# UCSF

**UC San Francisco Previously Published Works**

**Title**

A Qualitative Change in the Transcriptome Occurs after the First Cell Cycle and Coincides with Lumen Establishment during MDCKII Cystogenesis

**Permalink**

https://escholarship.org/uc/item/0692q5c2

**Journal**

iScience, 23(10)

**ISSN**

2589-0042

**Authors**

Wang, Tianfang
Kwon, Sang-Ho
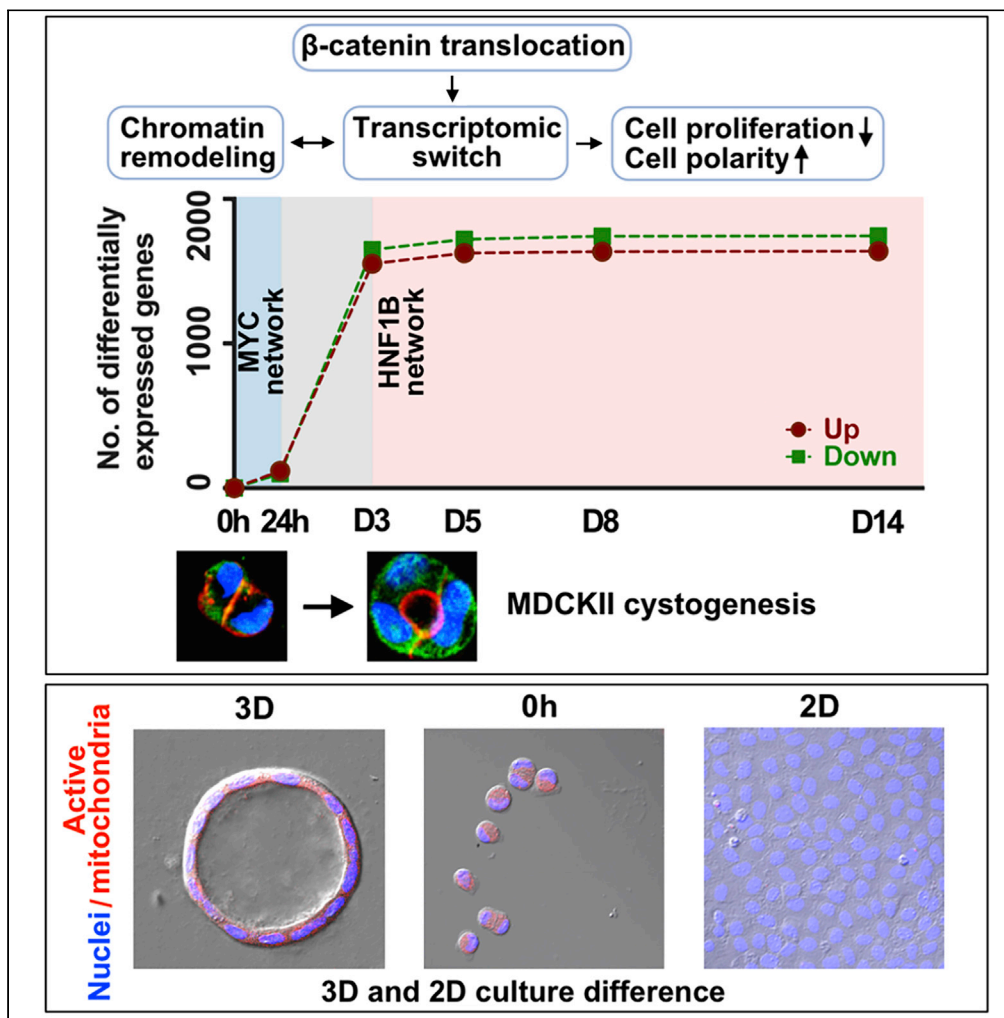Peng, Xiao
et al.

**Publication Date**

2020-10-01

**DOI**

10.1016/j.isci.2020.101629

Peer reviewed

**Article**

# A Qualitative Change in the Transcriptome Occurs after the First Cell Cycle and Coincides with Lumen Establishment during MDCKII Cystogenesis

Tianfang Wang, Sang-Ho Kwon, Xiao Peng, ..., Stephen Dalton, Keith E. Mostov, Shaying Zhao

szhao@uga.edu

**HIGHLIGHTS**

The transcriptome switches after the first cell cycle and during MDCKII lumenogenesis

The transcriptome switch is linked to β-catenin translocation and HNF1B activation

Chromatin accessibility decreases during MDCKII cystogenesis

Active mitochondria are maintained in 3D, but not 2D, epithelial morphogenesis

## Article

# A Qualitative Change in the Transcriptome Occurs after the First Cell Cycle and Coincides with Lumen Establishment during MDCKII Cystogenesis

Tianfang Wang,[1] Sang-Ho Kwon,[2,3] Xiao Peng,[2,5] Severine Urdy,[2,6] Zefu Lu,[4] Robert J. Schmitz,[4] Stephen Dalton,[1] Keith E. Mostov,[2] and Shaying Zhao[1,7,*]

## SUMMARY

**Madin-Darby canine kidney II (MDCKII) cells are widely used to study epithelial morphogenesis. To better understand this process, we performed time course RNA-seq analysis of MDCKII 3D cystogenesis, along with polarized 2D cells for comparison. Our study reveals a biphasic change in the transcriptome that occurs after the first cell cycle and coincides with lumen establishment. This change appears to be linked to translocation of β-catenin, supported by analyses with *AVL9*- and *DENND5A*-knockdown clones, and regulation by HNF1B, supported by ATAC-seq study. These findings indicate a qualitative change model for transcriptome remodeling during epithelial morphogenesis, leading to cell proliferation decrease and cell polarity establishment. Furthermore, our study reveals that active mitochondria are retained and chromatin accessibility decreases in 3D cysts but not in 2D polarized cells. This indicates that 3D culture is a better model than 2D culture for studying epithelial morphogenesis.**

## INTRODUCTION

Madin-Darby canine kidney (MDCK) cells represent a well-established cell line model for epithelial morphogenesis research (Dukes et al., 2011; Elia and Lippincott-Schwartz, 2009; Datta et al., 2011; McRae et al., 2017; Sottocornola et al., 2010; Pillman et al., 2018). Indeed, tens of thousands of papers have been published on using MDCK cells to study epithelial cell polarity, epithelial-mesenchymal transition (EMT) and cancer, tissue engineering, and other subjects. There are multiple strains of MDCK (Dukes et al., 2011). Among them, MDCKII is the most widely used and has been recommended for most studies (Dukes et al., 2011). We have been using MDCKII cells to validate our cancer driver candidates identified by dog-human comparisons (Tang et al., 2014; Li et al., 2014).

MDCKII can be cultured in both 3D and 2D conditions. In both cases, epithelial cell apico-basolateral polarity is established. However, 3D morphogenesis is a process of cystogenesis, forming a cyst that consists of a central lumen surrounded by a single layer of polarized cells (Li et al., 2014). 2D morphogenesis simply forms a monolayer of polarized cells, without lumenogenesis (Li et al., 2014).

Previous research has indicated the importance of both intracellular trafficking and cell cycle in MDCK 3D epithelial morphogenesis. Specifically, *de novo* lumen formation is regulated by Rab11a- and Cdc42-directed networks, which control mitotic spindle orientation and apical transport (Bryant et al., 2010; Rodriguez-Fraticelli et al., 2010). Modulation of cell proliferation rate contributes to the maintenance of the single-lumen phenotype of normal cysts (Cerruti et al., 2013).

Resembling cell differentiation at late stages (e.g., transit-amplifying cells to terminally differentiated cells in colonic crypt development when epithelial cell polarity is establishing [Sheaffer and Kaestner, 2012; Snippert et al., 2010]), gene expression plays a key role in MDCKII epithelial morphogenesis. A deep understanding of changes in the transcriptome over the course is thus important. Our literature search finds several relevant microarray studies. Two studies identify differentially expressed genes between 3D and 2D cells at 36 h (Galvez-Santisteban et al., 2012) and at 8 days after seeding (Wells et al., 2013), indicating the importance of synaptotagmin-like proteins in *de novo* lumenogenesis and interleukin-8 in 3D epithelial morphogenesis, respectively. Three additional studies investigate gene expression changes induced by

[1]Department of Biochemistry and Molecular Biology, Institute of Bioinformatics, University of Georgia, Athens, GA 30602, USA

[2]Department of Anatomy, University of California, San Francisco, CA 94143-2140, USA

[3]Department of Cellular Biology and Anatomy, Augusta University, Medical College of Georgia, 1460 Laney Walker Boulevard, CB2820A, Augusta, GA 30912, USA

[4]Department of Genetics, University of Georgia, Athens, GA 30602, USA

[5]Present address: PPD Laboratories, GMP Lab, Middleton, WI 53562, USA

[6]Present address: Univ. Lyon, ENS de Lyon, CNRS, Université Claude Bernard Lyon 1, Institut de Génomique Fonctionnelle deLyon, UMR 5242, 46 allée d'Italie, F-69364 Lyon Cedex 07, France

[7]Lead Contact

*Correspondence: szhao@uga.edu

https://doi.org/10.1016/j.isci. 2020.101629

hepatocyte growth factor (HGF) in 3D, 2D, or 2.5D culture conditions (Balkovetz et al., 2004; Kwon et al., 2011; Chacon-Heszele et al., 2014). HGF plays a role in epithelial tubulogenesis, where the cells initially undergo a partial EMT (Chacon-Heszele et al., 2014; O'Brien et al., 2004).

Although these microarray studies provide insightful information, several fundamental questions remain unanswered. For example, during 3D cystogenesis, does the transcriptome gradually change over the course or suddenly switch at a certain point? 3D cystogenesis can be divided into three stages: lumen establishing, lumen enlarging, and lumen maintenance (Li et al., 2014). What are the gene expression changes at each stage?

To answer these questions, we set out to perform full time course RNA sequencing (RNA-seq) analysis of MDCKII cystogenesis in 3D culture. Cells were seeded as a sparse single cell suspension to capture *de novo* lumenogensis, which initiates during the first cell division (Li et al., 2014). Fully polarized MDCKII cells in 2D culture were also included in the study for comparison.

## RESULTS

### Time Course RNA-Seq Analysis of MDCKII Cystogenesis

We conducted full time course analysis of MDCKII cystogenesis. Samples were taken at the time of seeding (0h), as well as culturing on top of Matrigel for 24 h, and 3, 5, 8, and 14 days after seeding (Figures 1A and S1). This design captures the three stages of cystogenesis established by our previous work (Li et al., 2014). These include: (1) lumen establishing, from seeding to the two- or more-cell stage (24 h to day 3); (2) lumen enlarging, with active oriented cell divisions (primarily from day 3 to day 8); and (3) lumen maintenance, with most cells ceasing to divide (after 8 days).

Once the lumen is established, apico-basolateral cell polarity is also built. This is illustrated with E-cadherin staining, which marks the lateral membrane, and F-actin staining, which marks the apical region (Figure 1A). Once built, the cell polarity is maintained throughout cystogenesis.

For comparison purposes, we also studied fully polarized MDCKII 2D cells. Specifically, MDCKII cells were cultured on Transwell filters (Figure 1B) until they reached over-confluence, where and when apico-basolateral polarity is well established (Figure 1B).

We included three biological repeats, which consist of three separate experiments starting with cells cultured at different times (Tables S1A and S1B), and performed quality control (QC) of each sample. First, we conducted qRT-PCR experiments with a panel of genes that include proliferation markers *MYC* and *CCND1*, apoptosis marker *BAX*, transcription factor and polarity marker *CTNNB1*, and cytoskeleton marker *ACTB*, as well as *PTEN* and *ERBB2*, which are important to lumen formation (Muthuswamy et al., 2001; Martin-Belmonte et al., 2007) (Figure 1C; Table S1C). Samples with expected expression patterns (e.g., higher *MYC* expression but lower *CTNNB1* expression for early time course samples but the opposite for late time course samples) were then subjected to RNA-seq analysis. Second, we performed a comprehensive QC of the RNA-seq data (Table S1B; Figure S1B), including examination of sequencing amount, base quality, GC content, sequence duplication, mapping rate and quality, rRNA and mitochondria contamination, and gene expression distribution. A total of 19 samples have passed our QC measures, with which we quantified the expression level of each gene and performed the analysis below.

### RNA-Seq Analysis Reveals a Qualitative Change in the Transcriptome during Lumen Establishment

We first performed a principal component analysis (PCA) with all 19 samples. This was done using only genes with FPKM (fragments per kilobase of transcript per million mapped) > 0 in all samples (making the expression distribution more normal) and FPKM >1 (expressed) in at least one sample (Table S2A), which yield 11,246 genes in total. The 19 samples are unambiguously ($p < 0.001$) grouped into three clusters, which are conveniently named early, 2D, and 3D (Figure 2A). The "early" cluster consists of all the early-stage samples, including the cells at seeding (0h) and 3D culture at 24 h afterward (3D24h). All other 3D samples make up the "3D" cluster, and all 2D samples constitute the "2D" cluster (Figure 2A). The PCA result is supported by two clustering analyses. First, the non-negative matrix factorization (NMF) strategy (Brunet et al., 2004) identifies three metagenes, with which the samples are unambiguously grouped into the same three clusters (Figure 2A). Second, unsupervised hierarchical clustering using the top-most
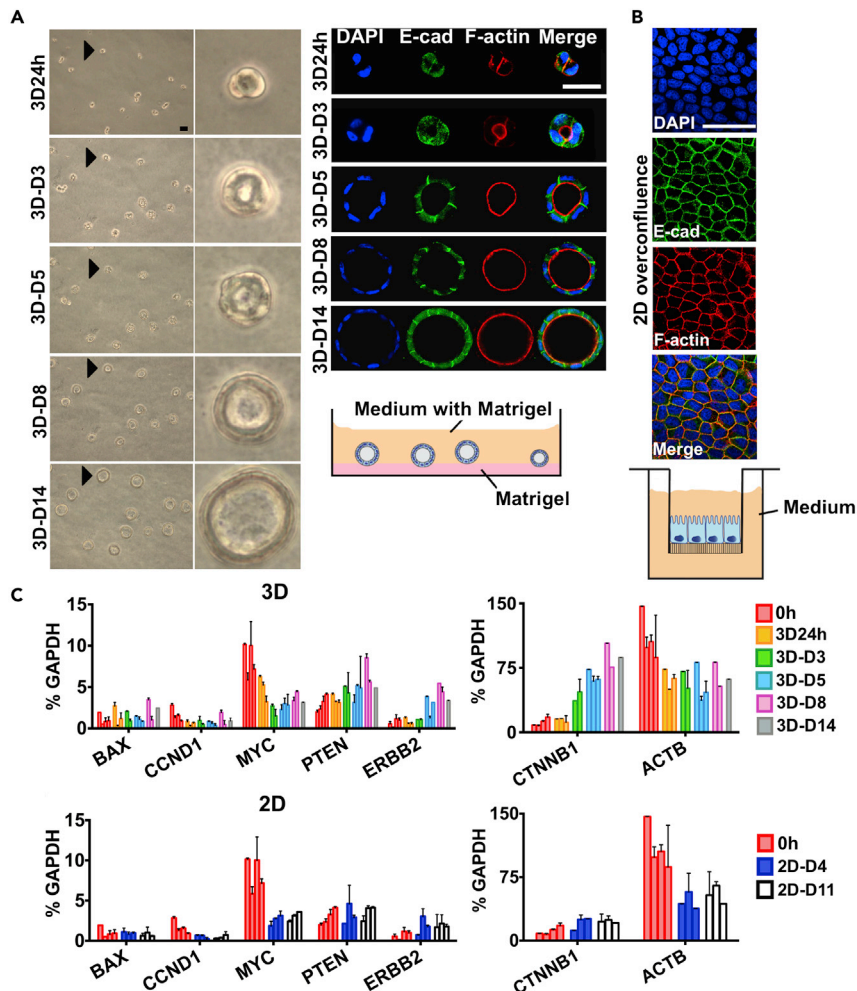
**Figure 1. Time Course RNA-Seq Analysis of MDCKII Cystogenesis**

(A) Time course 3D culture. Left: Representative bright-field images indicate time course 3D culture used for RNA-seq. Each arrow points to a cyst that is enlarged on the right. Right top: representative confocal images of the corresponding 3D culture shown in A, with markers indicated. Right bottom: 3D cell culture was performed as illustrated in the cartoon modified from a previous publication (Shamir and Ewald, 2014). Scale bar, 20 μm.

(B) Representative confocal images of fully polarized and over-confluent MDCKII 2D cells, cultured on Transwell filters as illustrated by the cartoon. Scale bar, 20 μm.

(C) Sample quality control by qRT-PCR with genes shown. Normalized gene expression against GAPDH is indicated for each condition specified. Data are represented as mean ± SD with three technical replicates. "0h" represents cells at seeding, "3D24h" represents 3D culture harvested at 24 h after seeding and so on.

See Figure S1 and Table S1.

variably expressed genes among the samples consistently identifies the same three clusters with high stability (Figures 2A and S2).

