonate (TEAB); 45°C, 30 min] and alkylation (55 mM 2-chloroacetamide in 0.1 M TEAB; room temperature, 30 min in dark) of cysteines, the proteins were digested overnight with trypsin (1:25; enzyme:protein) at 37°C, with constant mixing using a thermomixer. Proteolysis was stopped by adding 0.2% trifluoroacetic acid and peptides were desalted using SepPak C18 cartridge (Waters Corp). The desalted peptides were dried in Vacufuge (Eppendorf) and reconstituted in 100 μ l of 0.1 M TEAB. The TMT 10-plex reagents were dissolved in 41 μ l of anhydrous acetonitrile, and labeling was performed by transferring the entire digest to TMT reagent vial and incubating at room temperature for 1 hour. The reaction was quenched by adding 8 μ l of 5% hydroxyl amine and further 15-min incubation. Labeled samples were mixed together and dried using a vacufuge. An offline fractionation of the combined sample into six fractions was performed using a high-pH reversed-phase peptide fractionation kit according to the manufacturer's protocol (Pierce; catalog no. 8488). Fractions were dried and reconstituted in 12 μ l of 0.1% formic acid/2% acetonitrile in preparation for LC-MS/MS analysis. Sample-to-TMT channel information is shown in Fig. 2B.

For superior quantitation accuracy, we used multinotch-MS3 (92), which minimizes the reporter ion ratio distortion resulting from fragmentation of co-isolated peptides during MS analysis. Orbitrap Fusion (Thermo Fisher Scientific) and RSLC Ultimate 3000 nano-UPLC (Dionex) were used to acquire the data. The sample (2 μ l) was resolved on a PepMap RSLC C18 column (75 μ m inside diameter \times 50 cm; Thermo Fisher Scientific) at a flow rate of 300 nl/min using 0.1% formic acid/acetonitrile gradient system (2 to 22% acetonitrile in 110 min; 22 to 40% acetonitrile in 25 min; 6-min wash at 90% followed by 25-min re-equilibration) and directly sprayed onto the mass spectrometer using EasySpray source (Thermo Fisher Scientific). The mass spectrometer was set to collect one MS1 scan (Orbitrap; 120,000 resolution; AGC target, 2×10^5 ; max IT, 50 ms) followed by data-dependent, "Top Speed" (3 s) MS2 scans (collision-induced dissociation; ion trap; NCD 35; AGC, 5×10^3 ; max IT, 100 ms). For multinotch-MS3, top 10 precursors from each MS2 were fragmented by higher-energy-collisional-dissociation (HCD) followed by Orbitrap analysis [NCE 55; 60,000 resolution; AGC, 5×10^4 ; max IT, 120 ms; 100 to 500 m/z (mass/charge ratio) scan range].

Ratiometric analysis of proteomic data

Proteome Discoverer (v2.1; Thermo Fisher Scientific) was used for initial data analyses. MS2 spectra were searched against the SwissProt

human protein database (downloaded on 4 December 2018; 20331 reviewed entries) using the following search parameters: MS1 and MS2 tolerance were set to 10 parts per million and 0.6 Da, respectively; carbamidomethylation of cysteines (57.02146 Da) and TMT labeling of lysine and N termini of peptides (229.16293 Da) were considered static modifications; oxidation of methionine (15.9949 Da) and deamidation of asparagine and glutamine (0.98401 Da) were considered variable. Identified proteins and peptides were filtered to retain only those that passed $\leq 1\%$ FDR threshold and ≥ 2 unique peptides. Quantitation was performed using high-quality MS3 spectra using the Reporter Ion Quantifier Node of Proteome Discoverer (average signal-to-noise ratio of 10 and $< 30\%$ isolation interference).

Specific TMT ratios were normalized using known 495 mitochondria matrix soluble proteins (18): 198 were found in the hPSC-cyst dataset for ratiometric analyses in filter 1 [PODXL-APEX2 #1/negative (126/131), PODXL-APEX2 #2/negative (127N/131), EZR-APEX2/negative (127C/131), APEX2-ATP1B1 #1/negative (128N/131), APEX2-ATP1B1 #2/negative (128C/131), and SDC1-APEX2/negative (129N/131)], filter 2 [PODXL-APEX2 #1/APEX2-NES #1 (126/129C), PODXL-APEX2 #2/APEX2-NES #2 (127N/130N), EZR-APEX2/APEX2-NES #3 (127C/130C), APEX2-ATP1B1 #1/APEX2-NES #1 (128 N/129C), APEX2-ATP1B1 #2/APEX2-NES #2 (128C/130N), and SDC1-APEX2/APEX2-NES #3 (127C/129N)], and filters 3 and 4 [PODXL-APEX2 #1/APEX2-ATP1B1 #1 (126/128N), PODXL-APEX2 #2/APEX2-ATP1B1 #2 (127N/128C), and EZR-APEX2/SDC1-APEX2 (127C/129 N)]. In each TMT ratio, the median ratio of 198 mitochondrial matrix soluble proteins was calculated (PODXL-APEX2 #1/negative, 8.294; PODXL-APEX2 #2/negative, 7.786; EZR-APEX2/negative, 16.925; APEX2-ATP1B1 #1/negative, 14.876; APEX2-ATP1B1 #2/negative, 17.286; SDC1-APEX2/negative, 12.214; PODXL-APEX2 #1/APEX2-NES #1, 0.280; PODXL-APEX2 #2/APEX2-NES #2, 0.271; EZR-APEX2/APEX2-NES #3, 0.568; APEX2-ATP1B1 #1/APEX2-NES #1, 0.493; APEX2-ATP1B1 #2/APEX2-NES #2, 0.607; SDC1-APEX2/APEX2-NES #3, 0.397; PODXL-APEX2 #1/APEX2-ATP1B1 #1, 0.566; PODXL-APEX2 #2/APEX2-ATP1B1 #2, 0.447; EZR-APEX2/SDC1-APEX2, 1.450); all proteins in each TMT ratio were divided using these values to generate normalized hPSC-cyst dataset (ratios in \log_2 scale).

