A QUALITATIVE CHANGE IN THE TRANSCRIPTOME DURING MDCKII EPITHELIAL MORPHOGENESIS

by

TIANFANG WANG

(Under the Direction of Shaying Zhao)

ABSTRACT

Madin-Darby canine kidney II (MDCKII) cells, originally derived from canine kidney epithelial tissues, are widely employed in various research areas, including the study of cell polarity, epithelial morphogenesis, kidney disease, EMT, etc. In MDCKII 3D culture, epithelial morphogenesis takes place as a single cell develops into a cyst consisting of a monolayer of cells enclosing a central lumen through tightly regulated cell proliferation and establishment of cell polarity. In contrast, MDCKII cells develop into a monolayer with well-established cell polarity but without lumen formation in 2D culture. During MDCKII epithelial morphogenesis, transcriptome reprogramming may play a critical role in the regulation of differentiation, as observed in other cell differentiation systems. However, the complete transcriptome profile during the process is largely unknown. Moreover, fundamental questions remain unanswered about the transcriptome changes underlying the process. For instance, does the transcriptome gradually change or suddenly switch at a time point? What are the transcriptome similarities and differences between 3D and 2D epithelial morphogenesis?

To answer the questions, we performed a full time-course RNA-seq analysis for MDCKII 3D cystogenesis, together with fully polarized 2D cells. Our analysis revealed that a qualitative transcriptome change occurs after the first cell division and in parallel with lumen formation. Further analysis of *DENND5A*- and *AVL9*- knockdown cells indicated intracellular trafficking is related to the transcriptome changes. Specifically, β-catenin translocation from nucleus to cell-cell junctions during the first cell division likely leads to the down-regulation of MYC and other cell cycle genes. Integration of RNA-seq and ATAC-seq data uncovered a possible mechanism that HNF1B activation leads to transcriptome changes that contribute to cell polarity establishment. Based on these findings, our study proposed a model that redistribution of transcription factors during cell division may drive transcriptome remodeling. In addition, we identified some differences between 3D and 2D epithelial morphogenesis. First, chromatin accessibility decreases more drastically during 3D epithelial morphogenesis. Second, active mitochondria are retained only during 3D epithelial morphogenesis. Our work deepens the understanding of MDCKII epithelial morphogenesis and provides insight into mechanisms underlying the qualitative transcriptome change during differentiation.

INDEX WORDS: MDCKII, epithelial morphogenesis, transcriptome, intracellular trafficking, chromatin, β-catenin, HNF1B, RNA-seq, ATAC-seq

A QUALITATIVE CHANGE IN THE TRANSCRIPTOME DURING MDCKII EPITHELIAL MORPHOGENESIS

by

TIANFANG WANG

B.S., City University of Hong Kong, China, 2014

A Dissertation Submitted to the Graduate Faculty of The University of Georgia in Partial Fulfillment of the Requirements for the Degree

DOCTOR OF PHILOSOPHY

ATHENS, GEORGIA

2020

A QUALITATIVE CHANGE IN THE TRANSCRIPTOME DURING MDCKII EPITHELIAL MORPHOGENESIS

by

TIANFANG WANG

Major Professor:	Shaying Zhao
Committee:	Stephen Dalton
	Hang Yin
	Kevin Dobbin

Electronic Version Approved:

Ron Walcott

Dean of the Graduate School

The University of Georgia

December 2020

DEDICATION

This dissertation is dedicated to my loving family and friends.

ACKNOWLEDGEMENTS

First and foremost, I would like to express my sincere gratitude to my advisor, Dr. Shaying Zhao for providing me with the opportunity to work in her lab as a graduate student, and for supporting me during the past years. With her mentorship, I learned so much about how scientific research is conducted. The dissertation work would not have been possible without her insightful guidance and persistent encouragement.

Besides, I would like to express my gratitude to the present and previous lab members. Particularly, I would like to thank our previous technician, Mrs. Yaping Li, who did phenomenal work to establish the cell lines useful to my project and provided helpful suggestions on cell culturing and other experiments when I just joined the lab.

I would like to thank my committee members: Drs. Stephen Dalton, Hang Yin, and Kevin Dobbin, for their help on my projects. Dr. Yin has provided a lot of helpful suggestions for my projects and experiments. Dr. Dobbin has been very patient in answering all kinds of questions in statistics with clear and thorough explanations. Dr. Dalton has provided insightful input into my projects, which is very helpful in the long term.