The results indicate a biphasic, rather than a linear, change among the 3D time course samples (Figure 2A). To further confirm this, we performed the same analysis but excluding all 2D samples. The aforementioned three methods (PCA, hierarchical, and NMF clustering) all group the time course samples into two clusters (Figure 2B), namely, the same early and 3D clusters as described above (Figure 2A).

To determine when the observed biphasic event takes place (Figure 2B), we identified genes that are differentially expressed (DE) between any two adjacent time points during 3D cystogenesis. Remarkably, 10 times more DE genes (>1,000 each for up- and down-regulation) were found between the 24 h and the day 3 samples, compared with other time intervals (Figure 2C; Table S2B). Moreover, the transcriptome
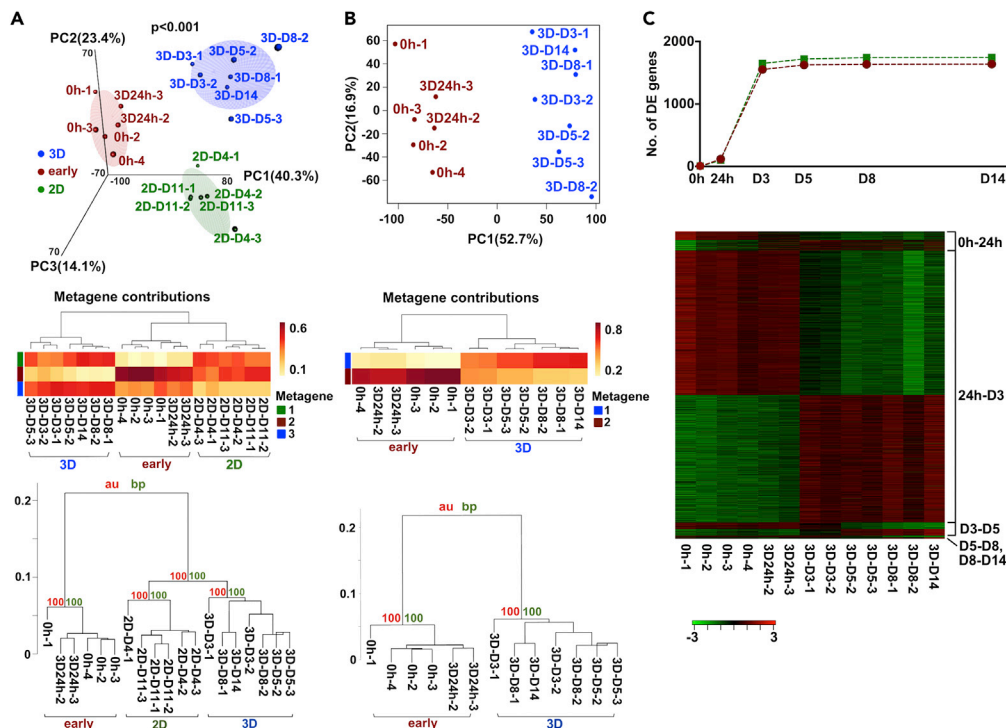
**Figure 2. RNA-Seq Analysis Reveals a Qualitative Change in the Transcriptome during Lumen Establishment**

(A) MDCKII samples were clustered into three groups (early, 3D, and 2D) by PCA (top), NMF (middle), and hierarchical clustering (bottom). PCA and NMF were performed with 11,246 genes, with each having FPKM>1 in at least one sample and FPKM>0 across all samples. The p value in the PCA plot was calculated via sample permutation, and the consensus clustering is shown for NMF. Hierarchical clustering was performed with top 2,000 variably expressed genes across the samples, with approximately unbiased (au) and bootstrap probability (bp) values shown in red and green, respectively. Bootstrapping was performed with permutations (see Transparent Methods).

(B) Time course 3D samples were grouped into two clusters by PCA (top), NMF (middle), and hierarchical clustering analysis (bottom).

(C) A qualitative change in the transcriptome occurs during the 24 h to day 3 transition period of 3D cystogenesis, marked by thousands of up- or down-regulated genes. The plot (top) indicates the total number of changed genes within each time period specified, whereas the heatmap (bottom) shows the expression levels of these genes, with up-regulation in red and down-regulation in green. The genes were identified with DESeq (Anders and Huber, 2010) and FDR <0.01. See Figure S2 and Table S2.

is largely stable before 24 h and after day 3 (Figure 2C). As the 24 h to day 3 period is when the lumen is becoming established (Figures 1A and S1), our data indicate that a qualitative change, which remodels the transcriptome, occurs during early lumenogenesis and immediately following the first cell division (Li et al., 2014).

## Each of the Early, 2D, and 3D Sample Clusters Harbors Distinct Molecular Features

To determine molecular features that characterize each of the three sample clusters indicated in Figure 2, we performed single sample gene set enrichment analysis (ssGSEA) with known cell polarity gene sets compiled from published studies (Bryant et al., 2010; Bryant and Mostov, 2008; Assemat et al., 2008; McCaffrey and Macara, 2012) (Table S3A). We noted that many are significantly up-regulated in 2D cells, including genes of cell-cell junctions and cell polarity complex (Figure 3A; Table S3A). Exceptions are cytoskeleton genes, which are up-regulated in early samples (likely due to very active cell division), and apical trafficking, which are up-regulated in 3D cells (likely because of lumen formation).

We also performed ssGSEA with major pathways and functions of epithelial morphogenesis compiled from databases (Table S3A) and publications (Kosinski et al., 2007; Fevr et al., 2007; Van der Flier et al., 2007; Benita et al., 2009; Koinuma et al., 2009; Mjelle et al., 2015). These include cell cycle, intracellular trafficking,
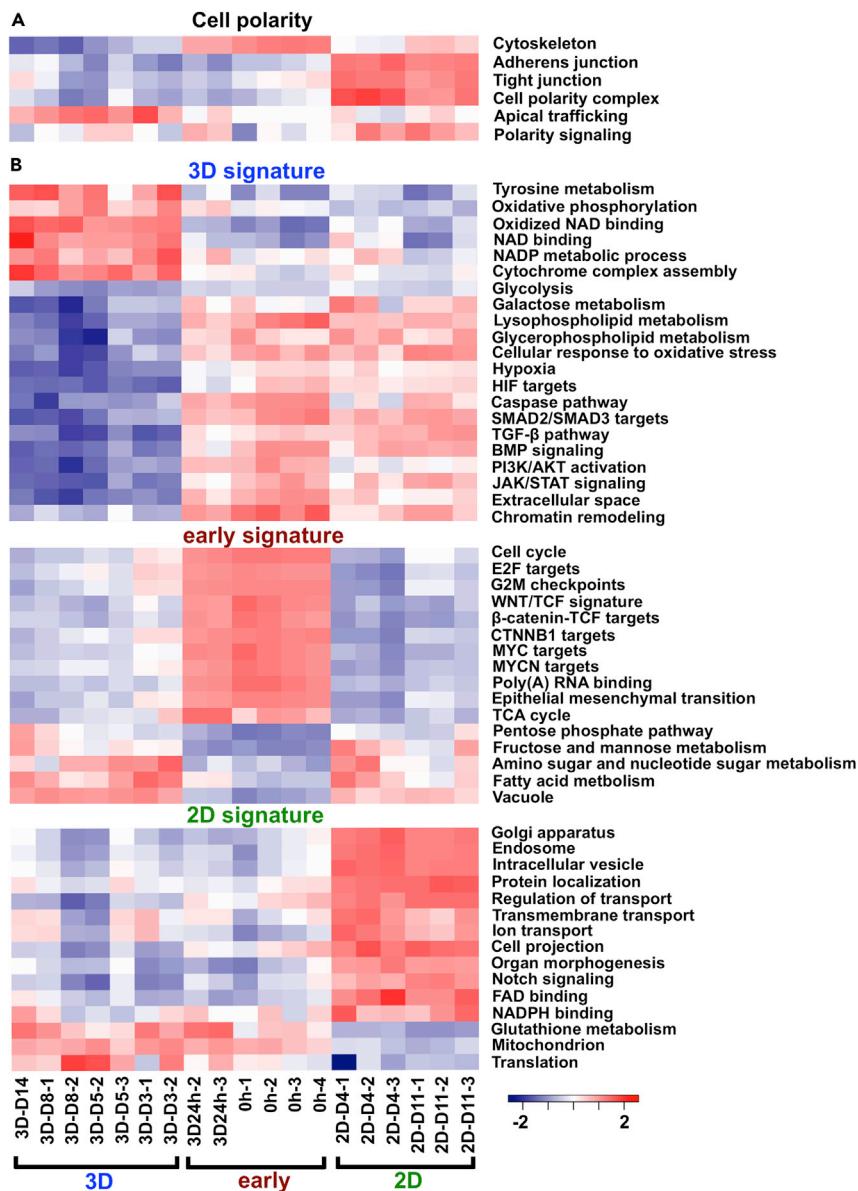
**Figure 3. Each of the Early, 2D, and 3D Sample Clusters Harbors Distinct Molecular Features**

(A) More functional groups of epithelial apico-basolateral cell polarity are up-regulated in 2D samples, compared with 3D samples and early samples. The heatmap indicates the ssGSEA scores (obtained with predefined gene sets), with up-regulation in red and down-regulation in blue. Only functional groups that significantly differ between any two sample clusters are shown, with p ≤ 0.05 by Wilcoxon tests.

(B) Each sample cluster harbors significantly up- or down-regulated functional groups that are related to epithelial morphogenesis, compared with the other two groups.

See Figure S3 and Table S3.

metabolism, redox regulation, and signaling (Table S3A). The analysis indicates that each cluster has distinct molecular features (Figure 3B). Early samples are marked by up-regulation of pathways and functions that promote cell cycle, as well as by down-regulation of the pentose phosphate pathway and the metabolism of sugars, amino acids, and fatty acids (Figure 3B). 2D samples are characterized by up-regulation of trafficking-related functions and down-regulation of mitochondria-associated genes (Figure 3B). 3D samples feature more down-regulated functions, including glycolysis, hypoxia, TGF-β signaling, and chromatin remodeling (Figure 3B). However, oxidative phosphorylation and redox-related functions are up-regulated in 3D samples.

## Cell Cycle Genes Are Transcriptionally Down-Regulated, whereas Genes of More Diverse Functions Are Up-Regulated during Epithelial Morphogenesis

To determine transcriptomic change-associated functions that contribute to epithelial morphogenesis, we identified DE genes between early and 3D or 2D sample groups, among 12,056 genes with FPKM >1 in at least one sample. Among ~2,000 DE genes found (Table S4A), the most prominent feature is down-regulation of cell cycle genes in both 3D and 2D samples, compared with early samples (Figure 4A). Specifically, ~63.5% of down-regulated genes overlap between 3D and 2D clusters (Figure 4B). Among them, ~50% are associated with cell cycle, many being targets of E2F and/or MYC (Figure 4B; Table S4B). These observations are consistent with the molecular characteristics of early samples (Figure 3B).

We identified ~1,000 up-regulated genes in each of the 3D and 2D groups (Figure 4A; Table S4A), among which 71% overlap (Figure 4B). These genes are, however, more diverse in their enriched functions. Among these functions are cell projection, trafficking, and others related to organ morphogenesis, none of which is, however, as prominent as down-regulation of cell cycle genes (Figure 4B).

We validated several up- or down-regulated genes at the protein level. These include *MYC* (Figure 5A), *VIM* (encoding the mesenchymal marker vimentin) (Figure S4E), and *DENND5A* (encoding a Rab11-interacting intracellular trafficking protein) (Figure S4F).

Lastly, down-regulated genes are more likely to be located in the nucleus and to be MYC targets, compared with up-regulated genes. However, we found no differences in the fraction of genes with miRNA or transcriptional factor-binding sites (Figure 4C; Table S4C).

Our analysis indicates that during MDCKII epithelial morphogenesis, cell cycle genes are down-regulated while trafficking-related genes are up-regulated (Figures 4A and 4B). To validate this finding, we investigated relevant microarray data published previously (Galvez-Santisteban et al., 2012; Wells et al., 2013; Chacon-Heszele et al., 2014). First, we compared up- or down-regulated genes during 3D or 2D epithelial morphogenesis from our study, indicated in Figure 4A, with DE genes identified between 36 h (Wells et al., 2013) and 8 day (Galvez-Santisteban et al., 2012) microarray studies. We found that the overlap is significant for down-regulated genes, with 22% and 56% genes shared for 3D and 2D, respectively (Figure S6A). Moreover, like our study (Figure 4A), down-regulation of cell cycle genes is also a prominent feature found for these microarray studies (Figure S6A). Up-regulated genes have much smaller overlaps (Figure S6A), but one consistency is up-regulation of trafficking-related genes in 2D condition (Figures 4A and S6A). Second, we compared our genes (Figure 4A) with the DE genes before and after HGF treatment from another microarray study (Chacon-Heszele et al., 2014) (Figure S6B). This is because HGF treatment has been shown to disrupt cell polarity (Chacon-Heszele et al., 2014; O'Brien et al., 2002), a reverse of epithelial morphogenesis of our study. Indeed, our results indicate corresponding gene expression changes, with 25%–33% of DE genes shared and enriched functions matched (Figure S6B). In summary, our findings are largely consistent with those obtained from microarray data.

## Transcriptomic Changes Are Related to Intracellular Trafficking

Our observed qualitative change in the transcriptome (Figure 2C), including down-regulation of cell cycle genes and up-regulation of trafficking-related genes (Figures 3 and 4), occurs in the same time frame (Figures 1 and 2) when a Rab11a-directed network regulates *de novo* MDCK lumenogenesis (Bryant et al., 2010). To test if our qualitative transcriptomic change is related to this intracellular trafficking, we investigated knockdown (KD) clones of *AVL9* and *DENND5A*, which encoded Rab11-interacting proteins and participate in intracellular trafficking in MDCKII cells (Li et al., 2014). Indeed, these KD clones display abnormal trafficking starting at the first cell cycle, as shown by diffused E-cadherin staining (Figure 4D), and accelerated cell proliferation after the first cell cycle when compared with the wild-type cells, resulting in abnormal lumenogenesis (Li et al., 2014). Thus, investigating the transcriptomic changes in these KD clones will be informative.

We performed RNA-seq analysis of two *AVL9*-KD clones and two *DENND5A*-KD clones, 10 samples total consisting of cells before seeding (0h) and over-confluent 2D cells (D4 and D7) that were cultured under the same condition as the wild-type cells (Figures S3 and S4B–S4D; Tables S1B and S3B). The results are consistent with their phenotype described above. First, a direct comparison between KD cells and wild-type cells indicates that, in KD cells, intracellular trafficking genes are significantly down-regulated, whereas cell
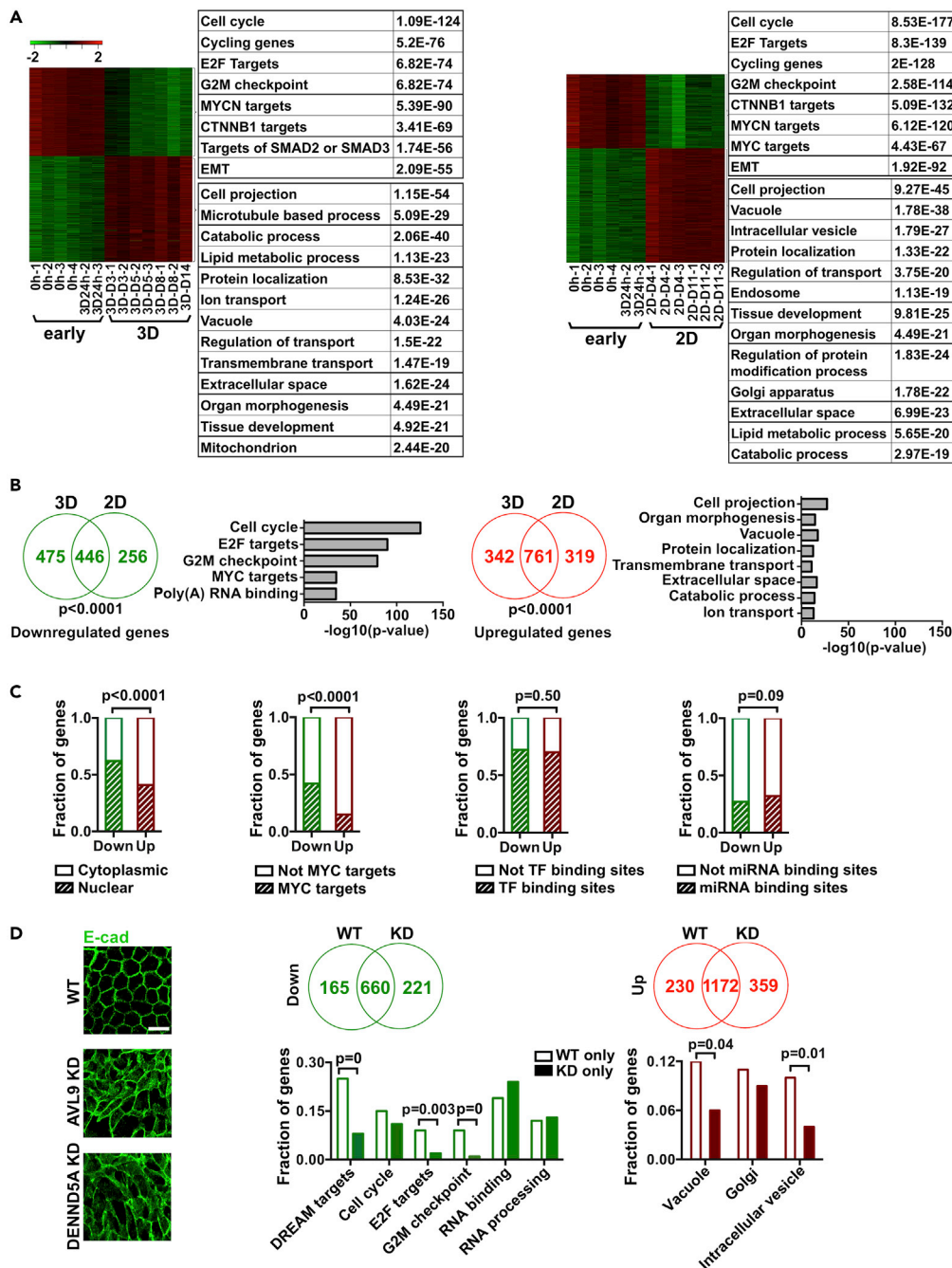
**Figure 4. Cell Cycle Genes are Transcriptionally Down-Regulated, whereas Genes of More Diverse Functions, Including Intracellular Trafficking, are Up-Regulated during Epithelial Morphogenesis**

(A) Heatmaps indicate the expression levels of up- and down-regulated genes (FDR <0.01 and >2-fold change), in red and green, respectively, between early and 3D (left) or 2D (right) sample groups. Their enriched functions and associated p values, identified with the GSEA web tool (see Transparent Methods), are also shown.