In filter 1 (F1) and filter 2 (F2), true positive (TP-F1 and TP-F2), proteins with UniProt "cell membrane" and "plasma membrane" annotations and false positive [FP, 198 mitochondrial matrix soluble proteins (filter 1) and all proteins in the list lacking UniProt "cell membrane" and "plasma membrane" annotations (filter 2)] were defined. TP-F1 and TP-F2 are proteins that are known to localize in the membrane territory; FP-F1 are proteins that are predicted to be nonbiotinylated by APEX2 fusion constructs in this study; FP-F2 consists of proteins that can be biotinylated by our APEX2 constructs but are not predicted to be proximal to membrane territories. True-positive rate (TPR) and false-positive rate (FPR) were calculated for each ratio [TPR = TP/(TP + FN); FPR = FP/(FP + TN); FN, false negative; TN, true negative]; these values were used to generate ROC curves to test the suitability of TP and FP for each ratiometric analysis based on the area under the curve (AUC; a commonly used statistic that calculates the area under the ROC curve and quantifies the probability in which a randomly chosen positive case outranks a randomly chosen negative case), as well as to determine cutoffs at which the largest difference between TPR and FPR was observed {PODXL-APEX2 #1/negative [$\log_2(126/131) - \text{AUC} = 0.74$, cutoff = 0.4]; PODXL-APEX2 #2/negative [$\log_2(127N/131) - \text{AUC} = 0.74$,

cutoff = 0.477]; EZR-APEX2/negative [$\log_2(127C/131) - AUC = 0.77$, cutoff = 0.476]; APEX2-ATP1B1 #1/negative [$\log_2(128N/131) - AUC = 0.77$, cutoff = 0.361]; APEX2-ATP1B1 #2/negative [$\log_2(128C/131) - AUC = 0.77$, cutoff = 0.435]; SDC1-APEX2/negative [$\log_2(129N/131) - AUC = 0.75$, cutoff = 0.529]; PODXL-APEX2 #1/APEX2-NES #1 [$\log_2(126/129C) - AUC = 0.63$, cutoff = 0.505]; PODXL-APEX2 #2/APEX2-NES #2 [$\log_2(127N/130N) - AUC = 0.62$, cutoff = 0.727]; EZR-APEX2/APEX2-NES #3 [$\log_2(127C/130C) - AUC = 0.67$, cutoff = 0.359]; APEX2-ATP1B1 #1/APEX2-NES #1 [$\log_2(128N/129C) - AUC = 0.76$, cutoff = 0.229]; APEX2-ATP1B1 #2/APEX2-NES #2 [$\log_2(128C/130N) - AUC = 0.78$, cutoff = 0.244]; SDC1-APEX2/APEX2-NES #3 [$\log_2(129N/130C) - AUC = 0.69$, cutoff = 0.36]} (Fig. 2, D and E, and fig. S2, A to D). TPR-FPR is equivalent to the Youden index (93), which is a statistic commonly used to represent the performance of a dichotomous test. Larger values of the index mean better performance. For example, a value of 1 would mean the performance is perfect as there are no false positives or false negatives. During the analysis of filter 2, proteins that passed filter 1 were used.

Given a threshold parameter T , we have AUC and cutoff (CO) formulas as

$$\text{TPR}(T): T \rightarrow y(x)$$

$$\text{FPR}(T): T \rightarrow x$$

$$\text{AUC} = \int_{x=0}^1 \text{TPR}(\text{FPR}^{-1}(x)) dx$$

$$\text{CO} = \arg \max_t (\text{TPR}(t) - \text{FPR}(t))$$

In filters 3 and 4, FDR was used to determine cutoffs for proteins that passed filter 2 and filter 1, respectively. Apical (AP: proteins with UniProt “apical cell membrane” annotation plus known apical proteins: PRKCI, PRKCZ, CDC42, RAB11B, RDX, and PARD6) and basolateral (BL: proteins with UniProt “basolateral cell membrane” annotation plus known basolateral proteins: CDH1, CTN-NB1, ITGB1, LLGL1, DLG3, and SCRIB) membrane proteins were defined. Proteins with FDR lower than 0.2 were kept, meaning 80% of these proteins are expected true discoveries.

Given a threshold parameter T , we defined the FDR formula for apical proteins as

$$\text{FDR}(T) = \frac{\sum_{i \in V} I\left(\left\{\frac{\text{AP}}{\text{BL}}\right\}_i > T\right)}{\sum_{j \in R} I\left(\left\{\frac{\text{AP}}{\text{BL}}\right\}_j > T\right)}$$

where $I(\cdot)$ is an indicator function that takes value one when the statement is true and zero otherwise; V is the set of basolateral proteins, and R is the set of total proteins (apical plus basolateral proteins). The numerator is the number of basolateral proteins above a given apical/basolateral cutoff (T), and the denominator is the number of total proteins above the same given cutoff.

Similarly, we defined the FDR formula for basolateral proteins as

$$\text{FDR}(T) = \frac{\sum_{i \in W} I\left(\left\{\frac{\text{AP}}{\text{BL}}\right\}_i < T\right)}{\sum_{j \in R} I\left(\left\{\frac{\text{AP}}{\text{BL}}\right\}_j < T\right)}$$

where W is the set of apical proteins. The numerator is the number of apical proteins below a given apical/basolateral cutoff (T), and the denominator is the number of total proteins below the same given cutoff.

Following filter 3, apical or basolateral replicates were intersected to reveal the list of apical (250) and basolateral (252) membrane territory proteins. In addition, 139 proteins that were shared among PODXL-APEX2 #1 and #2 (likely due to unique vesicular localization) but were not included in the list of 250 apical proteins were added to the final apical list (389). The list of nonpolar proteins (30) was generated by identifying proteins found in both the apical (389) and basolateral (252) lists. To generate curated lists of proteins after filter 4, proteins that were also found in the list after filter 3 [apical (389) and basolateral (252) proteins; table S2, B and C] were excluded from the original lists of proteins after filter 4: These lists were then intersected to identify the lists of post-filter 4 apical (1628), basolateral (171), and nonpolar (597) proteins (table S3).