I am also grateful to my previous advisors during my undergraduate study, Drs. Eddie Chi-Him Ma and Hon-Yeung Cheung, who provided me with wonderful opportunities to gain research experience in their labs and helped me with my application to the graduate study.

Last but not least, I would like to express my deepest and warmest gratitude to my family and friends. There were many ups and downs during the past few years. What kept me going are their endless support, patience, and love.

TABLE OF CONTENTS

Page

CHAPTER 1

INTRODUCTION AND LITERATURE REVIEW

INTRODUCTION TO MDCK CELL MODELS

Madin-Darby canine kidney (MDCK) cells were primarily derived from the normal kidney cortex of a female Cocker Spaniel dog in 1958. Later studies suggested that proximal/distal tubule and collecting duct are likely the origin of MDCK cells (Chuman et al., 1982; Jalal et al., 1991). There are several strains derived from the parental MDCK cell line. Among all the MDCK subtypes, MDCKII is one of the most commonly used strains (Dukes et al., 2011). As prototypical epithelial cells, MDCKII cells have well-established epithelial cell polarity in 3D and 2D culture. Researchers have extensively employed MDCKII cells as a model to study epithelial morphogenesis and lumenogenesis (O'Brien et al., 2002). The understanding of epithelial tissue development, epithelial cell trafficking, and cell polarity is heavily based on experiments using MDCKII cells (Dukes et al., 2011).

In conventional MDCK 2D culture, the cells are grown on a culture dish with a regular plastic substrate or a transwell with a filter membrane to form a monolayer of cells. The filter membrane allows the basal side of MDCK cells to be accessible to factors and nutrients (O'Brien et al., 2002). With the filter membrane as support, MDCK cells can form well-established cell polarity when grown over-confluence (Figure 1.1).

Typically, the cells do not undergo extensive dome formation when grown on a filter membrane as opposed to cell culture using a regular plastic substrate.

Compared with 2D culture, MDCK 3D culture better recapitulates the physiological conditions of epithelia and thus has become a more powerful model to study epithelial morphogenesis and architecture. The cystogenesis in 3D culture is triggered by collagen gel or other extracellular matrix (ECM)-based gels, such as Matrigel. Developed from a single cell, an MDCK cyst forms through lumen formation in the center, oriented cell division, and establishment of cell polarity. As a consequence, a mature cyst consists of a monolayer of cells enclosing a central lumen that offers a microenvironment in addition to the ECM in which the cyst is embedded (Figure 1.1).

Both 2D and 3D culture of MDCK cells serve as useful tools for various areas of research, which will be discussed in the next section.

APPLICATIONS OF MDCK CELL MODELS

Cell models for studying epithelial cell polarity

MDCK cell models have been extensively utilized for dissecting the components of epithelial cell polarity and the mechanisms of cell polarity regulation. For instance, the functions of apical complexes have been elucidated using functional assays on MDCKII 3D culture. Using Cdc42 knockdown MDCK cysts, a study revealed the regulatory role of Cdc42 in normal lumen formation (Martin-Belmonte et al., 2008). Activated by Cdc42, the downstream Par polarity complex regulates cell polarity. Utilizing Par3 knockdown MDCK cysts, another study revealed that Par3 controls the orientation of the spindle and contributes to the formation of a single lumen through

aPKC signaling (Hao et al., 2010). Through further investigations from several groups, the molecular chains for the regulation of apical cell polarity have been identified, which will be discussed later.

Cell models for studying kidney diseases

Kidney diseases are very prevalent and affect more than tens of million people in the USA. Cell line models are used as a supplement of *in vivo* models to gain a better understanding of these diseases. As a cell line originated from kidney epithelial tissue, MDCK cells are employed as disease models for understanding various kidney diseases, including polycystic kidney disease and drug-induced nephrotoxicity, (Desrochers et al., 2014). They serve as useful platforms to test drug candidates for treating these diseases.