(B) Venn diagrams indicate significant overlap between 3D and 2D samples for down- (left) and up-regulated (right) genes shown in (A), as indicated by p values obtained via permutation tests. Enriched functions among overlapped genes and their $\log_{10}$ transformed p values are also shown, identified with the GSEA web tool.

(C) Down-regulated genes overlapping between 3D and 2D samples (the same genes as in B) are significantly more enriched in nuclear location and MYC targets but are similarly enriched in miRNA and transcriptional factor (TF) binding sites, compared with up-regulated genes. p Values were calculated with Fisher's exact tests.

**Figure 4. *Continued***

(D) During epithelial morphogenesis, wild-type (WT) cells have more cell cycle genes down-regulated and more intracellular trafficking genes up-regulated, compared with *AVL9*- and *DENND5A*-knockdown (KD) clones (Li et al., 2014). Confocal images (left) indicate abnormal intracellular trafficking of E-cadherin (E-cad) in KD cells. Scale bar, 10 μm. Venn diagrams (middle and right top) indicate the overlap of down- or up-regulated genes during epithelial morphogenesis between WT and KD cells (with *AVL9*-KD and *DENND5A*-KD samples combined). p Values were calculated with Fisher's exact tests. Plots below show enriched functions of down- or up-regulated genes unique to WT cells (165 and 230 genes) or to KD cells (221 and 359 genes).

See Figure S4 and Table S4.

proliferation genes are significantly up-regulated (Figures S3 and S4D). Second, we compared the up- and down-regulated genes identified during epithelial morphogenesis of each cell type (Figure 4D; Table S4D). Down-regulated genes unique to wild-type cells are at least three times more enriched in cell cycle-associated functions, particularly DREAM (dimerization partner, RB-like, E2F and multi-vulval class B) targets, E2F targets, and G2M check point (Figure 4D). Up-regulated genes unique to wild-type are more enriched in intracellular trafficking-relevant functions, including vacuoles and intracellular vesicles. These results support our hypothesis that the observed transcriptomic changes (Figure 4A) are related to intracellular trafficking.
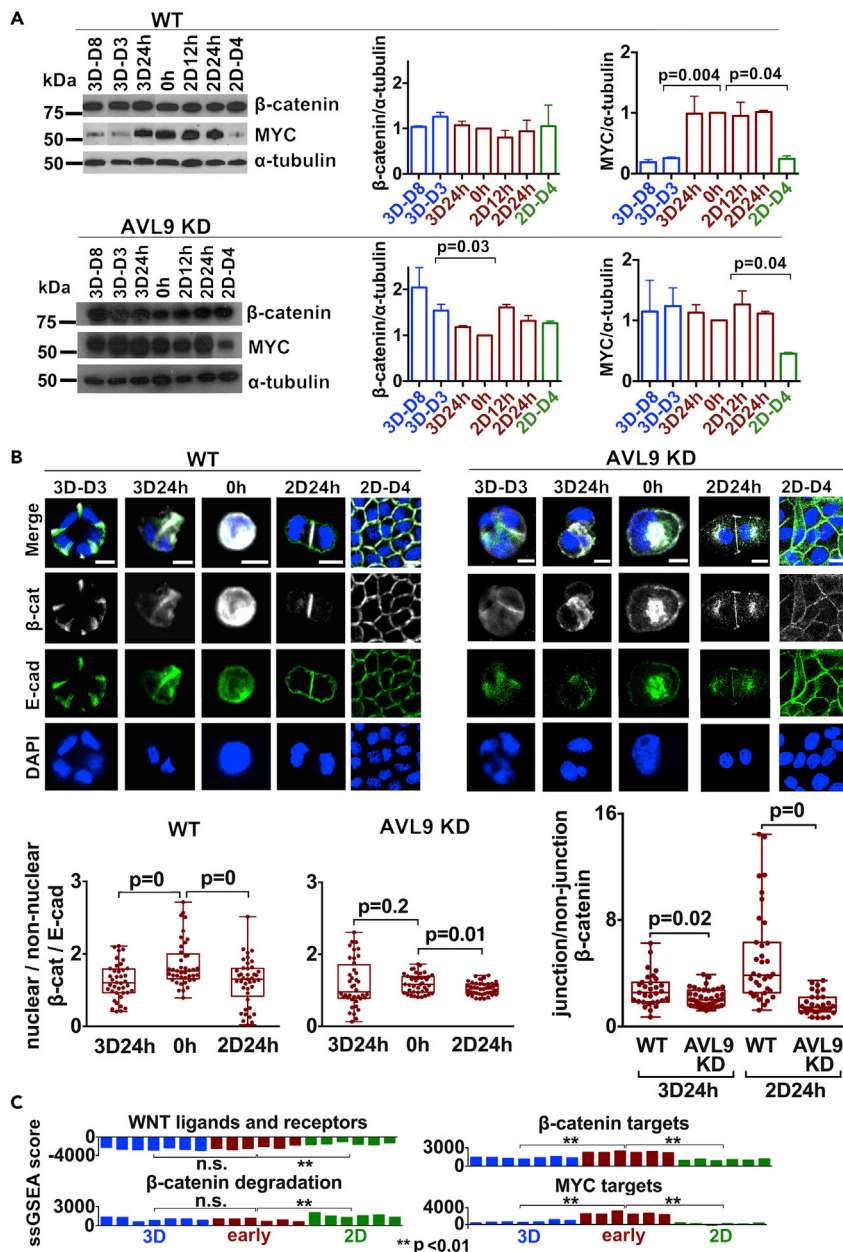
## β-Catenin Is Depleted in the Nucleus and Accumulated at the Cell-Cell Junctions upon the Completion of the First Cell Cycle

β-Catenin is a transducer of the WNT signaling pathway. When located in the nucleus, β-catenin is a transcription factor, regulating the expression of *MYC* and other target genes. When located in the cell-cell junctions, β-catenin is a component of adherens junctions, contributing to epithelial cell polarity establishment and maintenance (Baum and Georgiou, 2011). Importantly, our analysis reveals that, during MDCKII epithelial morphogenesis, β-catenin targets, including *MYC*, are down-regulated, whereas cell polarity genes, including those of adherens junctions, are up-regulated (Figures 3 and 4). Hence, we investigated β-catenin.

Resembling its transcript level change (Figure S5A), the amount of the MYC protein is significantly decreased during MDCKII epithelial morphogenesis (Figure 5A). Meanwhile, the amount of β-catenin remains relatively constant throughout the process (Figure 5A), despite differences in the levels of transcripts (Figure S5A). We did, however, observe a change in the cellular location of β-catenin. In both 3D and 2D culture, using E-cadherin as a reference (as it is not known to locate in the nucleus), our analysis indicates that β-catenin is located in the nucleus before the first cell division (Figure 5B) but depleted from the nucleus after the first cell cycle (Figure 5B). β-Catenin is mostly concentrated at the cell-cell junctions after the first cell division (Figure 5B). This differential distribution is statistically significant after examining >40 cells in each condition (Figure 5B).

To determine if intracellular trafficking affects β-catenin localization, we investigated an *AVL9*-KD clone of MDCKII, in which early intracellular trafficking and lumenogenesis are disrupted (Li et al., 2014). We noted several differences, compared with wild-type cells. First, in *AVL9*-KD cells, β-catenin is clearly visible near the nucleus and in the cytoplasm after the first cell division, with smaller amount of β-catenin accumulating at the cell-cell junctions (Figure 5B). Second, during 3D cystogenesis of *AVL9*-KD cells, the total amount of β-catenin increases (Figure 5A), which is accompanied by mostly constant levels of the *MYC* transcript (Figure S5A) and protein (Figure 5A). Indeed, MYC target genes are significantly up-regulated in *AVL9*-KD cells, compared with wild-type cells (Figure S5B). Similar observations were made for *DENND5A*-KD clones (Figure S5C).

In an attempt to assess if WNT signaling contributes to the observed depletion of nuclear β-catenin in normal morphogenesis (Figure 5B), we divided WNT signaling genes into three groups: (1) WNT ligands and receptors, (2) β-catenin degradation complex, and (3) β-catenin targets (Table S5). We performed ssGSEA of each group with RNA-seq data. Interestingly, we observed no significant changes for either WNT ligand and receptor genes or β-catenin degradation complex genes during 3D cystogenesis (Figure 5C). This differs from 2D epithelial morphogenesis, where WNT signaling appears to be up-regulated (Figure 5C). As expected, β-catenin targets (and MYC targets) are down-regulated for both 3D and 2D (Figure 5C). Hence, it is possible that no significant change occurs to WNT signaling during 3D cystogenesis.

**Figure 5. β-Catenin is Depleted in the Nucleus and Enriched at the Cell-Cell Junctions, whereas MYC is Down-Regulated, upon the Completion of the First Cell Cycle**

(A) In WT cells, β-catenin stays constant, whereas MYC is significantly reduced in both 3D and 2D samples, compared with the early group. In AVL9-KD cells, β-catenin increases significantly, whereas MYC stays constant during 3D cystogenesis. Bar plots quantify western blot experiments shown on the left, with p values of Wilcoxon tests indicated. Data are represented as mean ± SD with three biological replicates.

(B) Depletion of β-catenin in the nucleus and enrichment of β-catenin at the cell-cell junctions occur during the first cell division in WT cells. These changes are, however, less clear in *AVL9*-KD cells. Representative confocal images are shown at the top. Scale bar, 5 μm. Bar plots below indicate the relative amount of nuclear β-catenin normalized by E-cadherin, and of cell junction β-catenin normalized by non-junction β-catenin, quantified from at least 40 cells for each condition specified and with p values of Wilcoxon tests shown.

(C) ssGSEA scores (calculated with RNA-seq data) indicate no significant change for WNT ligands and receptors, and β-catenin degradation complex during 3D cystogenesis, unlike β-catenin targets and MYC targets. All four functional groups are significantly changed in 2D epithelial morphogenesis. p Values are from Wilcoxon tests. **, p < 0.01; n.s., p > 0.05. See also Figure S5 and Table S5.
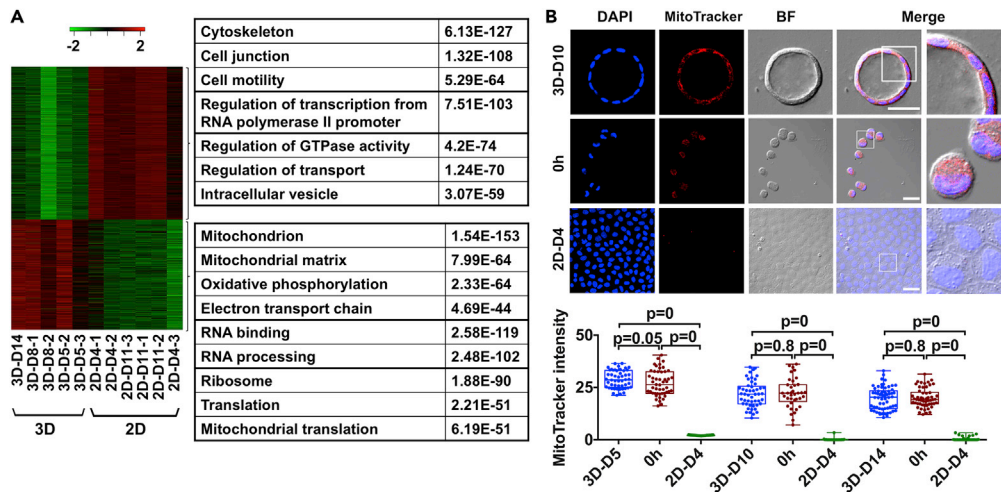
**Figure 6. Active Mitochondria are Maintained in 3D, but not 2D, Epithelial Morphogenesis**

(A) The heatmap indicates the expression levels of up- and down-regulated genes (FDR <0.01), in red and green, respectively, between 3D and 2D samples. Enriched functions of these genes, including those related to mitochondria, and their associated p values, obtained via GSEA, are also shown.

(B) Representative confocal images of MitoTracker staining, as an indicator of active mitochondria, in 3D, 2D, and 0h cells. Scale bar, 20 μm. Box plots at the bottom indicate the quantification of MitoTracker intensity in 3D, 2D, and 0h cells, collected from at least 40 cells for each cell type. p values were calculated using Wilcoxon tests.

See also Figure S6 and Table S6.

These observations support that intracellular trafficking may possibly play a role in depleting β-catenin from the nucleus and enriching β-catenin at the cell-cell junctions during early 3D epithelial morphogenesis. This in turn may down-regulate *MYC* transcription and the MYC network.

## Active Mitochondria Are Maintained in 3D, but Not 2D, Epithelial Morphogenesis

Besides shared features between 3D and 2D epithelial morphogenesis described above, we also investigated their differences. By identifying DE genes between 3D and 2D cells, we found that trafficking-related genes are up-regulated in 2D cells, whereas mitochondria-related genes are up-regulated in 3D cells (Figure 6A; Table S6). These are consistent with up-regulated genes unique to 3D or 2D epithelial morphogenesis (Figure S4A). To validate this finding, we quantified active mitochondria in 3D, 2D, and 0h cells via MitoTracker staining. We found that active mitochondria are maintained in 3D, but not 2D, epithelial morphogenesis (Figure 6B).

We also analyzed published MDCKII microarray data (Wells et al., 2013; Chacon-Heszele et al., 2014) for 2D and 3D cell difference. Surprisingly, the results are inconsistent, with no significant overlap in DE genes and no significant matches in enriched functions between our study and the microarray data (Figure S6C). This may be due to different 3D culture conditions, as type 1 fibrillar collagen gel was used in these microarray studies, unlike ours where Matrigel was used.

## Chromatin Becomes Less Accessible during 3D Epithelial Morphogenesis, Especially at Intergenic Regions

To test if the qualitative transcriptomic change (Figures 2C and 4A) is associated with chromatin remodeling, we performed ATAC-seq experiments with 3D, 2D, and 0h samples. We generated 120–200 million paired-end reads of 36 bp per sample and investigated the open chromatin with reads aligned to the nuclear genome (Table S7A).

Our analysis reveals that the chromatin becomes less accessible during 3D epithelial morphogenesis. First, 3D samples contain nearly 15,000 fewer open chromatin regions (OCRs), identified as ATAC-seq peaks, compared with 0h samples (Figure S7A; Table S7B). As a comparison, the corresponding number is only ~1,000 for 2D samples (Figure S7A). Supporting this, differential peak analysis reveals that OCRs have

shrunken or disappeared in 3D samples by ∼3-fold, compared with 0h samples (Figure 7A; Table S7B). The corresponding number is significantly smaller for 2D samples (Figure 7A; Table S7B). Moreover, 3D samples harbor more *de novo* changes (i.e., newly emerged or disappeared OCRs) than 2D samples (Figure 7A), indicating that 3D epithelial morphogenesis involves more extensive chromatin remodeling. This may be related to the lumen, which harbors a microenvironment that differs from the culture media and could restrict the cell's access to growth factors from the culture media.

To determine the functional implication of the OCRs, we examined their distribution in the genome. About 15%–16% of OCRs are located within the promoter regions (Figure S7B; Table S7B), defined as 2 kb upstream or downstream of each transcription start site (TSS) and with many validated via H3K4me3 enrichment analysis (Villar et al., 2015) (Figure S7B). The remaining OCRs are located in exons (15%–16%), introns (34%–35%), and intergenic regions (34%–36%) (Figure S7B; Table S7B), with the majority within 20 kb upstream or downstream of the TSS (Figure S7B). These numbers are similar to those reported in human ATAC-seq studies (Pastor et al., 2018; Corces et al., 2018; Ackermann et al., 2016).