In Fig. 3A (top), the percentage of proteins with “plasma membrane” or “cell membrane” UniProt annotations (table S1B) were identified for each category: entire human proteome (20367, Swiss-Prot), apical proteome (389), and basolateral proteome (252). In Fig. 3A (bottom) and table S2 (E and F), proteins with evidence for apical or basolateral localization were identified based on UniProt and GO.

GO enrichment analysis

The final apical (389 or 250) or basolateral (252) proteomes were uploaded to the STRING database [string-db.org (94), analysis performed on 8 May 2020]: The top 10 GO terms (ranked by FDR) on Cellular Component, Biological Process, and Molecular Function were plotted (Fig. 3, C to E, and fig. S3, A to C).

Protein network analysis

The apical (389) and basolateral (252) proteome lists were uploaded to the STRING database (analysis performed on 26 May 2020) to generate a protein network based on the protein interaction and confidence scores, which were then imported to Cytoscape (v.3.8.0) for clustering (Markov clustering, inflation value: 3.5 for apical proteome and 3 for basolateral proteome) and generating network diagrams (Fig. 3, F and G). For simplification, the intercluster interactions and clusters with equal to or less than four proteins are not shown in Fig. 3 (F and G).

Statistical analyses

Graphs were generated using Prism 7 (GraphPad Software) or R, and statistical analyses were performed using Prism 7. In Fig. 4, at least 50 aggregates were counted per sample from three independent experiments (total of >150 aggregates): χ^2 test was performed. In Fig. 5G, 293 (control), 298 (KO), and 206 (KO rescue) aggregates from three independent experiments were examined for luminal shape, which were then statistically analyzed using χ^2 test. In Fig. 5E, a total of 116, 153, and 133 cells were counted for control, KO, and rescue groups, respectively, for analysis using one-way analysis of variance (ANOVA) with a Turkey’s multiple comparison test. In Fig. 6B and fig. S8A (middle), at least 50 cells were counted from three independent biological samples (total of >150 cells) per time point per cell line (analyzed using two-way ANOVA with Turkey’s multiple comparisons test). In fig. S8A (right), 176 control, 154 control expressing PODXL-mCherry, and 156 *AP1G1*-KO expressing PODXL-mCherry cells from three independent experiments were examined

for apicosome size, which were then statistically analyzed using χ^2 test. *, significant ($P < 0.0001$); NS, not significant. All experiments were repeated at least three times.

SUPPLEMENTARY MATERIALS

Supplementary material for this article is available at <http://advances.sciencemag.org/cgi/content/full/7/17/eabd8407/DC1>

[View/request a protocol for this paper from Bio-protocol.](#)