Polycystic kidney disease is developed due to dysregulation of cell proliferation and excessive fluid secretion, which is likely a consequence of abnormal cystic fibrosis transmembrane regulator and mTOR signaling, which are both negatively regulated by the AMPK pathway (Bergmann et al., 2018). A study utilized MDCK cells to test an AMPK stimulating drug and examined the effect of activation of the AMPK pathway on cystogenesis (Takiar et al., 2011). Indeed, the drug shows effectiveness in mitigating excessive cystogenesis via activation of the AMPK pathway, and thus indicates its potential use in treating the disease. In addition, other studies used the MDCK cyst model to test and screen various drug candidates for treating polycystic kidney disease (He et al., 2020; Yang et al., 2008; Zhou et al., 2012).

MDCK cells are also used to study the mechanism of the development of kidney diseases and to identify potential drug targets. In a recent study, it has been found that NHA2, an ion transporter, has with its high expression correlated with increased cyst size

in both patients and the MDCK cell model (Prasad et al., 2019). This study further uncovered the pathway upstream of NHA2, which could be a potential drug target.

Study of mechanisms underlying EMT

Another active research area that extensively utilizes MDCK cell models is the study of EMT. EMT is a critical process involved in the development process and tumor progression, which occurs with a loss of epithelial features, especially the cell polarity, and a gain of migration and invasiveness of mesenchymal cells.

There are several types of MDCK cell models, including those induced by TGF-β, LEF1, or Ras, to recapitulate EMT on culture dish (Carmona et al., 2014; Domaschenz et al., 2017; Kobayashi and Ozawa, 2013; Shukla et al., 2015). A comprehensive miRNA and mRNA profiling study used Ras-transformed MDCK cells to reveal novel miRNAs, genes, and signaling pathways that are linked to EMT (Shukla et al., 2015). With the use of TGF-β induced model as a comparison, another study uncovered the key regulatory role of H2A.Z in controlling EMT, using H2A.Z knockdown MDCK cells (Domaschenz et al., 2017). In another study, Carmona and colleagues generated DNA methylation profiles for MDCK cells that undergo EMT (Carmona et al., 2014). Further, they identified differentially methylated regions compared with control MDCK cells, which could be translated into a human EMT cell line model and breast cancer (Carmona et al., 2014).

In summary, EMT models based on MDCK cells are useful for investigating the mechanisms underlying EMT, with *in vivo* implications.

Validation and study of tumor suppressor genes

MDCK cell models have been extensively utilized to validate tumor suppressor genes and to understand the related mechanisms. Using a dog and human comparison strategy, more than 70% of genes harboring deletion in colorectal cancer were found related to cell polarity (Tang et al., 2014). Two of the discovered candidate tumor suppressor genes, *DENND5A* and *AVL9*, are related to intracellular trafficking and cell polarity (Li et al., 2014). Abnormal lumen formation during cystogenesis as well as aberrant regulation of cell division and cell polarity were observed in knockdown cells of either of the two genes (Figure 1.2). Intriguingly, the multi-lumen phenotype observed in both knockdown cells is reminiscent of colorectal polyps, with the feature of disrupted coordination between cell proliferation and cell polarity (Figure 1.2). The further mechanistic study indicated that spindle orientation during mitosis, in addition to dysregulation of membrane trafficking, contributes to the abnormal lumen formation (Li et al., 2014).

In another study, the functions in the regulation of cell polarity and epithelial morphogenesis were elucidated for a well-studied tumor suppressor, *APC* (Lesko et al., 2015). Lesko and colleagues showed a filled-lumen phenotype in *APC*-knockdown cells in 3D culture conditions, where the cells completely lose cell polarity. Further experiments using β-catenin/TCF transcriptional reporter assay and examination of β-catenin expression revealed that the abnormal phenotype of *APC*-knockdown cells is independent of the Wnt signaling pathway (Lesko et al., 2015).

In addition to the aforementioned studies, other studies also used MDCK cell models to study tumor suppressor genes, many of which were found to be important in

regulating epithelial cell polarity (Astrinidis et al., 2002; Masuda et al., 2002; McLaughlin et al., 2007; Musch et al., 2002; Wei et al., 2012; Yamashita et al., 2015; Zhang et al., 2015).