We also noted that OCRs in 3D samples are more likely to reside in promoters and exonic regions and less likely to reside in intergenic regions, compared with OCRs in 0h samples (Figure S7B). No such difference was observed for 2D samples (Figure S7B). For OCRs that are changed during epithelial morphogenesis, 3D and 2D samples follow the same trend. Specifically, expanded or newly emerged OCRs tend to locate in promoter or exonic regions, whereas shrunken or disappeared OCRs tend to locate in introns and intergenic regions (Figures 7A and S8).

To determine if any OCRs are enhancers, we identified genomic regions that are enriched in H3K27 acetylation, an active enhancer marker (Villar et al., 2015). The putative enhancers discovered are mostly concentrated within 5 kb upstream or downstream of the TSS (Figure S7B), consistent with findings in human studies (Villar et al., 2015). We noted that, in all three groups of samples, a substantial fraction of OCRs, especially those located in exons and introns, are putative enhancers (Figure S7B; Table S7B). For OCRs that are changed during epithelial morphogenesis and within non-promoter regions, those expanded or newly emerged contain more putative enhancers, compared with those that are shrunken or disappeared, for both 3D and 2D culture (Figures 7A and S8).

### Transcriptomic Changes during MDCKII Epithelial Morphogenesis Are Associated with Chromatin Remodeling

Among up- or down-regulated genes during 3D or 2D epithelial morphogenesis, a significant portion of promoters contain OCRs that are expanded or shrunken accordingly (Figure 7B). This is exemplified by the following genes. *CDC6*, which functions in DNA replication and cell cycle checkpoints (Borlado and Mendez, 2008), is down-regulated and has promoter OCRs shrunken during both 3D and 2D epithelial morphogenesis (Figure 7C). The opposite was observed for SLC12A2 (Figure 7C), which encodes a basolateral $Na^+$-$K^+$-$2Cl^-$ cotransporter participating in cation reabsorption in kidney (Manohar and Leung, 2018). Other examples are provided in Figure S9A. These observations support that chromatin remodeling underlies the observed transcript abundance changes for a significant fraction of genes (Figures 2C, 4A, and 7B).

To identify what transcription factors are important for epithelial morphogenesis, we analyzed motif enrichment in changed OCRs and observed the same trend in both 3D and 2D cells. Specifically, among shrunken OCRs, the most enriched motifs are the binding sites for Fra1, Atf3, and other bZIP (basic leucine zipper) transcription factors (Table S7C), all of which have proliferative function. Among expanded OCRs, the HNF1B-binding motif is the most significantly enriched (Figures 7D and S9B; Table S7C). HNF1B is a transcription factor required for nephron development. Its binding motif is distributed in OCR-located promoters, exons, introns, and intergenic regions, with a substantial fraction being putative enhancers (Figure S9B). Note that HNF1A/B-binding sites are also among the top-most enriched motifs amid up-regulated genes during 3D and 2D epithelial morphogenesis (Figures 7D and S9B; Table S7C). Importantly, a significant overlap was observed between genes with either type of enrichment (Figures 7D and S9B). *HNF1B* expression also increases during the course of epithelial morphogenesis (Figures 7D and S9B). These results indicate that HNF1B is likely a key transcription factor for up-regulated genes during MDCKII epithelial morphogenesis, which is consistent with the importance of mitotic bookmarking by HNF1B in kidney development (Lerner et al., 2016; Verdeguer et al., 2010).
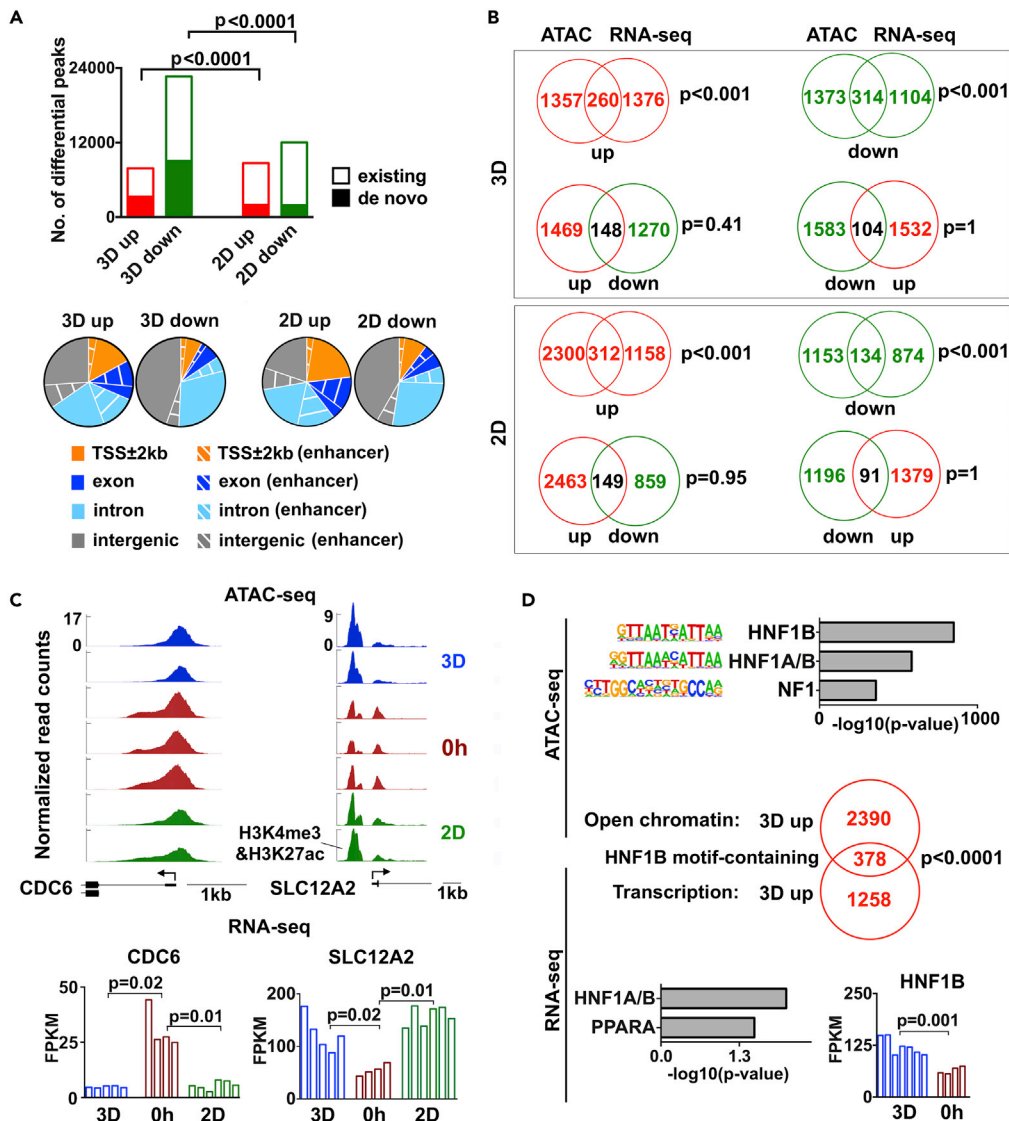
**Figure 7. Transcriptomic Changes during MDCKII Epithelial Morphogenesis are Associated with Chromatin Remodeling**

(A) Decrease in chromatin accessibility is more extensive during 3D epithelial morphogenesis, compared with 2D. The bar plot (top) indicates increased/decreased (existing) or emerged/disappeared (*de novo*) ATAC-seq peaks in 3D and 2D samples, when compared with 0h samples. p Values were calculated with Fisher's exact tests. Pie charts below indicate that increased or emerged open chromatin regions in 3D and 2D samples are likely to be located in promoter and exonic regions, compared with those that decreased or disappeared. Regions with dashed lines are putative enhancers identified by H3K27ac enrichment.

(B) Significant fractions of down- and up-regulated genes during 3D and 2D epithelial morphogenesis have undergone chromatin remodeling. Venn diagrams show the overlaps between differential expressed genes and genes with differential ATAC-seq peaks at their promoter regions. p Values were calculated with permutation tests.

(C) Two genes exemplify the corresponding changes between gene expression and promoter chromatin accessibility. Of note is that CDC6 is a target of MYC, whereas SLC12A2 is a target of HNF1B.

(D) HNF1B-binding sites are among the topmost enriched motifs in increased ATAC-seq peaks and the promoters of up-regulated genes during 3D epithelial morphogenesis. The analysis was performed using HOMER and Enrichr. Venn diagram indicates significant overlap between up-regulated genes and genes that harbor expanded/emerged and HNF1B motif-containing ATAC-seq peaks in their gene body and 20-kb flanking regions. p Values were calculated with permutation tests. Bar plot shows up-regulation of HNF1B.

See also Figures S7–S10 and Table S7.

**DISCUSSION**

**A Qualitative Change in the Transcriptome Is Linked to the First Cell Cycle and Lumen Establishment**

MDCKII is widely used to study epithelial morphogenesis and cell polarity establishment. Thus, it is important to understand transcriptomic changes during these processes. Our time course analysis of MDCKII 3D cystogenesis has revealed for the first time a single qualitative transcriptomic change, which consists of >1,000 up- or down-regulated genes each. The change occurs during the 24 h to day 3 transition time of MDCKII cystogenesis, coinciding with cell polarity establishment and lumen formation.

3D cystogenesis undergoes lumen-establishing, lumen-enlarging, and lumen-maintenance stages (Li et al., 2014). Our study reveals that the transcriptome switches during the lumen-establishing stage but stays largely the same during the other two stages. However, a significant fraction of the changes that contribute to this transcriptomic switch also occurs in 2D culture. As 2D culture lacks lumenogenesis, the shared changes, which include down-regulation of cell cycle genes and up-regulation of trafficking genes, are most likely related to cell polarity building. Note that the transcriptomic switch also consists of changes unique to 3D, as the transcriptome significantly differs between polarized 2D and 3D cells (which will be discussed later).

We propose that β-catenin and HNF1B may contribute to this qualitative transcriptomic change (Figure S10). During MDCK lumenogenesis, β-catenin is likely transported to the newly formed cell-cell junctions via a Rab11a-directed network (Bryant et al., 2010; Lock and Stow, 2005; Chen et al., 1999). Our data indicate the possibility that translocation of β-catenin from the nucleus to the cell-cell junctions takes place during the first cell cycle of MDCK epithelial morphogenesis. This down-regulates *MYC* transcription and in turn the MYC network that consists of many cell cycle genes (Valenta et al., 2012), thereby decreasing cell proliferation. HNF1B is a transcription factor required for kidney development and its target genes include those functioning in epithelial cell polarity building. Previous studies have shown mitotic bookmarking of HNF1B in MDCKII cells (Lerner et al., 2016; Verdeguer et al., 2010). We hypothesize that, because of HNF1B mitotic bookmarking and nuclear β-catenin depletion (Chan et al., 2019), the HNF1B-regulated transcriptional network is immediately activated upon the first cell cycle completion, leading to the establishment of cell polarity and the lumen. Of course, our model (Figure S10) needs experimental validation.

The qualitative change reported here may have implications for *in vivo* tissue development. For example, colonic crypts develop via two major differentiation steps: (1) colonic stem cells to transit-amplifying (TA) cells and (2) TA cells to terminally differentiated (TD) cells (Sheaffer and Kaestner, 2012; Snippert et al., 2010). The second differentiation (TA cells to TD cells) resembles our MDCKII epithelial morphogenesis, with cell proliferation decreasing and cell polarity being established. Furthermore, similar signatures, such as MYC network down-regulation, are also reported in this differentiation (Merlos-Suarez et al., 2011; Kosinski et al., 2007; Wang et al., 2018a, 2018b). However, two questions remain. First, is the transcriptomic change during the differentiation qualitative, as in our MDCKII system, or quantitative, i.e., gradual change? Second, more β-catenin molecules accumulate at the adherens junctions of TD cells, which have well-established apico-basolateral cell polarity, unlike TA cells. Is translocation of β-catenin also possibly responsible for the nuclear β-catenin depletion, besides WNT signaling regulation (Schepers and Clevers, 2012; Di Cecilia et al., 2016)?

Addressing these questions is also important in disease development. Loss of cell polarity is a driver of carcinogenesis of epithelial tissues including colon and breast (Tang et al., 2014; Li et al., 2014; Royer and Lu, 2011). Our MDCKII 3D epithelial morphogenesis is the reverse of epithelial carcinogenesis, i.e., a tumor suppression process. Enhancing the trafficking of β-catenin from the nucleus to the adherens junction and maintaining β-catenin at the junction will result in tumor suppression.

**3D Culture Is a Better Model Than 2D Culture in Studying Mitochondria in Kidney**

Our study indicates that mitochondria-related genes, e.g., those functioning in oxidative phosphorylation, are up-regulated in 3D cells, compared with 2D cells. Our study also reveals that active mitochondria are maintained in 3D, but not 2D, epithelial morphogenesis. The kidney has high mitochondrial abundance to support ATP generation via oxidative phosphorylation (Pagliarini et al., 2008). Particularly, proximal tubule cells, one of the major cell types in MDCKII cells, contain the most mitochondria and use the majority of

oxygen in the kidney to support physiological functions, including reabsorption of ions and glucose (Forbes, 2016). Our work indicates that 3D culture better resembles physiological conditions and therefore serves as a better cell model to study kidney biology, especially mitochondria and metabolism. We thus envision a wider use of MDCKII 3D culture in related studies.

### Limitations of the Study

Our study reveals that, during MDCKII epithelial morphogenesis, numerous genes are significantly changed at the transcript level, including down-regulation of many cell cycle genes (Figures 4A and 4B). However, we have not validated these changes at the protein level. Moreover, our study indicates that down-regulation of *MYC* and cell cycle genes may be caused by the translocation of β-catenin from the nucleus to the cell-cell junctions during the first cell cycle of MDCKII epithelial morphogenesis (Figures 5 and S10). However, we lack experimental data to support this. Future studies, such as those that disrupt β-catenin translocation, are needed to test this hypothesis. Our study, in combination with findings by others (Lerner et al., 2016; Verdeguer et al., 2010), indicates that mitotic bookmarking of HNF1B during the first cell division of MDCKII epithelial morphogenesis may lead to up-regulation of HNF1B target genes (Figures 7 and S9). However, experimental data are lacking to test this hypothesis. Functional studies, including HNF1B knockdown or knockout, may illustrate the mechanism through which HNF1B contributes to MDCKII epithelial morphogenesis. Our study also reveals that active mitochondria are maintained in 3D culture but not in 2D culture (Figure 6). However, the mechanism of this finding remains unknown, and future studies are required.

### Resource Availability

#### Lead Contact

Further information and requests for resources and reagents should be directed to and will be fulfilled by the Lead Contact, Shaying Zhao (szhao@uga.edu).

#### Materials Availability

This study did not generate new unique reagents.

#### Data and Code Availability

The accession number for the ATAC-seq and RNA-seq data reported in this paper is GEO: GSE129824. Custom scripts ATAC-seq and RNA-seq data analysis are available from the Lead Contact upon request.

### METHODS

All methods can be found in the accompanying Transparent Methods supplemental file.

### SUPPLEMENTAL INFORMATION

Supplemental Information can be found online at https://doi.org/10.1016/j.isci.2020.101629.

### AUTHOR CONTRIBUTIONS

Conceptualization, S.Z. and K.E.M.; Methodology, T.W., Z.L., S.Z., P.X., S.-H.K., and S.U.; Investigation, T.W. and S.Z.; Writing, S.Z., T.W., K.E.M., and R.J.S.; Visualization, T.W.; Resources, R.J.S. and S.D.; Supervision, S.Z. and K.E.M.; Funding Acquisition, S.Z., K.E.M., and S.-H.K.

## DECLARATION OF INTERESTS

The authors declare no competing interests.

## REFERENCES

Ackermann, A.M., Wang, Z., Schug, J., Naji, A., and Kaestner, K.H. (2016). Integration of ATAC-seq and RNA-seq identifies human alpha cell and beta cell signature genes. Mol. Metab. 5, 233–244.

Anders, S., and Huber, W. (2010). Differential expression analysis for sequence count data. Genome Biol. 11, R106.

Assemat, E., Bazellieres, E., Pallesi-Pocachard, E., Le Bivic, A., and Massey-Harroche, D. (2008). Polarity complex proteins. Biochim. Biophys. Acta 1778, 614–630.

Balkovetz, D.F., Gerrard, E.R., Li, S., Johnson, D., Lee, J., Tobias, J.W., Rogers, K.K., Snyder, R.W., and Lipschutz, J.H. (2004). Gene expression alterations during HGF-induced dedifferentiation of a renal tubular epithelial cell line (MDCK) using a novel canine DNA microarray. Am. J. Physiol. Ren. Physiol. 286, F702–F710.

Baum, B., and Georgiou, M. (2011). Dynamics of adherens junctions in epithelial establishment, maintenance, and remodeling. J. Cell Biol. 192, 907–917.

Benita, Y., Kikuchi, H., Smith, A.D., Zhang, M.Q., Chung, D.C., and Xavier, R.J. (2009). An integrative genomics approach identifies Hypoxia Inducible Factor-1 (HIF-1)-target genes that form the core response to hypoxia. Nucleic Acids Res. 37, 4587–4602.