REFERENCES AND NOTES

- K. Taniguchi, I. Heemskerck, D. L. Gumucio, Opening the black box: Stem cell-based modeling of human post-implantation development. *J. Cell Biol.* **218**, 410–421 (2019).
- M. N. Shahbazi, M. Zernicka-Goetz, Deconstructing and reconstructing the mouse and human early embryo. *Nat. Cell Biol.* **20**, 878–887 (2018).
- T. Boroviak, J. Nichols, Primate embryogenesis predicts the hallmarks of human naïve pluripotency. *Development* **144**, 175–186 (2017).
- M. N. Shahbazi, A. Jedrusik, S. Vuoristo, G. Recher, A. Hupalowska, V. Bolton, N. N. M. Fogarty, A. Campbell, L. Devito, D. Ilic, Y. Khalaf, K. K. Niakan, S. Fishel, M. Zernicka-Goetz, Self-organization of the human embryo in the absence of maternal tissues. *Nat. Cell Biol.* **18**, 700–708 (2016).
- A. Deglincerti, G. F. Croft, L. N. Pietila, M. Zernicka-Goetz, E. D. Siggia, A. H. Brivanlou, Self-organization of the *in vitro* attached human embryo. *Nature* **533**, 251–254 (2016).
- T. O'Leary, B. Heindryckx, S. Lierman, M. Van der Jeught, G. Duggal, P. De Sutter, S. M. Chuva de Sousa, Derivation of human embryonic stem cells using a post-inner cell mass intermediate. *Nat. Protoc.* **8**, 254–264 (2013).
- T. O'Leary, B. Heindryckx, S. Lierman, D. van Bruggen, J. J. Goeman, M. Vandewoestyne, D. Deforce, S. M. de Sousa Lopes, P. De Sutter, Tracking the progression of the human inner cell mass during embryonic stem cell derivation. *Nat. Biotechnol.* **30**, 278–282 (2012).
- K. Taniguchi, Y. Shao, R. F. Townshend, Y.-H. Tsai, C. J. DeLong, S. A. Lopez, S. Gayen, A. M. Freddo, D. J. Chue, D. J. Thomas, J. R. Spence, B. Margolis, S. Kalantry, J. Fu, K. S. O'Shea, D. L. Gumucio, Lumen formation is an intrinsic property of isolated human pluripotent stem cells. *Stem Cell Rep.* **5**, 954–962 (2015).
- B. S. Freedman, C. R. Brooks, A. Q. Lam, H. Fu, R. Morizane, V. Agrawal, A. F. Saad, M. K. Li, M. R. Hughes, R. V. Werff, D. T. Peters, J. Lu, A. Baccei, A. M. Siedlecki, M. T. Valerius, K. Musunuru, K. M. McNagny, T. I. Steinman, J. Zhou, P. H. Lerou, J. V. Bonventre, Modelling kidney disease with CRISPR-mutant kidney organoids derived from human pluripotent epiblast spheroids. *Nat. Commun.* **6**, 8715 (2015).
- Y. Zheng, X. Xue, Y. Shao, S. Wang, S. N. Eshfahani, Z. Li, J. M. Muncie, J. N. Lakins, V. M. Weaver, D. L. Gumucio, J. Fu, Controlled modelling of human epiblast and amnion development using stem cells. *Nature* **573**, 421–425 (2019).
- L. M. Orre, M. Vesterlund, Y. Pan, T. Arslan, Y. Zhu, A. F. Woodbridge, O. Frings, E. Fredlund, J. Lehtio, SubCellBarCode: Proteome-wide mapping of protein localization and relocalization. *Mol. Cell* **73**, 166–182.e7 (2019).
- D. N. Itzhak, S. Tyanova, J. Cox, G. H. Borner, Global, quantitative and dynamic mapping of protein subcellular localization. *eLife* **5**, e16950 (2016).
- E. Lundberg, G. H. H. Borner, Spatial proteomics: A powerful discovery tool for cell biology. *Nat. Rev. Mol. Cell Biol.* **20**, 285–302 (2019).
- D. N. Itzhak, C. Davies, S. Tyanova, A. Mishra, J. Williamson, R. Antrobus, J. Cox, M. P. Weekes, G. H. H. Borner, A mass spectrometry-based approach for mapping protein subcellular localization reveals the spatial proteome of mouse primary neurons. *Cell Rep.* **20**, 2706–2718 (2017).
- Y. Brunner, D. Schwartz, Y. Couté, J.-C. Sanchez, Proteomics of regulated secretory organelles. *Mass Spectrom. Rev.* **28**, 844–867 (2009).
- P. S. Caceres, D. Gravotta, P. J. Zager, N. Dephoure, E. Rodriguez-Boulant, Quantitative proteomics of MDCK cells identify unrecognized roles of clathrin adaptor AP-1 in polarized distribution of surface proteins. *Proc. Natl. Acad. Sci. U.S.A.* **116**, 11796–11805 (2019).
- B. Tan, S. M. J. M. Yatim, S. Peng, J. Gunaratne, W. Hunziker, A. Ludwig, The mammalian crumbs complex defines a distinct polarity domain apical of epithelial tight junctions. *Curr. Biol.* **30**, 2791–2804.e6 (2020).
- H.-W. Rhee, P. Zou, N. D. Udeshi, J. D. Martell, V. K. Mootha, S. A. Carr, A. Y. Ting, Proteomic mapping of mitochondria in living cells via spatially restricted enzymatic tagging. *Science* **339**, 1328–1331 (2013).
- G. Mayer, M. Bendayan, Biotinyl-tyramide: A novel approach for electron microscopic immunocytochemistry. *J. Histochem. Cytochem.* **45**, 1449–1454 (1997).
- K. Taniguchi, Y. Shao, R. F. Townshend, C. L. Cortez, C. E. Harris, S. Meshinchi, S. Kalantry, J. Fu, K. S. O'Shea, D. L. Gumucio, An apicosome initiates self-organizing morphogenesis of human pluripotent stem cells. *J. Cell Biol.* **216**, 3981–3990 (2017).
- M. N. Shahbazi, A. Scialdone, N. Skorupska, A. Weberling, G. Recher, M. Zhu, A. Jedrusik, L. G. Devito, L. Noli, I. C. Macaulay, C. Buecker, Y. Khalaf, D. Ilic, T. Voet, J. C. Marioni, M. Zernicka-Goetz, Pluripotent state transitions coordinate morphogenesis in mouse and human embryos. *Nature* **552**, 239–243 (2017).
- A. I. McClatchey, ERM proteins at a glance. *J. Cell Sci.* **127**, 3199–3204 (2014).
- R. Kishore, G. Qin, C. Luedemann, E. Bord, A. Hanley, M. Silver, M. Gavin, Y. S. Yoon, D. Goukassian, D. W. Losordo, The cytoskeletal protein ezrin regulates EC proliferation and angiogenesis via TNF- α -induced transcriptional repression of cyclin A. *J. Clin. Invest.* **115**, 1785–1796 (2005).
- T. Jöns, H. K. Heim, U. Kistner, G. Ahnert-Hilger, SAP 97 is a potential candidate for basolateral fixation of ezrin in parietal cells. *Histochem. Cell Biol.* **111**, 313–318 (1999).
- R. Zhou, L. Zhu, A. Kodani, P. Hauser, X. Yao, J. G. Forte, Phosphorylation of ezrin on threonine 567 produces a change in secretory phenotype and repolarizes the gastric parietal cell. *J. Cell Sci.* **118**, 4381–4391 (2005).
- N. Dard, S. Louvet-Vallée, A. Santa-Maria, B. Maro, Phosphorylation of ezrin on threonine T567 plays a crucial role during compaction in the mouse early embryo. *Dev. Biol.* **271**, 87–97 (2004).
- R. A. Orlando, T. Takeda, B. Zak, S. Schmieder, V. M. Benoit, T. McQuistan, H. Furthmayr, M. G. Farquhar, The glomerular epithelial cell anti-adhesin podocalyxin associates with the actin cytoskeleton through interactions with ezrin. *J. Am. Soc. Nephrol.* **12**, 1589–1598 (2001).
- S. Maday, E. Anderson, H. C. Chang, J. Shorter, A. Satoh, J. Sfakianos, H. Fölsch, J. M. Anderson, Z. Walther, I. Mellman, A PDZ-binding motif controls basolateral targeting of syndecan-1 along the biosynthetic pathway in polarized epithelial cells. *Traffic* **9**, 1915–1924 (2008).
- J. Füllekrug, A. Shevchenko, A. Shevchenko, K. Simons, Identification of glycosylated marker proteins of epithelial polarity in MDCK cells by homology driven proteomics. *BMC Biochem.* **7**, 8 (2006).
- U. Fischer, J. Huber, W. C. Boelens, I. W. Mattaj, R. Lührmann, The HIV-1 Rev activation domain is a nuclear export signal that accesses an export pathway used by specific cellular RNAs. *Cell* **82**, 475–483 (1995).
- Y. Shao, K. Taniguchi, K. Gurdziel, R. F. Townshend, X. Xue, K. M. A. Yong, J. Sang, J. R. Spence, D. L. Gumucio, J. Fu, Self-organized amniogenesis by human pluripotent stem cells in a biomimetic implantation-like niche. *Nat. Mater.* **16**, 419–425 (2017).
- Y. Shao, K. Taniguchi, R. F. Townshend, T. Miki, D. L. Gumucio, J. Fu, A pluripotent stem cell-based model for post-implantation human amniotic sac development. *Nat. Commun.* **8**, 208 (2017).
- V. Hung, S. S. Lam, N. D. Udeshi, T. Svinikina, G. Guzman, V. K. Mootha, S. A. Carr, A. Y. Ting, Proteomic mapping of cytosol-facing outer mitochondrial and ER membranes in living human cells by proximity biotinylation. *eLife* **6**, e24463 (2017).
- V. Hung, N. D. Udeshi, S. S. Lam, K. H. Loh, K. J. Cox, K. Pedram, S. A. Carr, A. Y. Ting, Spatially resolved proteomic mapping in living cells with the engineered peroxidase APEX2. *Nat. Protoc.* **11**, 456–475 (2016).
- J. Li, S. Han, H. Li, N. D. Udeshi, T. Svinikina, D. R. Mani, C. Xu, R. Guajardo, Q. Xie, T. Li, D. J. Luginbuhl, B. Wu, C. N. McLaughlin, A. Xie, P. Kaewsapsak, S. R. Quake, S. A. Carr, A. Y. Ting, L. Luo, Cell-surface proteomic profiling in the fly brain uncovers wiring regulators. *Cell* **180**, 373–386.e15 (2020).
- V. Hung, P. Zou, H.-W. Rhee, N. D. Udeshi, V. Cracan, T. Svinikina, S. A. Carr, V. K. Mootha, A. Y. Ting, Proteomic mapping of the human mitochondrial intermembrane space in live cells via ratiometric APEX tagging. *Mol. Cell* **55**, 332–341 (2014).
- C. Stark, B.-J. Breitkreutz, T. Reguly, L. Boucher, A. Breitkreutz, M. Tyers, BioGRID: A general repository for interaction datasets. *Nucleic Acids Res.* **34**, D535–D539 (2006).
- F. Martin-Belmonte, A. Gassama, A. Datta, W. Yu, U. Rescher, V. Gerke, K. Mostov, PTEN-mediated apical segregation of phosphoinositides controls epithelial morphogenesis through Cdc42. *Cell* **128**, 383–397 (2007).
- M. A. Schlüter, C. S. Pfarr, J. Pieczynski, E. L. Whiteman, T. W. Hurd, S. Fan, C.-J. Liu, B. Margolis, Trafficking of Crumbs3 during cytokinesis is crucial for lumen formation. *Mol. Biol. Cell* **20**, 4652–4663 (2009).
- M. A. Schlüter, B. Margolis, Apical lumen formation in renal epithelia. *J. Am. Soc. Nephrol.* **20**, 1444–1452 (2009).
- E. Rodriguez-Boulant, I. G. Macara, Organization and execution of the epithelial polarity programme. *Nat. Rev. Mol. Cell Biol.* **15**, 225–242 (2014).
- A. Yilmaz, C. Braverman-Gross, A. Bialer-Tsypin, M. Peretz, N. Benvenisty, Mapping gene circuits essential for germ layer differentiation via loss-of-function screens in haploid human embryonic stem cells. *Cell Stem Cell* **27**, 679–691.e6 (2020).
- F. Nakatsu, K. Hase, H. Ohno, The role of the clathrin adaptor AP-1: Polarized sorting and beyond. *Membranes* **4**, 747–763 (2014).
- A. Sanger, J. Hirst, A. K. Davies, M. S. Robinson, Adaptor protein complexes and disease at a glance. *J. Cell Sci.* **132**, jcs222992 (2019).
- W. Li, R. Puertollano, J. S. Bonifacio, P. A. Overbeek, E. T. Everett, Disruption of the murine *Ap2 β* gene causes nonsyndromic cleft palate. *Cleft Palate Craniofac. J.* **47**, 566–573 (2010).
- E. E. Heldwein, E. Macia, J. Wang, H. L. Yin, T. Kirchhausen, S. C. Harrison, Crystal structure of the clathrin adaptor protein 1 core. *Proc. Natl. Acad. Sci. U.S.A.* **101**, 14108–14113 (2004).