Borlado, L.R., and Mendez, J. (2008). CDC6: from DNA replication to cell cycle checkpoints and oncogenesis. Carcinogenesis 29, 237–243.

Brunet, J.P., Tamayo, P., Golub, T.R., and Mesirov, J.P. (2004). Metagenes and molecular pattern discovery using matrix factorization. Proc. Natl. Acad. Sci. U S A 101, 4164–4169.

Bryant, D.M., Datta, A., Rodriguez-Fraticelli, A.E., Peranen, J., Martin-Belmonte, F., and Mostov, K.E. (2010). A molecular network for de novo generation of the apical surface and lumen. Nat. Cell Biol. 12, 1035–1045.

Bryant, D.M., and Mostov, K.E. (2008). From cells to organs: building polarized tissue. Nat. Rev. Mol. Cell Biol. 9, 887–901.

Cerruti, B., Puliafito, A., Shewan, A.M., Yu, W., Combes, A.N., Little, M.H., Chianale, F., Primo, L., Serini, G., Mostov, K.E., et al. (2013). Polarity, cell division, and out-of-equilibrium dynamics control the growth of epithelial structures. J. Cell Biol. 203, 359–372.

Chacon-Heszele, M.F., Zuo, X., Hellman, N.E., Mckenna, S., Choi, S.Y., Huang, L., Tobias, J.W., Park, K.M., and Lipschutz, J.H. (2014). Novel MAPK-dependent and -independent tubulogenes identified via microarray analysis of 3D-cultured Madin-Darby canine kidney cells. Am. J. Physiol. Ren. Physiol. 306, F1047–F1058.

Chan, S.C., Zhang, Y., Pontoglio, M., and Igarashi, P. (2019). Hepatocyte nuclear factor-1 beta regulates Wnt signaling through genome-wide competition with beta-catenin/lymphoid enhancer binding factor. Proc. Natl. Acad. Sci. U S A 116, 24133–24142.

Chen, Y.T., Stewart, D.B., and Nelson, W.J. (1999). Coupling assembly of the E-cadherin/beta-catenin complex to efficient endoplasmic reticulum exit and basal-lateral membrane targeting of E-cadherin in polarized MDCK cells. J. Cell Biol. 144, 687–699.

Corces, M.R., Granja, J.M., Shams, S., Louie, B.H., Seoane, J.A., Zhou, W., Silva, T.C., Groeneveld, C., Wong, C.K., Cho, S.W., et al. (2018). The chromatin accessibility landscape of primary human cancers. Science 362, eaav1898.

Datta, A., Bryant, D.M., and Mostov, K.E. (2011). Molecular regulation of lumen morphogenesis. Curr. Biol. 21, R126–R136.

Di Cecilia, S., Zhang, F., Sancho, A., Li, S.D., Aguilo, F., Sun, Y.F., Rengasamy, M., Zhang, W.J., Del Vecchio, L., Salvatore, F., and Walsh, M.J. (2016). RBM5-AS1 is critical for self-renewal of colon cancer stem-like cells. Cancer Res. 76, 5615–5627.

Dukes, J.D., Whitley, P., and Chalmers, A.D. (2011). The MDCK variety pack: choosing the right strain. BMC Cell Biol. 12, 43.

Elia, N., and Lippincott-Schwartz, J. (2009). Culturing MDCK cells in three dimensions for analyzing intracellular dynamics. Curr. Protoc. Cell Biol. 4, 4 22.

Fevr, T., Robine, S., Louvard, D., and Huelsken, J. (2007). Wnt/beta-catenin is essential for intestinal homeostasis and maintenance of intestinal stem cells. Mol. Cell. Biol. 27, 7551–7559.

Forbes, J.M. (2016). Mitochondria-power players in kidney function? Trends Endocrinol. Metab. 27, 441–442.

Galvez-Santisteban, M., Rodriguez-Fraticelli, A.E., Bryant, D.M., Vergarajauregui, S., Yasuda, T., Banon-Rodriguez, I., Bernascone, I., Datta, A., Spivak, N., Young, K., et al. (2012). Synaptotagmin-like proteins control the formation of a single apical membrane domain in epithelial cells. Nat. Cell Biol. 14, 838–849.

Koinuma, D., Tsutsumi, S., Kamimura, N., Taniguchi, H., Miyazawa, K., Sunamura, M., Imamura, T., Miyazono, K., and Aburatani, H. (2009). Chromatin immunoprecipitation on microarray analysis of Smad2/3 binding sites reveals roles of ETS1 and TFAP2A in transforming growth factor beta signaling. Mol. Cell. Biol. 29, 172–186.

Kosinski, C., Li, V.S.W., Chan, A.S.Y., Zhang, J., Ho, C., Tsui, W.Y., Chan, T.L., Mifflin, R.C., Powell, D.W., Yuen, S.T., et al. (2007). Gene expression patterns of human colon tops and basal crypts and BMP antagonists as intestinal stem cell niche factors. Proc. Natl. Acad. Sci. U S A 104, 15418–15423.

Kwon, S.-H., Nedvetsky, P.I., and Mostov, K.E. (2011). Transcriptional profiling identifies TNS4 function in epithelial tubulogenesis. Curr. Biol. 21, 161–166.

Lerner, J., Bagattin, A., Verdeguer, F., Makinistoglu, M.P., Garbay, S., Felix, T., Heidet, L., and Pontoglio, M. (2016). Human mutations affect the epigenetic/bookmarking function of HNF1B. Nucleic Acids Res. 44, 8097–8111.

Li, Y., Xu, J., Xiong, H., Ma, Z., Wang, Z., Kipreos, E.T., Dalton, S., and Zhao, S. (2014). Cancer driver candidate genes AVL9, DENND5A and NUPL1 contribute to MDCK cystogenesis. Oncoscience 1, 854–865.

Lock, J.G., and Stow, J.L. (2005). Rab11 in recycling endosomes regulates the sorting and basolateral transport of E-cadherin. Mol. Biol. Cell 16, 1744–1755.

Manohar, S., and Leung, N. (2018). Cisplatin nephrotoxicity: a review of the literature. J. Nephrol. 31, 15–25.

Martin-Belmonte, F., Gassama, A., Datta, A., Yu, W., Rescher, U., Gerke, V., and Mostov, K. (2007). PTEN-mediated apical segregation of phosphoinositides controls epithelial morphogenesis through Cdc42. Cell 128, 383–397.

McCaffrey, L.M., and Macara, I.G. (2012). Signaling pathways in cell polarity. Cold Spring Harb. Perspect. Biol. 4, a009654.

McRae, R., Lapierre, L.A., Manning, E.H., and Goldenring, J.R. (2017). Rab11-FIP1 phosphorylation by MARK2 regulates polarity in MDCK cells. Cell. Logist. 7, e1271498.

Merlos-Suarez, A., Barriga, F.M., Jung, P., Iglesias, M., Cespedes, M.V., Rossell, D., Sevillano, M., Hernando-Momblona, X., Da Silva-Diz, V., Munoz, P., et al. (2011). The intestinal stem cell signature identifies colorectal cancer stem cells and predicts disease relapse. Cell Stem Cell 8, 511–524.

Mjelle, R., Hegre, S.A., Aas, P.A., Slupphaug, G., Drablos, F., Saetrom, P., and Krokan, H.E. (2015).

Cell cycle regulation of human DNA repair and chromatin remodeling genes. DNA Repair (Amst) *30*, 53–67.

Muthuswamy, S.K., Li, D., Lelievre, S., Bissell, M.J., and Brugge, J.S. (2001). ErbB2, but not ErbB1, reinitiates proliferation and induces luminal repopulation in epithelial acini. Nat. Cell Biol. *3*, 785–792.

O'brien, L.E., Tang, K., Kats, E.S., Schutz-Geschwender, A., Lipschutz, J.H., and Mostov, K.E. (2004). ERK and MMPs sequentially regulate distinct stages of epithelial tubule development. Dev. Cell *7*, 21–32.

O'brien, L.E., Zegers, M.M.P., and Mostov, K.E. (2002). Building epithelial architecture: insights from three-dimensional culture models. Nat. Rev. Mol. Cell Biol. *3*, 531–537.

Pagliarini, D.J., Calvo, S.E., Chang, B., Sheth, S.A., Vafai, S.B., Ong, S.-E., Walford, G.A., Sugiana, C., Boneh, A., Chen, W.K., et al. (2008). A mitochondrial protein compendium elucidates complex I disease biology. Cell *134*, 112–123.

Pastor, W.A., Liu, W., Chen, D., Ho, J., Kim, R., Hunt, T.J., Lukianchikov, A., Liu, X., Polo, J.M., Jacobsen, S.E., and Clark, A.T. (2018). TFAP2C regulates transcription in human naive pluripotency by opening enhancers. Nat. Cell Biol. *20*, 553–564.

Pillman, K.A., Phillips, C.A., Roslan, S., Toubia, J., Dredge, B.K., Bert, A.G., Lumb, R., Neumann, D.P., Li, X., Conn, S.J., et al. (2018). miR-200/375 control epithelial plasticity-associated alternative splicing by repressing the RNA-binding protein Quaking. EMBO J. *37*, e99016.

Rodriguez-Fraticelli, A.E., Vergarajauregui, S., Eastburn, D.J., Datta, A., Alonso, M.A., Mostov, K., and Martin-Belmonte, F. (2010). The Cdc42

GEF Intersectin 2 controls mitotic spindle orientation to form the lumen during epithelial morphogenesis. J. Cell Biol. *189*, 725–738.

Royer, C., and Lu, X. (2011). Epithelial cell polarity: a major gatekeeper against cancer? Cell Death Differ. *18*, 1470–1477.

Schepers, A., and Clevers, H. (2012). Wnt signaling, stem cells, and cancer of the gastrointestinal tract. Cold Spring Harb. Perspect. Biol. *4*, a007989.

Shamir, E.R., and Ewald, A.J. (2014). Three-dimensional organotypic culture: experimental models of mammalian biology and disease. Nat. Rev. Mol. Cell Biol. *15*, 647–664.

Sheaffer, K.L., and Kaestner, K.H. (2012). Transcriptional networks in liver and intestinal development. Cold Spring Harb Perspect. Biol. *4*, a008284.

Snippert, H.J., Van Der Flier, L.G., Sato, T., Van Es, J.H., Van Den Born, M., Kroon-Veenboer, C., Barker, N., Klein, A.M., Van Rheenen, J., Simons, B.D., and Clevers, H. (2010). Intestinal crypt homeostasis results from neutral competition between symmetrically dividing Lgr5 stem cells. Cell *143*, 134–144.

Sottocornola, R., Royer, C., Vives, V., Tordella, L., Zhong, S., Wang, Y., Ratnayaka, I., Shipman, M., Cheung, A., Gaston-Massuet, C., et al. (2010). ASPP2 binds Par-3 and controls the polarity and proliferation of neural progenitors during CNS development. Dev. Cell *19*, 126–137.

Tang, J., Li, Y., Lyon, K., Camps, J., Dalton, S., Ried, T., and Zhao, S. (2014). Cancer driver-passenger distinction via sporadic human and dog cancer comparison: a proof-of-principle study with colorectal cancer. Oncogene *33*, 814–822.

Valenta, T., Hausmann, G., and Basler, K. (2012). The many faces and functions of β-catenin. EMBO J. *31*, 2714–2736.

van Der Flier, L.G., Sabates-Bellver, J., Oving, I., Haegebarth, A., De Palo, M., Anti, M., Van Gijn, M.E., Suijkerbuijk, S., Van De Wetering, M., Marra, G., and Clevers, H. (2007). The intestinal Wnt/TCF signature. Gastroenterology *132*, 628–632.

Verdeguer, F., Le Corre, S., Fischer, E., Callens, C., Garbay, S., Doyen, A., Igarashi, P., Terzi, F., and Pontoglio, M. (2010). A mitotic transcriptional switch in polycystic kidney disease. Nat. Med. *16*, 106–110.

Villar, D., Berthelot, C., Aldridge, S., Rayner, T.F., Lukk, M., Pignatelli, M., Park, T.J., Deaville, R., Erichsen, J.T., Jasinska, A.J., et al. (2015). Enhancer evolution across 20 mammalian species. Cell *160*, 554–566.

Wang, J., Wang, T., Bishop, M.A., Edwards, J.F., Yin, H., Dalton, S., Bryan, L.K., and Zhao, S. (2018a). Collaborating genomic, transcriptomic and microbiomic alterations lead to canine extreme intestinal polyposis. Oncotarget *9*, 29162–29179.

Wang, J., Wang, T., Sun, Y., Feng, Y., Kisseberth, W.C., Henry, C.J., Mok, I., Lana, S.E., Dobbin, K., Northrup, N., et al. (2018b). Proliferative and invasive colorectal tumors in pet dogs provide unique insights into human colorectal cancer. Cancers (Basel) *10*, 330.

Wells, E.K., Yarborough, O., 3rd, Lifton, R.P., Cantley, L.G., and Caplan, M.J. (2013). Epithelial morphogenesis of MDCK cells in three-dimensional collagen culture is modulated by interleukin-8. Am. J. Physiol. Cell Physiol. *304*, C966–C975.

**Supplemental Information**

**A Qualitative Change in the Transcriptome Occurs**

**after the First Cell Cycle and Coincides**

**with Lumen Establishment during MDCKII Cystogenesis**

Tianfang Wang, Sang-Ho Kwon, Xiao Peng, Severine Urdy, Zefu Lu, Robert J. Schmitz, Stephen Dalton, Keith E. Mostov, and Shaying Zhao

## Supplemental Figures and Legends

**A**



**B**
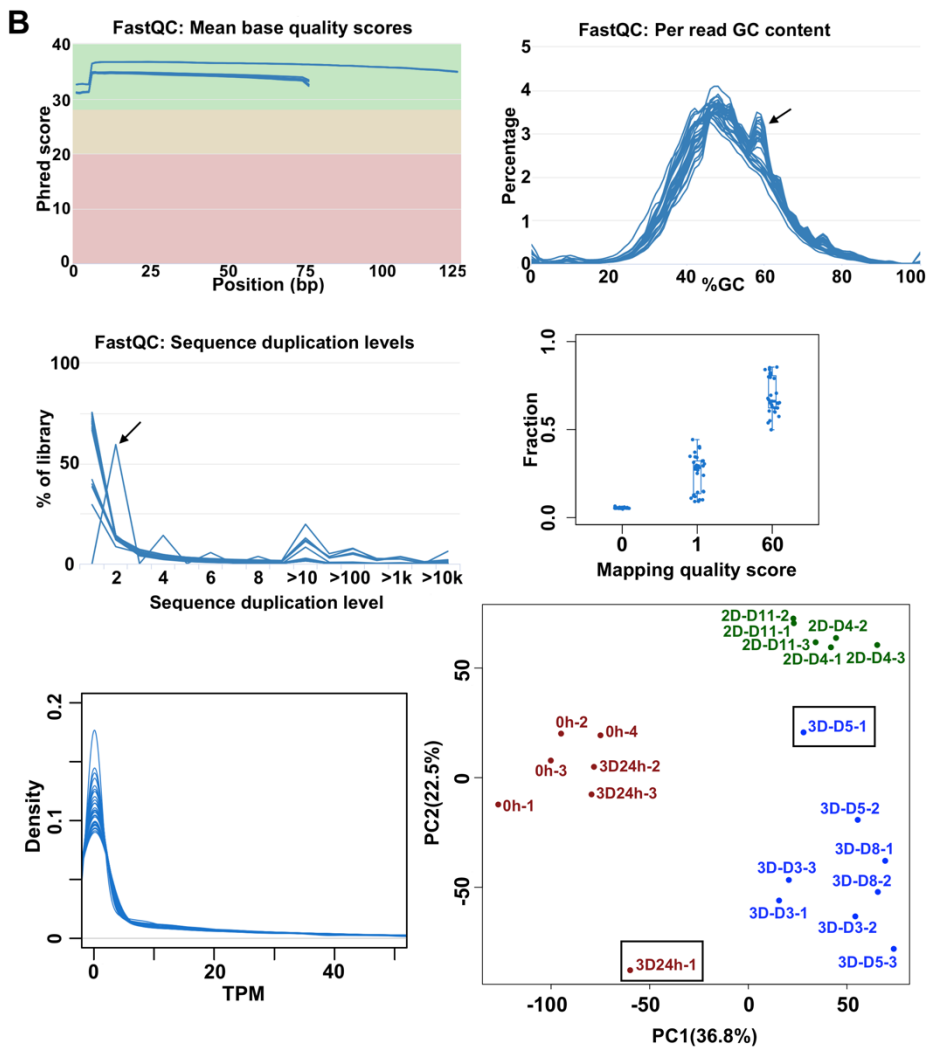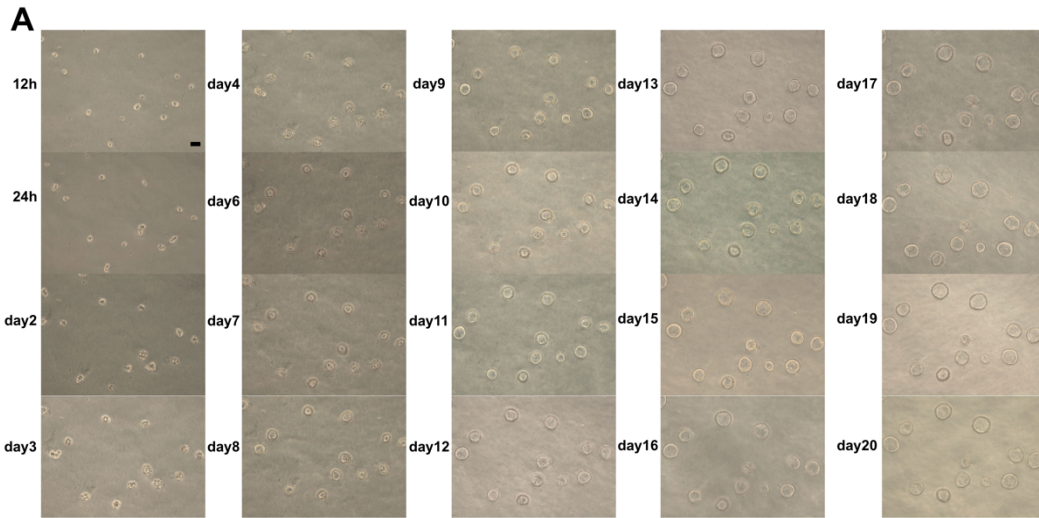
**Figure S1. Time-course RNA-seq experiments. Related to Figure 1**.