47. J. Huerta-Cepas, D. Szklarczyk, D. Heller, A. Hernández-Plaza, S. K. Forslund, H. Cook, D. R. Mende, I. Letunic, T. Rattei, L. J. Jensen, C. von Mering, P. Bork, eggNOG 5.0: A hierarchical, functionally and phylogenetically annotated orthology resource based on 5090 organisms and 2502 viruses. *Nucleic Acids Res.* **47**, D309–D314 (2019).
48. D. Zizioli, C. Meyer, G. Guhde, P. Saftig, K. von Figura, P. Schu, Early embryonic death of mice deficient in γ -adaplin. *J. Biol. Chem.* **274**, 5385–5390 (1999).
49. F. Etoc, J. Metzger, A. Ruzo, C. Kirst, A. Yoney, M. Z. Ozair, A. H. Brivanlou, E. D. Siggia, A balance between secreted inhibitors and edge sensing controls gastruloid self-organization. *Dev. Cell* **39**, 302–315 (2016).
50. M. Klingeborn, W. M. Dismuke, N. P. Skiba, U. Kelly, W. D. Stamer, C. Bowes Rickman, Directional exosome proteomes reflect polarity-specific functions in retinal pigmented epithelium monolayers. *Sci. Rep.* **7**, 4901 (2017).
51. A. Paradelo, S. B. Bravo, M. Henriquez, G. Riquelme, F. Gavilanes, J. M. Gonzalez-Ros, J. P. Albar, Proteomic analysis of apical microvillous membranes of syncytiotrophoblast cells reveals a high degree of similarity with lipid rafts. *J. Proteome Res.* **4**, 2435–2441 (2005).
52. T. Laux, K. Fukami, M. Thelen, T. Golub, D. Frey, P. Caroni, GAP43, MARCKS, and CAP23 modulate PI(4,5)P₂ at plasmalemmal rafts, and regulate cell cortex actin dynamics through a common mechanism. *J. Cell Biol.* **149**, 1455–1472 (2000).
53. D. Frey, T. Laux, L. Xu, C. Schneider, P. Caroni, Shared and unique roles of CAP23 and GAP43 in actin regulation, neurite outgrowth, and anatomical plasticity. *J. Cell Biol.* **149**, 1443–1454 (2000).
54. F. Grammer, S. K. Ramakrishnan, M. M. Rinschen, A. A. Larionov, M. Syed, H. Khatib, M. Roerden, J. O. Sass, M. Helmstaedt, D. Osenberg, L. Kuhne, O. Kretz, N. Wanner, F. Jouret, T. Benzing, F. Artunc, T. B. Huber, F. Theilig, mTOR regulates endocytosis and nutrient transport in proximal tubular cells. *J. Am. Soc. Nephrol.* **28**, 230–241 (2017).
55. F. Steinberg, M. Gallon, M. Winfield, E. C. Thomas, A. J. Bell, K. J. Heesom, J. M. Tavare, P. J. Cullen, A global analysis of SNX27-retromer assembly and cargo specificity reveals a function in glucose and metal ion transport. *Nat. Cell Biol.* **15**, 461–471 (2013).
56. P. Temkin, B. Lauffer, S. Jager, P. Cimermancic, N. J. Krogan, M. von Zastrow, SNX27 mediates retromer tubule entry and endosome-to-plasma membrane trafficking of signalling receptors. *Nat. Cell Biol.* **13**, 715–721 (2011).
57. Y. Sancak, L. Bar-Peled, R. Zoncu, A. L. Markhard, S. Nada, D. M. Sabatini, Ragulator-Rag complex targets mTORC1 to the lysosomal surface and is necessary for its activation by amino acids. *Cell* **141**, 290–303 (2010).
58. P. S. Mrozowska, M. Fukuda, Regulation of podocalyxin trafficking by Rab small GTPases in 2D and 3D epithelial cell cultures. *J. Cell Biol.* **213**, 355–369 (2016).
59. K. Klinkert, M. Rocancourt, A. Houdusse, A. Echard, Rab35 GTPase couples cell division with initiation of epithelial apico-basal polarity and lumen opening. *Nat. Commun.* **7**, 11166 (2016).
60. V. Singh, J. Yang, J. Yin, R. Cole, M. Tse, D. E. Berman, S. A. Small, G. Petsko, M. Donowitz, Cholera toxin inhibits SNX27-retromer-mediated delivery of cargo proteins to the plasma membrane. *J. Cell Sci.* **131**, jcs218610 (2018).
61. V. Singh, J. Yang, B. Cha, T. E. Chen, R. Sarker, J. Yin, L. R. Avula, M. Tse, M. Donowitz, Sorting nexin 27 regulates basal and stimulated brush border trafficking of NHE3. *Mol. Biol. Cell* **26**, 2030–2043 (2015).
62. K. Bannert, P. Berlin, J. Reiner, H. Lemcke, R. David, R. Engelmann, G. Lamprecht, SNX27 regulates DRA activity and mediates its direct recycling by PDZ-interaction in early endosomes at the apical pole of Caco2 cells. *Am. J. Physiol. Gastrointest. Liver Physiol.* **318**, G854–G869 (2020).
63. K. Ling, S. F. Bairstow, C. Carbonara, D. A. Turbin, D. G. Huntsman, R. A. Anderson, Type I γ phosphatidylinositol phosphate kinase modulates adherens junction and E-cadherin trafficking via a direct interaction with μ 1B adaptin. *J. Cell Biol.* **176**, 343–353 (2007).
64. J. S. Bonifacino, Adaptor proteins involved in polarized sorting. *J. Cell Biol.* **204**, 7–17 (2014).
65. G. A. Castillon, P. Burriat-Couleru, D. Abegg, N. Criado Santos, R. Watanabe, Clathrin and AP1 are required for apical sorting of glycosyl phosphatidyl inositol-anchored proteins in biosynthetic and recycling routes in Madin-Darby canine kidney cells. *Traffic* **19**, 215–228 (2018).
66. N. D. Dwyer, C. E. Adler, J. G. Crump, N. D. L'Etoile, C. I. Bargmann, Polarized dendritic transport and the AP-1 μ 1 clathrin adaptor UNC-101 localize odorant receptors to olfactory cilia. *Neuron* **31**, 277–287 (2001).
67. G. Gillard, M. Shafaq-Zadah, O. Nicolle, R. Damaj, J. Pecreaux, G. Michaux, Control of E-cadherin apical localisation and morphogenesis by a SOAP-1/AP-1/clathrin pathway in *C. elegans* epidermal cells. *Development* **142**, 1684–1694 (2015).
68. D. Gravotta, A. Perez Bay, C. T. H. Jonker, P. J. Zager, I. Benedicto, R. Schreiner, P. S. Caceres, E. Rodriguez-Boulant, Clathrin and clathrin adaptor AP-1 control apical trafficking of megalin in the biosynthetic and recycling routes. *Mol. Biol. Cell* **30**, 1716–1728 (2019).
69. M. Shafaq-Zadah, L. Brocard, F. Solari, G. Michaux, AP-1 is required for the maintenance of apico-basal polarity in the *C. elegans* intestine. *Development* **139**, 2061–2070 (2012).
70. I. Bedzhov, M. Zernicka-Goetz, Self-organizing properties of mouse pluripotent cells initiate morphogenesis upon implantation. *Cell* **156**, 1032–1044 (2014).
71. A. C. Hick, A. S. Delmarcelle, M. Bouquet, S. Klotz, T. Copetti, C. Forez, P. Van Der Smissen, P. Sonveaux, J. F. Collet, O. Feron, P. J. Courtoy, C. E. Pierreux, Reciprocal epithelial:endothelial paracrine interactions during thyroid development govern follicular organization and C-cells differentiation. *Dev. Biol.* **381**, 227–240 (2013).
72. G. Kesavan, F. W. Sand, T. U. Greiner, J. K. Johansson, S. Kobberup, X. Wu, C. Brakebusch, H. Semb, Cdc42-mediated tubulogenesis controls cell specification. *Cell* **139**, 791–801 (2009).
73. A. Q. Ryan, C. J. Chan, F. Graner, T. Hiragi, Lumen expansion facilitates epiblast-primitive endoderm fate specification during mouse blastocyst formation. *Dev. Cell* **51**, 684–697.e4 (2019).
74. C. Zeng, S. Vanoni, D. Wu, J. M. Caldwell, J. C. Wheeler, K. Arora, T. K. Noah, L. Waggoner, J. A. Besse, A. N. Yamani, J. Uddin, M. Rochman, T. Wen, M. Chehade, M. H. Collins, V. A. Mulkada, P. E. Putnam, A. P. Naren, M. E. Rothenberg, S. P. Hogan, Solute carrier family 9, subfamily A, member 3 (SLC9A3)/sodium-hydrogen exchanger member 3 (NHE3) dysregulation and dilated intercellular spaces in patients with eosinophilic esophagitis. *J. Allergy Clin. Immunol.* **142**, 1843–1855 (2018).
75. S. Shenolikar, J. W. Voltz, R. Cunningham, E. J. Weinman, Regulation of ion transport by the NHERF family of PDZ proteins. *Physiology* **19**, 362–369 (2004).
76. O. A. Weisz, E. Rodriguez-Boulant, Apical trafficking in epithelial cells: Signals, clusters and motors. *J. Cell Sci.* **122**, 4253–4266 (2009).
77. H. Folsch, Role of the epithelial cell-specific clathrin adaptor complex AP-1B in cell polarity. *Cell Logist.* **5**, e1074331 (2015).
78. K. D. Walton, A. M. Freddo, S. Wang, D. L. Gumucio, Generation of intestinal surface: An absorbing tale. *Development* **143**, 2261–2272 (2016).
79. A. S. Grosse, M. F. Pressprich, L. B. Curley, K. L. Hamilton, B. Margolis, J. D. Hildebrand, D. L. Gumucio, Cell dynamics in fetal intestinal epithelium: Implications for intestinal growth and morphogenesis. *Development* **138**, 4423–4432 (2011).
80. R. Santos, O. Ursu, A. Gaulton, A. P. Bento, R. S. Donadi, C. G. Bologna, A. Karlsson, B. Al-Lazikani, A. Hersey, T. I. Oprea, J. P. Overington, A comprehensive map of molecular drug targets. *Nat. Rev. Drug Discov.* **16**, 19–34 (2017).
81. S. I. Kim, F. Ocegueda-Yanez, C. Sakurai, M. Nakagawa, S. Yamanaka, K. Woltjen, Inducible transgene expression in human iPSC cells using versatile All-in-One piggyBac Transposons. *Methods Mol. Biol.* **1357**, 111–131 (2016).
82. S. S. Lam, J. D. Martell, K. J. Kamer, T. J. Deerink, M. H. Ellisman, V. K. Mootha, A. Y. Ting, Directed evolution of APEX2 for electron microscopy and proximity labeling. *Nat. Methods* **12**, 51–54 (2015).
83. F. Chen, J. LoTurco, A method for stable transgenesis of radial glia lineage in rat neocortex by piggyBac mediated transposition. *J. Neurosci. Methods* **207**, 172–180 (2012).
84. J. Betschinger, J. Nichols, S. Dietmann, P. D. Corrin, P. J. Paddison, A. Smith, Exit from pluripotency is gated by intracellular redistribution of the bHLH transcription factor Tfe3. *Cell* **153**, 335–347 (2013).
85. A. Lacoste, F. Berenshteyn, A. H. Brivanlou, An efficient and reversible transposable system for gene delivery and lineage-specific differentiation in human embryonic stem cells. *Cell Stem Cell* **5**, 332–342 (2009).
86. A. Ferrer-Vaquer, A. Piliszek, G. Tian, R. J. Aho, D. Dufort, A. K. Hadjantonakis, A sensitive and bright single-cell resolution live imaging reporter of Wnt/ss-catenin signaling in the mouse. *BMC Dev. Biol.* **10**, 121 (2010).
87. K. T. Beier, M. E. Samson, T. Matsuda, C. L. Cepko, Conditional expression of the TVA receptor allows clonal analysis of descendants from Cre-expressing progenitor cells. *Dev. Biol.* **353**, 309–320 (2011).
88. M. C. Subauste, M. Von Herrath, V. Benard, C. E. Chamberlain, T. H. Chuang, K. Chu, G. M. Bokoch, K. M. Hahn, Rho family proteins modulate rapid apoptosis induced by cytotoxic T lymphocytes and Fas. *J. Biol. Chem.* **275**, 9725–9733 (2000).
89. P. D. Allaire, M. Seyed Sadr, M. Chaineau, E. Seyed Sadr, S. Konefal, M. Fotouhi, D. Maret, B. Ritter, R. F. Del Maestro, P. S. McPherson, Interplay between Rab35 and Arf6 controls cargo recycling to coordinate cell adhesion and migration. *J. Cell Sci.* **126**, 722–731 (2013).
90. F. A. Ran, P. D. Hsu, J. Wright, V. Agarwala, D. A. Scott, F. Zhang, Genome engineering using the CRISPR-Cas9 system. *Nat. Protoc.* **8**, 2281–2308 (2013).
91. C. J. Zysnarski, S. Lahiri, F. T. Javed, J. Y. Martinez-Marquez, J. W. Trowbridge, M. C. Duncan, Adaptor protein complex-1 (AP-1) is recruited by the HEATR5 protein Laa1 and its co-factor Laa2 in yeast. *J. Biol. Chem.* **294**, 1410–1419 (2019).
92. G. C. McAlister, D. P. Nusinow, M. P. Jedrychowski, M. Wuhr, E. L. Huttlin, B. K. Erickson, R. Rad, W. Haas, S. P. Gygi, MultiNotch MS3 enables accurate, sensitive, and multiplexed detection of differential expression across cancer cell line proteomes. *Anal. Chem.* **86**, 7150–7158 (2014).
93. W. J. Youden, Index for rating diagnostic tests. *Cancer* **3**, 32–35 (1950).
94. D. Szklarczyk, A. L. Gable, D. Lyon, A. Junge, S. Wyder, J. Huerta-Cepas, M. Simonovic, N. T. Doncheva, J. H. Morris, P. Bork, L. J. Jensen, C. V. Mering, STRING v11: Protein-protein