(A)      Bright-field images of time course MDCKII 3D epithelial morphogenesis.  Scale bar, 20 µm.

(B)      Quality control for MDCKII wild type (WT) and knockdown (KD) RNA-seq samples.  Top: mean base quality scores at each base position and distributions of per read GC content for 31 samples (21 WT and 10 KD) determined by FastQC.  Arrow indicates a peak corresponding to ~60% GC content for samples sequenced in one batch.  Middle: sequence duplication levels determined by FastQC, and distribution of mapping quality score of all the reads of each sample.  Arrow indicates the *AVL9*-KD 0h sample that was sequenced twice and the data was pooled.  Bottom: distributions of TPM values of 31 samples, and the PCA plot of 21 WT samples.  Highlighted samples are distant from other samples of the same group (outliers) and are thus excluded from downstream analysis.
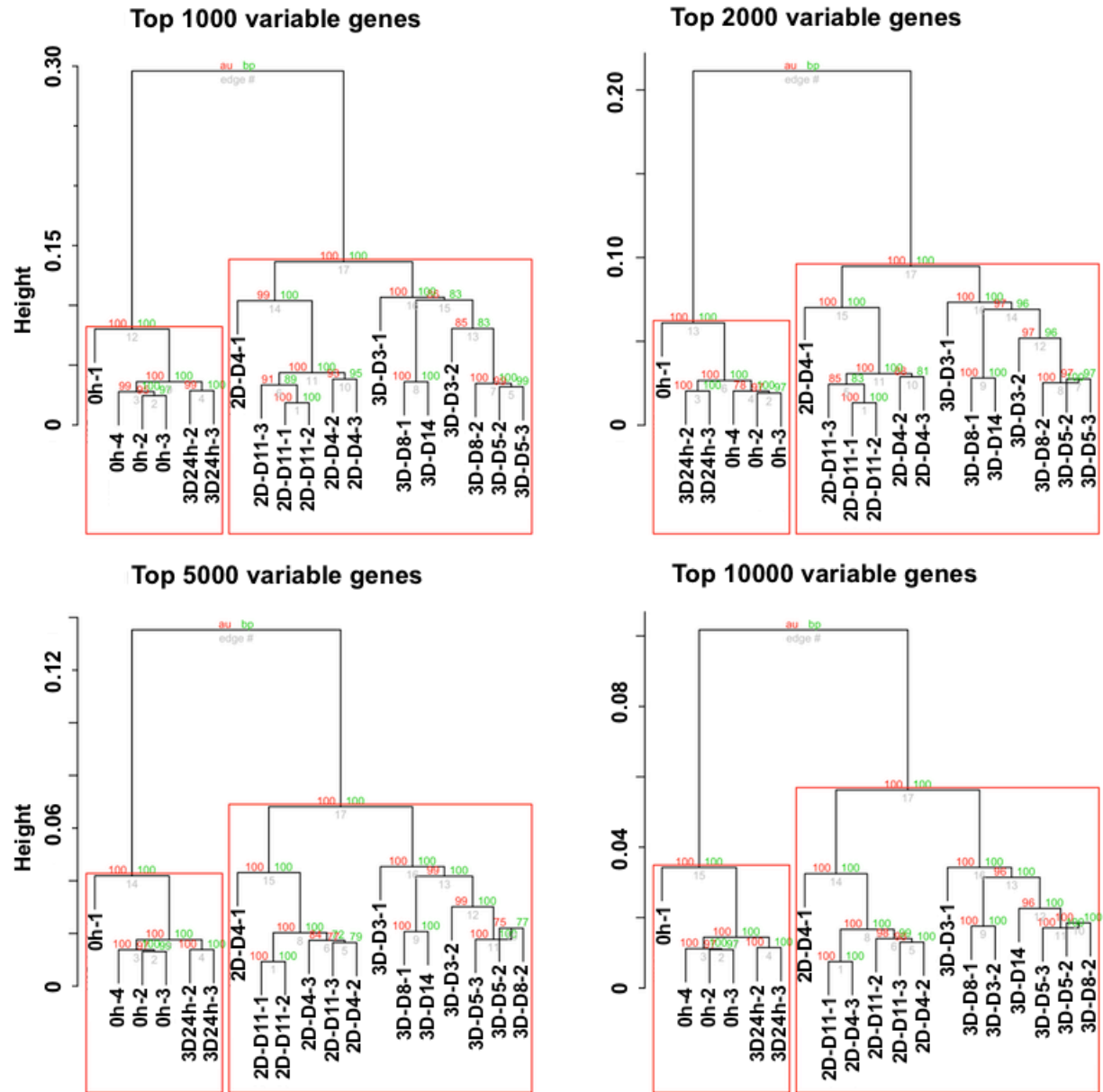
**Figure S2. Hierarchical clustering with top variably expressed genes across samples supports three sample clusters (3D, early and 2D). Related to Figure 2.**
Bootstrapping was performed with permutations (see Transparent Methods). Approximately unbiased (au) and bootstrap probability (bp) values are shown in red and green, respectively.

Heatmap columns (0h): 0h-1 0h, 0h-2 0h, 0h-3 0h, 0h-4 0h | AVL9-1 0h, AVL9-4 0h, DENND5A-1 0h, DENND5A-2 0h — WT | KD

Heatmap columns (2D): 2D-D4-1, 2D-D4-2, 2D-D4-3, 2D-D11-1, 2D-D11-2, 2D-D11-3 | AVL9-1 2D-D4, AVL9-4 2D-D4, DENND5A-1 2D-D4, DENND5A-2 2D-D4, DENND5A-1 2D-D7, DENND5A-2 2D-D7 — WT | KD

Row labels (top to bottom):
Cytoskeleton
Adherens junction
Tight junction
Cell polarity complex
Apical trafficking
Polarity signaling

Tyrosine metabolism
Oxidative phosphorylation
Oxidized NAD binding
NAD binding
NADP metabolic process
Cytochrome complex assembly
Glycolysis
Galactose metabolism
Lysophospholipid metabolism
Glycerophospholipid metabolism
Cellular response to oxidative stress
Hypoxia
HIF targets
Caspase pathway
SMAD2/SMAD3 targets
TGF-β pathway
BMP signaling
PI3K/AKT activation
JAK/STAT signaling
Extracellular space
Chromatin remodeling

Cell cycle
E2F targets
G2M checkpoints
WNT/TCF signature
β-catenin-TCF targets
CTNNB1 targets
MYC targets
MYCN targets
Poly(A) RNA binding
Epithelial mesenchymal transition
TCA cycle
Pentose phosphate pathway
Fructose and mannose metabolism
Amino sugar and nucleotide sugar metabolism
Fatty acid metbolism
Vacuole

Golgi apparatus
Endosome
Intracellular vesicle
Protein localization
Regulation of transport
Transmembrane transport
Ion transport
Cell projection
Organ morphogenesis
Notch signaling
FAD binding
NADPH binding
Glutathione metabolism
Mitochondrion
Translation

Legend:
-2 ▬ 2
★ Down in KD (p<0.05)
★ Up in KD (p<0.05)
▭ Trafficking related

**Figure S3. The molecular characterization of MDCKII WT, *AVL9-* and *DENND5A*-KD cells with RNA-seq data. Related to Figures 3 and 6.**
Heatmap shows ssGSEA enrichment scores of MDCKII WT, *AVL9-* and *DENND5A*-KD cells, indicating intracellular trafficking and other functions related to cell polarity are suppressed in KD cells. The stars indicate gene sets with enrichment scores significantly different between WT and KD cells, with p < 0.05, identified by Wilcoxon tests. The same gene sets are used as in Figure 3.

# A

**3D down unique (475 genes)**

Cytoskeleton organization
Cell motility
Cell junction
Cell projection organization

0　　　　30

**2D down unique (256 genes)**

DREAM targets
Cell cycle
E2F targets
Embryonic stem cell

0　　　75　　　150

**3D up unique (341 genes)**

Intracellular transport
Mitochondrion
Ribosome
Translation

0　　　-log10 (p-value)　　　30

**2D up unique (317 genes)**

Cell cell signaling
Regulation of cell differentiation
Endosome
Cell projection organization

0　　　-log10 (p-value)　　　30

# B

○ WT-3D　　　◇ AVL9-KD-0h
△ WT-3D24h　　▽ AVL9-KD-2D
+ WT-0h　　　⊠ DENND5A-KD-0h
✕ WT-2D　　　✳ DENND5A-KD-2D

t-SNE_2 vs t-SNE_1

# C

○ WT-2D
△ AVL9-KD-2D
+ DENND5A-KD-2D

t-SNE_2 vs t-SNE_1

# D



Color Key
-3 -2 -1 0 1 2 3
Row Z-Score

| GO_CELL_PROJECTION | 9.45E-13 |
| GO_REGULATION_OF_TRANSPORT | 5.90E-12 |
| GO_GOLGI_APPARATUS | 8.96E-12 |
| GO_CELL_JUNCTION | 1.63E-10 |
| GO_INTRACELLULAR_VESICLE | 2.22E-10 |

| GO_REGULATION_OF_CELL_DEATH | 2.98E-13 |
| GO_REGULATION_OF_CELL_DIFFERENTIATION | 4.56E-13 |
| GO_TISSUE_DEVELOPMENT | 7.82E-13 |
| GO_REGULATION_OF_CELL_PROLIFERATION | 2.15E-12 |
| GO_CELL_JUNCTION | 2.15E-12 |

2D-D4-1　2D-D4-2　2D-D4-3　2D-D11-1　2D-D11-2　2D-D11-3　AVL9-1 2D-D4　AVL9-4 2D-D4　DENND5A-1 2D-D4　DENND5A-2 2D-D4　DENND5A-1 2D-D7　DENND5A-2 2D-D7

WT　　　　KD

# E

VIM

3D　early　2D

-2　0　2

kDa　3D-D8　3D-D3　3D24h　0h　2D-D4

50 — Vimentin

50 — α-tubulin

Vimentin/α-tubulin
p=0.002
p=0.02
3D-D8　3D-D3　3D24h　0h　2D-D4

# F

DENND5A

early　2D

-2　0　2

**2D-24h**　　**2D-D4**

DENND5A

DAPI

**Figure S4. Comparison between MDCKII WT and KD cells indicates the downregulation of intracellular trafficking genes and upregulation of cell proliferation genes in KD cells. Related to Figure 4.**

(A)      Functional enrichment of up- or down-regulated genes unique to 3D and 2D epithelial morphogenesis.

(B)      t-SNE analysis (11,090 genes) of MDCKII WT and KD samples shows three clusters: 3D, 2D and early.

(C)      t-SNE analysis (11,090 genes) of MDCKII WT and KD 2D over-confluent cells shows KD samples cluster distinctly from WT samples.

(D)      Heatmap shows expression of DE genes (FDR<0.01 and >2-fold change) between MDCKII WT and KD over-confluent cells, with their enriched functions and associated p-values also shown.

(E)      Protein level validation of transcriptional downregulation of *VIM* in both 3D and 2D samples.  Top heatmap shows RNA-seq data, with red denoting up-regulation and green denoting down-regulation.  Bar plot quantifies western blot experiments shown on the left, with p-values of Wilcoxon tests indicated.  Data are represented as mean ± SD with three biological replicates.

(F)      Protein level validation of transcriptional upregulation of *DENND5A* in 2D samples.  Top heatmap shows RNA-seq data, with red denoting up-regulation and green denoting down-regulation.  Bottom confocal images indicate MDCKII cells at 2D 24h and 2D day4, stained for DENND5A and DAPI.  Scale bar, 10 μm.

**A**

WT

CTNNB1
MYC

3D    early    2D

-1.5    1.5

AVL9 KD

CTNNB1
MYC

D0  2D

**B**

73/156 genes
upregulated in
AVL9 KD

14/156 genes
upregulated in
WT

Adhesion
Ligand/
receptor
Structural
Channels
Translation
Trafficking

Metabolism

Proteolysis

Signal transduction

Anti/proapoptosis

Nuclear regulatory

Transcription

DNA repair

3D    early    2D    0h  2D

WT        AVL9 KD

-3    3

p=0.03

AVL9 KD    WT

**C**

20x    WT  0h    AVL9 KD  0h    DENND5A KD  0h    WT 2D 24h    AVL9 KD  2D 24h    DENND5A KD 2D 24h

β-catenin

DAPI

Merge

40x

β-catenin

DAPI

Merge

63x z-stack

Merge

**Figure S5.  *MYC* target genes are more abundantly expressed during epithelial morphogenesis of *AVL9*-KD cells, with β-catenin detected in the nucleus and cytoplasm after the first cell cycle, unlike WT cells.  Related to Figure 5.**

(A)       At the transcript level, *CTNNB1* is upregulated while *MYC* is downregulated in both 3D and 2D samples, compared to the early group, in WT cells.  In *AVL9*-KD cells, *MYC* stays relatively constant. Heatmaps show RNA-seq data, with red denoting up-regulation and green denoting down-regulation, while bar plots show qRT-PCR data.  Data are represented as mean ± SD with three biological replicates.

(B)       Heatmap shows expression of *MYC* target genes grouped based on their functions, up-regulation in red and down-regulation in green.  Right plot indicates the *MYC* target enrichment scores and p-value of *AVL9*-KD cells compared to WT cells, calculated from GSEA.

(C)       Representative confocal images of WT, *AVL9*- and *DENND5A*-KD cells at 0h and 2D 24h, stained for β-catenin and DAPI.  Scale bar, 20 μm.