association networks with increased coverage, supporting functional discovery in genome-wide experimental datasets. *Nucleic Acids Res.* **47**, D607–D613 (2019).

95. A. Chapman-Smith, J. E. Cronan Jr., In vivo enzymatic protein biotinylation. *Biomol. Eng.* **16**, 119–125 (1999).

Acknowledgments: We thank K. Woltjen (Kyoto University) for the piggyBac DOX-inducible vector (Addgene, #80477 and 80478), D. Louvard (Curie Institute Paris) for pDEST-hEZRIN, J. Loturco (University of Connecticut) for the constitutively active piggyBac backbone vector, A.-K. Hadjantonakis (Sloan Kettering Institute) for TCF/Lef:H2B-GFP (Addgene, #32610), G. Bokoch (Scripps) for pCDNA3-EGFP-Cdc42-wt (Addgene, #12975), C. Cepko (Harvard) for pQC membrane TdTomato IX (Addgene, #37351), A. Smith (Cambridge) for pPB-CAG-empty-pgk-hph (Addgene, #48753), K. Inoki (University of Michigan) for pRK7_CMV-hBASP1-HA, G. Waldo (Los Alamos National Laboratory) and M. Davidson (Florida State University) for sfGFP-N1 (Addgene, #54737), P. McPherson (McGill University) for human RAB35 (Addgene #47424), F. Zhang (MIT) for SpCas9(BB)-T2A-Puro (Addgene, #62988), A. Ting (Stanford) for NES (Addgene, #49386), and A. Brivanlou (Rockefeller University) for ePiggyBac transposase and DOX-inducible ePiggyBac constructs. We thank the University of Michigan Microscopy Core and Flow Cytometry Core, as well as the Medical College of Wisconsin Electron Microscopy Core Facility (director, C. Wells). We also thank A. Carleton (DrawBioMedicine program) for hand-drawn schematics. **Funding:** This work was supported by NIH grants R01-HD098231 (K.T.), R01-HD102496 (K.T.), R01-GM129255 (M.C.D.), and R01-DK089933 (D.L.G.); MCW CBNA Start-up funds (K.T.); the University of Michigan Phi-Kappa-Phi Honor Society Student Grant (R.T.); the University of Michigan Synergy Grant (M.C.D. and D.L.G.); and the University of

Michigan Medical School Proteomics Resource Facility Pilot Project Program Grant (K.T.). Microscopic analyses were supported by 3R01GM120735-03S1, 3R01GM067180-14S1, and 3R35GM119544-03S1 (awarded, in the respective order, to A. Hudson, B. Hill, and M. Scaglione, Medical College of Wisconsin). S.W. was partially supported by the University of Michigan Mechanical Engineering Faculty Fellowship. **Author contributions:** S.W., P.Z., J.F., D.L.G., M.C.D., and K.T. designed experiments; S.W., C.-W.L., C.L.C., A.E.C., L.E.T., N.S., R.F.T., V.B., M.C.D., and K.T. performed experiments; S.W., J.F., D.L.G., M.C.D., and K.T. analyzed data and wrote the manuscript; J.F., D.L.G., M.C.D., and K.T. supervised the project; all authors contributed to the manuscript. **Competing interests:** The authors declare that they have no competing interest. **Data and materials availability:** All data needed to evaluate the conclusions in the paper are present in the paper and/or the Supplementary Materials. Additional information and reagents can be provided by the authors pending scientific review and a completed materials and transfer agreement, where appropriate: Requests should be directed to M.D. (mcduncan@med.umich.edu) and K.T. (ktaniguchi@mcw.edu).

Submitted 16 July 2020

Accepted 5 March 2021

Published 23 April 2021

10.1126/sciadv.abd8407

Citation: S. Wang, C.-W. Lin, A. E. Carleton, C. L. Cortez, C. Johnson, L. E. Taniguchi, N. Sekulovski, R. F. Townshend, V. Basrur, A. I. Nesvizhskii, P. Zou, J. Fu, D. L. Gumucio, M. C. Duncan, K. Taniguchi, Spatially resolved cell polarity proteomics of a human epiblast model. *Sci. Adv.* **7**, eabd8407 (2021).

Spatially resolved cell polarity proteomics of a human epiblast model

Sicong Wang, Chien-Wei Lin, Amber E. Carleton, Chari L. Cortez, Craig Johnson, Linnea E. Taniguchi, Nikola Sekulovski, Ryan F. Townshend, Venkatesha Basrur, Alexey I. Nesvizhskii, Peng Zou, Jianping Fu, Deborah L. Gumucio, Mara C. Duncan and Kenichiro Taniguchi

Sci Adv 7 (17), eabd8407.
DOI: 10.1126/sciadv.abd8407

ARTICLE TOOLS	http://advances.sciencemag.org/content/7/17/eabd8407
SUPPLEMENTARY MATERIALS	http://advances.sciencemag.org/content/suppl/2021/04/19/7.17.eabd8407.DC1
REFERENCES	This article cites 95 articles, 29 of which you can access for free http://advances.sciencemag.org/content/7/17/eabd8407#BIBL
PERMISSIONS	http://www.sciencemag.org/help/reprints-and-permissions

Use of this article is subject to the [Terms of Service](#)

Science Advances (ISSN 2375-2548) is published by the American Association for the Advancement of Science, 1200 New York Avenue NW, Washington, DC 20005. The title *Science Advances* is a registered trademark of AAAS.

Copyright © 2021 The Authors, some rights reserved; exclusive licensee American Association for the Advancement of Science. No claim to original U.S. Government Works. Distributed under a Creative Commons Attribution NonCommercial License 4.0 (CC BY-NC).