# A

## 3D36h-D8_up

3D_up  3D36h-D8_up  3D_down

our study 941 | 162 | GSE32495 GSE28381 896 | 71 | 850

p-value: p<0.0001   p=0.0005
fraction: 0.15   0.08

Cell projection organization
HNF1A targets
Protein phosphorylation
Cell motility

0   20   40   60

## 3D36h-D8_down

3D_down  3D36h-D8_down  3D_up

778 | 143 | 508 | 13 | 1090

p-value: p<0.0001   p=1
fraction: 0.22   0.02

Embryonic stem cell
Cell cycle
E2F targets
DREAM targets

0   50   100   150   200

## 2D36h-D8_up

2D_up  2D36h-D8_up  2D_down

679 | 401 | 1146 | 3 | 699

p-value: p<0.0001   p=1
fraction: 0.37   0.004

Vacuole
Golgi apparatus
Intracellular transport
HNF1A targets

0   20   40   60

## 2D36h-D8_down

2D_down  2D36h-D8_down  2D_up

307 | 395 | 1044 | 12 | 1068

p-value: p<0.0001   p=1
fraction: 0.56   0.01

Embryonic stem cell
E2F targets
Cell cycle
DREAM targets

0   100   200   300
-log10 (p-value)

# B

## 3D-HGF down

3D_up  3D-HGF_down  3D_down

779 | 324 | GSE49518 1220 | 36 | 885

p-value: p<0.0001   p=1
fraction: 0.29   0.04

Golgi apparatus
HNF1A targets
Intracellular transport
Vacuole

0   30   60

## 3D-HGF up

3D_down  3D-HGF_up  3D_up

621 | 300 | 1068 | 4 | 1099

p-value: p<0.0001   p=1
fraction: 0.33   0.004

Cell cycle
Embryonic stem cell
RNA binding
DREAM targets

0   100   200

## 2D-HGF down

2D_up  2D-HGF_down  2D_down

805 | 275 | 970 | 6 | 696

p-value: p<0.0001   p=1
fraction: 0.25   0.01

Golgi apparatus
Intracellular transport
Cilium
HNF1A targets

0   20   40

## 2D-HGF up

2D_down  2D-HGF_up  2D_up

487 | 215 | 972 | 18 | 1062

p-value: p<0.0001   p=1
fraction: 0.31   0.02

Embryonic stem cell
E2F targets
Cell cycle
DREAM targets

0   75   150
-log10 (p-value)

# C

## 3D vs. 2D

### Wells et al., 2013

3D_up  3D-D8_up

our study 1540 | 121 | GSE28381 1303

p-value: p=0.1088
fraction: 0.08

Cytoskeleton organization
E2F targets
Cell cycle
DREAM targets

0   75   150

2D_up  2D-D8_up

1958 | 66 | 701

p=0.7538
0.09

Ribosome
Intracellular transport
Cilium
Translation

0   12   24

### Chacon-Heszele et al., 2014

3D_up  3D-D10_up

our study 1576 | 85 | GSE49518 707

p-value: p=0.0008
fraction: 0.11

G2M checkpoint
E2F targets
Cell cycle
DREAM targets

0   90   180
-log10(p-value)

2D_up  2D-D6_up

1843 | 181 | 1361

p=0.0006
0.12

Mitochondrion
Oxidation reduction process
Endoplasmic reticulum
Intracellular transport
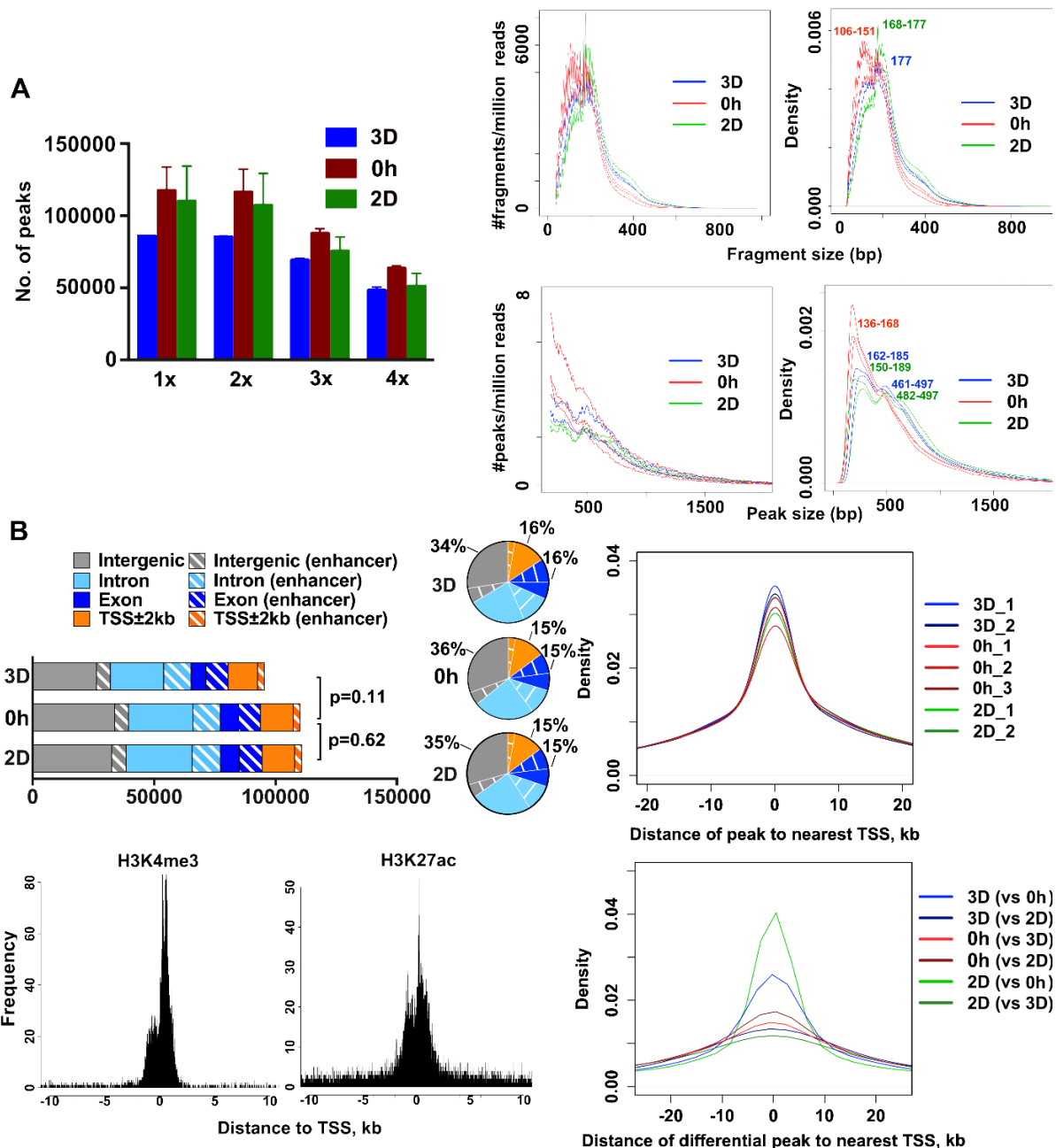
0   30   60
-log10(p-value)

**Figure S6. We found consistence between our study and previous microarray work for transcriptomic changes during epithelial morphogenesis, but not for transcriptomic differences between 2D and 3D culture. Related to Figures 4 and 6.**
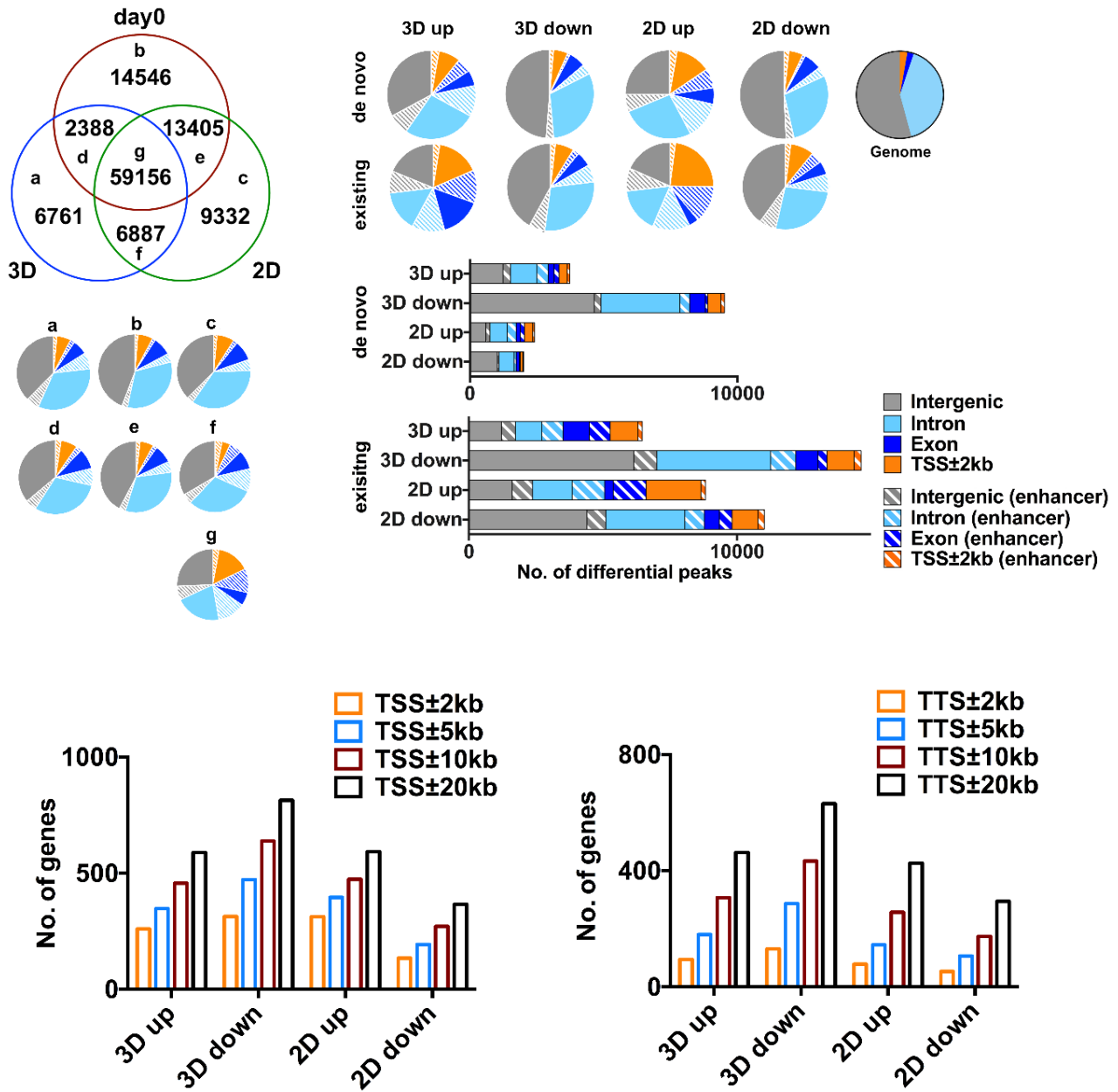
(A)      Comparison with DE genes (FDR< 0.05 and >2-fold change) between 36h and day8 samples from two microarray studies (Wells et al., 2013, Galvez-Santisteban et al., 2012).  The analysis was performed using limma R package.  Venn diagrams show overlap between these microarray DE genes and up-/down-regulated genes during 3D or 2D epithelial morphogenesis from our study (Figure 4A).  Fractions of overlap and p-values (calculated with 10,000 permutations) are indicated.  Bar plots on the right show the enriched functions of the microarray DE genes and associated p-values, identified with the GSEA web tool.

(B)      Comparison with DE genes (FDR<0.05 and >1.5-fold change) between with and without HGF treatment from another microarray study (Chacon-Heszele et al., 2014).  The analysis was performed using limma R package.  The findings are presented as described in panel A.  Fractions of overlap and p-values (calculated with 10,000 permutations) are indicated.  Bar plots on the right show the enriched functions of the microarray DE genes and associated p-values, identified with the GSEA web tool.

(C)      Comparison with 3D-2D DE genes from two microarray studies: 1) between day8 3D and 2D cells (FDR<0.05) (Wells et al., 2013); and 2) between 3D (day10) and 2D (day6) cells (FDR<0.05 and >1.5-fold change) (Chacon-Heszele et al., 2014).  The analysis was performed using limma R package.  DE genes for our study are indicated Figure 6A.  Venn diagrams indicate the overlap. Fractions of overlap and p-values (calculated with 10,000 permutations) are indicated. Bar plots indicate enriched functions of the microarray DE genes and associated p-values, identified with the GSEA web tool.

Note:  Different selection cutoffs were used for DE gene identification among the microarray studies, in order to keep the number of DE genes comparable to our study for each corresponding change for statistic tests.

**Figure S7. Chromatin accessibility decreases during MDCKII 3D epithelial morphogenesis, revealed by ATAC-seq analysis. Related to Figure 7.**
(A)     Left bar plot indicates that 3D cells have the fewest ATAC-seq peaks, identified with the software MACS2 (Feng et al., 2012) at FDR<0.05 and ≥1 to 4-fold enrichment (the ratio of the ATAC-seq peak signal versus the background signal).  Right plots show the distribution of ATAC-seq fragment size and peak size, indicating that 0h cells have more but smaller open chromatin regions, compared to 3D and 2D samples.  Data are represented as mean ± SD with biological replicates (see Table S7).
(B)     Top left bar plot and pie charts show the distribution of ATAC-seq peaks in the genome.  P-values were calculated with paired t-test using the total numbers of ATAC-seq peaks detected in each genomic region (intergenic, etc.) between 0h and 3D or 2D cells.  Bottom left plots indicate the locations of canine cell H3K4me3 and H3K27ac ChIP-seq peaks (Villar et al., 2015) relative to TSSs used for our analysis.  Right plots show the locations of ATAC-seq peaks or differential peaks relative to TSSs.

**Figure S8. Genomic regions with increased chromatin accessibility during MDCKII epithelial morphogenesis are more likely to be found in promoters and exonic regions. Related to Figure 7.**
Venn diagram at the top left shows the number of shared and unique ATAC-seq peaks in different samples, while the pie charts beneath indicate the distribution of shared or unique ATAC-seq peaks in the genome. Pie charts and bar plots at the top right show genomic locations of existing and de novo differential ATAC-seq peaks. Bar plots at the bottom indicate the number of DE genes that have differential ATAC-seq peaks located at their TSS and TTS regions as specified.
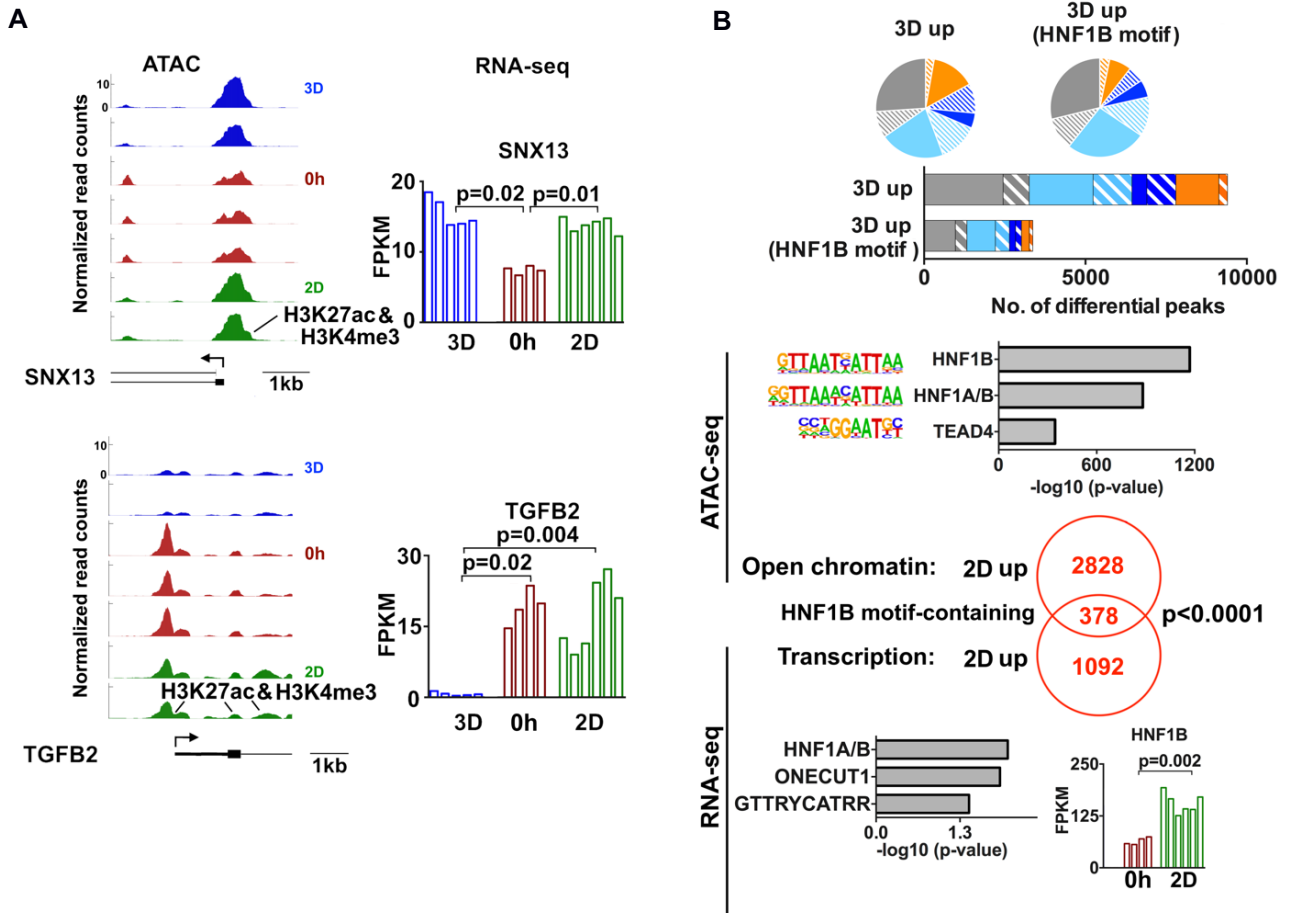
**Figure S9. Examples of chromatin accessibility correlating with gene expression, and upregulation of HNF1B target genes during MDCKII epithelial morphogenesis. Related to Figure 7.**

(A)    *SNX13*, up-regulated in both 2D and 3D samples, and *TGFB2*, down-regulated in 3D but not in 2D samples, both have corresponding chromatin accessibility changes in their promoter, revealed by ATAC-seq. P-values of Wilcoxon tests are indicated.

(B)    Top plots show the genomic distribution of expanded and HNF1B motif-containing ATAC-seq peaks in 3D samples. Plots below are for 2D samples. HNF1B binding sites are among the top-most enriched motifs in expanded ATAC-seq peaks and the promoters of upregulated genes during 2D epithelial morphogenesis. The analysis was performed using HOMER and Enrichr. Venn diagram indicates significant overlap between genes upregulated and genes with expanded and HNF1B-containing ATAC-seq peaks detected in their gene body or 20kb flanking regions. P-value for the overlap was calculated via 10,000 permutations. Bar plot shows upregulation of HNF1B, and the p-value of Wilcoxon test is indicated.
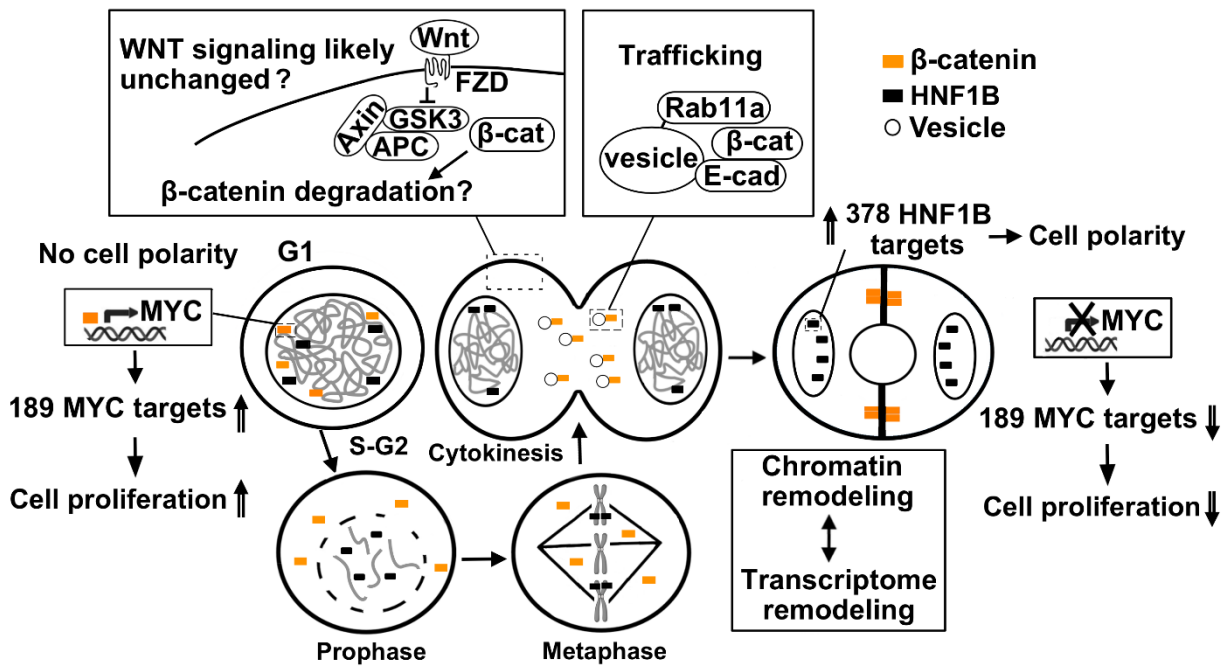
**Figure S10. Our model proposes that the qualitative change in the transcriptome during MDCKII 3D cystogenesis is a result of the redistribution of transcription factors during the first cell cycle.** Related to **Figures 2, 4, 5 and 7.** Our model is exemplified by β-catenin, which translocates to the cell-cell junctions, and HNF1B, which remains chromatin-bound via mitotic bookmarking, during the first cell cycle.

**Transparent Methods**

**MDCKII cell culture**
MDCKII cells (Li et al., 2014) were maintained in Eagle's minimal essential medium (Corning, #10-010-CM) with 5% fetal bovine serum (Atlanta Biological, #S11550), 10 mM HEPES, 2 mM glutamine, 100 U/ml penicillin, and 100 mg/ml streptomycin, at 37°C in an incubator with 5% $CO_2$. The media was changed every other day. Only cells that are healthy and when fully polarized, display well-defined cell-cell junction, pentagon shape and uniform size were used for experiments as below.

Cells grown at <70% confluence were harvested via trypsinization. Then, $1 \times 10^6$ cells were seeded on 100 mm plates until reaching ~10% confluence (e.g., single cell stage). The cells were harvested via trypsinization, with a portion used as 0h samples and the rest used for seeding 3D and 2D cultures. For 3D cultures, cells were grown on top of growth factor-reduced Matrigel (Corning, #354230) and in media containing 2% Matrigel, with seeding density at 60 cells/mm$^2$ and daily media change. For 2D cultures, cells were seeded on Transwell filters with 0.4 µm pore size (Corning) with the same seeding density as 3D cultures and grown for 11 days, or with seeding density of 3400 cells/mm$^2$ and grown for 4 days, to reach over-confluence.

**Immunofluorescence and confocal microscopy**
MDCKII cells were fixed with 4% PFA, then permeabilized with 0.5% Triton X-100 in PBS and blocked with 10% normal donkey serum and 0.1% Triton X-100 in PBS. Primary and secondary antibodies incubations were in 1% donkey serum and 0.1% Triton X-100. MDCKII cysts were stained as previously described (Li et al., 2014). The antibodies used include Goat polyclonal anti-E-cadherin (R&D Systems, #AF648), Rabbit polyclonal anti-β-catenin (Santa Cruz, #sc-7199), Rabbit polyclonal anti-DENND5A (Sigma-Aldrich, #HPA052923), Alexa Fluor 488 Donkey anti-goat IgG (Jackson ImmunoResearch, #705-545-147), and Alexa Fluor 647 Donkey anti-rabbit IgG (Jackson ImmunoResearch, #711-605-152). MitoTracker staining was performed using MitoTracker Red CMXRos (Thermo Fisher, #M7512) according to manufacturer's instruction. Confocal images were taken on Zeiss LSM 710 confocal microscope with 20x, 40x, 63x or 100x oil objectives. Bright-field images were taken using Leica DFC320.

Imaging analysis was performed using the software Fiji. To quantify the intensity of β-catenin and E-cadherin staining (Figure 5B), nuclei and cell-cell junctions were outlined based on the DAPI staining and bright-field images. After region of interest (ROI) selection, the Intermode Mode was used to choose a threshold of each image to adjust the intensity of ROI. Then, the Measure function was used to quantify the intensity of ROI. Similar procedures were performed to quantify MitoTracker staining intensity (Figure 6B).

**Western blot analysis**
3D cysts were isolated and collected from Matrigel using cold PBS and 2D cells were collected using cell scraper. Protein lysate of MDCK cells was prepared in RIPA buffer. Protein extracts were separated by SDS-PAGE (Bio-Rad) for 60 min at 100V, then transferred onto PVDF membranes (Bio-Rad) for 30 min at 100V, and blocked with 5% non-fat milk in 0.05% TBST for 2 hr. The membranes were then incubated with primary antibodies at 4°C overnight, and incubated with secondary antibodies after washing with TBST. The antibodies used include Rabbit monoclonal anti-MYC (Abcam, #ab32072), Rabbit monoclonal anti-Vimentin (Abcam, # Ab92547), Mouse monoclonal anti-α-Tubulin (Sigma-Aldrich, #T6199), Rabbit polyclonal anti-β-catenin (Santa Cruz, #sc-7199), Goat anti-rabbit IgG-HRP (Santa Cruz, #sc-2030), and Goat anti-mouse IgG-HRP (Santa Cruz, #sc-2031). The western ECL blotting substrate (Bio-Rad) was applied on the membrane before film exposure.

**Paired-end RNA-seq**
Cysts grown for 24 hours, 3, 5, 8, 14 days in 3D condition, cells for seeding (0h) and over-confluent cells grown in 2D condition were collected and their RNA was extracted using the AllPrep DNA/RNA Mini Kit (Qiagen, #80204). Only samples with ratio of absorbance at 260 nm and 280 nm of ~2.0 were subjected to further analyses. qRT-PCR of selected genes were performed for quality control. cDNA was made using iScript cDNA synthesis kit (Bio-Rad, #1708890) and qRT-PCR was performed with iQ SYBR Green Supermix (Bio-Rad, #1708880) and the primers previously described (Liu et al., 2014). RNA-seq libraries were constructed using KAPA Stranded mRNA-

Seq Kit.  The samples were subjected to 75 or 125 bp paired-end sequencing using Illumina HiSeq 2500 at BGI (the first biological replicate) or NextSeq 500 at Georgia Genomics Facility (the other biological replicates).

## ATAC-seq

ATAC-seq was performed as previously described (Lu et al., 2017).  2D cells were grown on 100 mm plate for 4 days, and 3D cysts were cultured on Matrigel for 19 days.  0h and 2D cells were collected using a cell scraper, and 3D cysts were isolated from Matrigel with cell recovery solution (Corning, #354253).  After washing with cold PBS, the cells were dissociated in lysis buffer (1% NP40, 10mM Tris pH 7.5, 10 mM NaCl, 3 mM $MgCl_2$, 1mM EDTA) using a syringe.  The samples were filtered with miracloth.  Crude nuclei were sorted by flow cytometer (Beckman Coulter MoFlo XDP) with DAPI.  The nuclei were pelleted and resuspended in the transposase reaction mix (25 µL 2xTagmentation buffer, 2.5 µL transposase and 22.5 µL nuclease-free water; Nextera Kit, Illumina, #FC-121-1030), and incubated at 37°C for 30 min.  Immediately following transposition, transposed DNA was purified with MinElute Reaction Cleanup Kit (Qiagen, #28204), and amplified using Phusion DNA polymerase (NEB, # M0530S) for 10-13 cycles, with sequencing indexing primer for ATAC library preparation described previously (Buenrostro et al., 2013).  Amplified libraries were purified with AMPure beads.  The libraries were subjected to 36 bp paired-end sequencing using Illumina NextSeq 500.

## RNA-seq data analysis

RNA-seq data analyses were performed as previously described (Wang et al., 2018a, Wang et al., 2018b).  Read pairs were aligned to the dog reference genome canFam3 with TopHat v2.0.14 or HISAT v2.1.0 (Pertea et al., 2016, Trapnell et al., 2012). Cufflinks v2.2.1 and canine gene annotation canFam3.1.84 were used to quantify gene expression as FPKM values using only uniquely mapped read pairs.  Clustering of the samples were performed using hierarchical clustering and NMF (Brunet et al., 2004).  For hierarchical clustering, the *pvclust* R package was used with top variably expressed genes and the centroid method.  Bootstrapping was performed with a total of 1,000 permutations.  The *NMF* R package was used with 30 runs for the rank determination and 200 runs for the final results.  DE genes were identified by using the *DESeq* R package, and their enriched functions were identified using the GSEA webtool at https://www.gsea-msigdb.org/gsea/index.jsp.  ssGSEA was performed with predefined genes (see Table S3 and Figure 3) using GenePattern (Reich et al., 2006).  Motif enrichment analysis was performed using the software Enrichr (Kuleshov et al., 2016).

## ATAC-seq data analysis

ATAC-seq data were analyzed as previously described (Milani et al., 2016).  Sequencing reads were mapped to the canFam3 genome using BWA v0.7.10 (Li and Durbin, 2009).  ATAC-seq peaks were identified using MACS v2.1.1 with FDR < 0.05 (Zhang et al., 2008).  The differential open chromatin regions were identified using the bdgdiff option of MACS.  Peaks were mapped to TSS±2kb and other regions using Bedops v2.4.30.  Peaks were visualized by uploading the data to the UCSC genome browser.  Motif enrichment analysis was performed using HOMER v4.10 (Heinz et al., 2010).

## Supplemental References

BRUNET, J. P., TAMAYO, P., GOLUB, T. R. & MESIROV, J. P. 2004. Metagenes and molecular pattern discovery using matrix factorization. *Proceedings of the National Academy of Sciences of the United States of America,* 101**,** 4164-4169.
BUENROSTRO, J. D., GIRESI, P. G., ZABA, L. C., CHANG, H. Y. & GREENLEAF, W. J. 2013. Transposition of native chromatin for fast and sensitive epigenomic profiling of open chromatin, DNA-binding proteins and nucleosome position. *Nat Meth,* 10**,** 1213-1218.
CHACON-HESZELE, M. F., ZUO, X., HELLMAN, N. E., MCKENNA, S., CHOI, S. Y., HUANG, L., TOBIAS, J. W., PARK, K. M. & LIPSCHUTZ, J. H. 2014. Novel MAPK-dependent and -independent tubulogenes identified via microarray analysis of 3D-cultured Madin-Darby canine kidney cells. *American journal of physiology. Renal physiology,* 306**,** F1047-F1058.
FENG, J., LIU, T., QIN, B., ZHANG, Y. & LIU, X. S. 2012. Identifying ChIP-seq enrichment using MACS. *Nat Protoc,* 7**,** 1728-40.
GALVEZ-SANTISTEBAN, M., RODRIGUEZ-FRATICELLI, A. E., BRYANT, D. M., VERGARAJAUREGUI, S., YASUDA, T., BANON-RODRIGUEZ, I., BERNASCONE, I., DATTA, A., SPIVAK, N., YOUNG, K., SLIM, C. L., BRAKEMAN, P. R., FUKUDA, M., MOSTOV, K. E. & MARTIN-BELMONTE, F. 2012.

Synaptotagmin-like proteins control the formation of a single apical membrane domain in epithelial cells. *Nat Cell Biol,* 14**,** 838-49.

HEINZ, S., BENNER, C., SPANN, N., BERTOLINO, E., LIN, Y. C., LASLO, P., CHENG, J. X., MURRE, C., SINGH, H. & GLASS, C. K. 2010. Simple Combinations of Lineage-Determining Transcription Factors Prime cis-Regulatory Elements Required for Macrophage and B Cell Identities. *Molecular Cell,* 38**,** 576-589.

KULESHOV, M. V., JONES, M. R., ROUILLARD, A. D., FERNANDEZ, N. F., DUAN, Q., WANG, Z., KOPLEV, S., JENKINS, S. L., JAGODNIK, K. M., LACHMANN, A., MCDERMOTT, M. G., MONTEIRO, C. D., GUNDERSEN, G. W. & MA'AYAN, A. 2016. Enrichr: a comprehensive gene set enrichment analysis web server 2016 update. *Nucleic Acids Res,* 44**,** W90-7.

LI, H. & DURBIN, R. 2009. Fast and accurate short read alignment with Burrows–Wheeler transform. *Bioinformatics,* 25**,** 1754-1760.

LI, Y., XU, J., XIONG, H., MA, Z., WANG, Z., KIPREOS, E. T., DALTON, S. & ZHAO, S. 2014. Cancer driver candidate genes AVL9, DENND5A and NUPL1 contribute to MDCK cystogenesis. *Oncoscience,* 1**,** 854-65.

LIU, D., XIONG, H., ELLIS, A. E., NORTHRUP, N. C., RODRIGUEZ, C. O., JR., O'REGAN, R. M., DALTON, S. & ZHAO, S. 2014. Molecular homology and difference between spontaneous canine mammary cancer and human breast cancer. *Cancer research,* 74**,** 5045-5056.

LU, Z., HOFMEISTER, B. T., VOLLMERS, C., DUBOIS, R. M. & SCHMITZ, R. J. 2017. Combining ATAC-seq with nuclei sorting for discovery of cis-regulatory regions in plant genomes. *Nucleic Acids Res,* 45**,** e41.

MILANI, P., ESCALANTE-CHONG, R., SHELLEY, B. C., PATEL-MURRAY, N. L., XIN, X. F., ADAM, M., MANDEFRO, B., SAREEN, D., SVENDSEN, C. N. & FRAENKEL, E. 2016. Cell freezing protocol suitable for ATAC-Seq on motor neurons derived from human induced pluripotent stem cells. *Scientific Reports,* 6.

PERTEA, M., KIM, D., PERTEA, G. M., LEEK, J. T. & SALZBERG, S. L. 2016. Transcript-level expression analysis of RNA-seq experiments with HISAT, StringTie and Ballgown. *Nature Protocols,* 11**,** 1650.

REICH, M., LIEFELD, T., GOULD, J., LERNER, J., TAMAYO, P. & MESIROV, J. P. 2006. GenePattern 2.0. *Nat Genet,* 38**,** 500-1.

TRAPNELL, C., ROBERTS, A., GOFF, L., PERTEA, G., KIM, D., KELLEY, D. R., PIMENTEL, H., SALZBERG, S. L., RINN, J. L. & PACHTER, L. 2012. Differential gene and transcript expression analysis of RNA-seq experiments with TopHat and Cufflinks. *Nature Protocols,* 7**,** 562-578.

WANG, J., WANG, T., BISHOP, M. A., EDWARDS, J. F., YIN, H., DALTON, S., BRYAN, L. K. & ZHAO, S. 2018a. Collaborating genomic, transcriptomic and microbiomic alterations lead to canine extreme intestinal polyposis. *Oncotarget,* 9**,** 29162-29179.

WANG, J., WANG, T., SUN, Y., FENG, Y., KISSEBERTH, W. C., HENRY, C. J., MOK, I., LANA, S. E., DOBBIN, K., NORTHRUP, N., HOWERTH, E. W. & ZHAO, S. 2018b. Proliferative and Invasive Colorectal Tumors in Pet Dogs Provide Unique Insights into Human Colorectal Cancer. *Cancers (Basel),* 10.

WELLS, E. K., YARBOROUGH, O., 3RD, LIFTON, R. P., CANTLEY, L. G. & CAPLAN, M. J. 2013. Epithelial morphogenesis of MDCK cells in three-dimensional collagen culture is modulated by interleukin-8. *Am J Physiol Cell Physiol,* 304**,** C966-75.

ZHANG, Y., LIU, T., MEYER, C. A., EECKHOUTE, J., JOHNSON, D. S., BERNSTEIN, B. E., NUSBAUM, C., MYERS, R. M., BROWN, M., LI, W. & LIU, X. S. 2008. Model-based Analysis of ChIP-Seq (MACS). *Genome Biology,* 9**,** R137.