MDCKII EPITHELIAL MORPHOGENESIS

Epithelial cell polarity

Much knowledge of epithelial cell polarity has been gained from the study of MDCKII cells. During epithelial morphogenesis, MDCKII cysts develop the features of epithelial architecture: 1) maintenance of cell adhesions; 2) exhibition of apicobasal polarity; and 3) orientation of mitotic spindle in the plane of epithelia (McCaffrey and Macara, 2011). In the next paragraphs, cell polarity in epithelial cells, such as MDCK cells, will be discussed briefly.

The polarity in epithelial cells is extremely complicated and intertwined with other aspects of cell biology, such as membrane traffic and cell division. A simplified diagram illustrates the main components of cell polarity in the typical epithelium (Figure 1.3). Briefly, each cell establishes apico-basolateral polarity, with the apical membrane facing the lumen and basal membrane interacting with ECM. Apical and basolateral surfaces have different compositions. Both adherens junction and tight junction are formed at the lateral membrane. Adherens junction (cadherin/catenin and nectin/afadin complexes) mediates cell-cell adhesion and tight junction (claudins, occludins, and adaptor proteins) and therefore restricts the movement of proteins and lipids between membrane domains. Both junctions serve as a barrier to seal the space between the cells (Lee and Vasioukhin, 2008; Martin-Belmonte and Perez-Moreno, 2011), which maintains

the epithelial architecture. Integrin receptors are located at the basal membrane to mediate signals from ECM (Bryant and Mostov, 2008). In addition, many other receptors that function in cell signaling as well as transporters for metabolites are localized in epithelial cells with polarity.

The main players in the regulation of apicobasal cell polarity are the Crumbs complex and Par complex at the apical membrane and Scribble complex at the basolateral membrane (Figure 1.3) (Royer and Lu, 2011). Activated by Cdc42, the Scribble complex inhibits the function of apical polarity complexes and thus maintains the identity of the basolateral membrane (Moreno-Bueno et al., 2008). The antagonizing effect also occurs as the Par complex inhibits the functions of proteins in the Scribble complex (Moreno-Bueno et al., 2008). These polarity complexes are required for correct assembly and localization of the aforementioned junction complexes (Coradini et al., 2011).

As previously mentioned, epithelial cell polarity is intertwined with other biological processes, such as polarized trafficking and cell division, which will be discussed in more detail in the next sections. The cells in mature MDCKII cysts or other epithelia display polarized cellular trafficking and polarized cytoskeleton and organelles (Rodriguez-Boulan and Macara, 2014). Importantly, the proper orientation of the mitotic spindle regulated by polarity proteins and complexes maintains the overall epithelial architecture, while the epithelial proliferation is tightly regulated (McCaffrey and Macara, 2011).

Intracellular trafficking in MDCKII epithelial morphogenesis

As previously mentioned, intracellular trafficking contributes to the establishment of epithelial cell polarity. Many studies have been dedicated to elucidating the

mechanisms of polarized trafficking using MDCKII cells. Of note, Rab11a-directed molecular network has been identified as a key player in the control of apical trafficking and initiation of the lumen, a process that marks the establishment of identities for apical and basolateral membranes, during MDCKII epithelial morphogenesis in 3D culture (Bryant et al., 2010). Through a series of gene knockdown and overexpression experiments in MDCKII 3D culture, the study revealed that Rab11a-Rab8a associated apical vesicles are responsible for trafficking apical determinants and components of apical polarity complex, including Cdc42 and Par6 (Bryant et al., 2010). The apically trafficked complex then contributes to the regulation of lumen formation and orientation of cell division, as well as maintenance of cell polarity. In another study, it has been shown that the apical targeting of Crumb3, a member of the Crumbs complex, is regulated by Rab11a associated vesicles, which occurs during the first cell division (Schluter et al., 2009). Further functional assays indicated the normal lumen formation and epithelial morphogenesis require the apical trafficking of Crumbs3 through Rab11a regulation (Schluter et al., 2009).

In addition to apical trafficking, Rab11a also plays a critical role in basolateral trafficking. Utilizing MDCK or Hela 2D cell culture with the expression of wildtype, constitutively active or dominant negative form of Rab11, it has been uncovered that E-cadherin basolateral trafficking is regulated by Rab11 vesicles in the depolarized cells (Lock and Stow, 2005). In another study, in which Rab11 mutants were utilized, it was reported that Rab11 is required for continued basolateral targeting and maintenance of E-cadherin throughout the entire process of epithelial morphogenesis in both MDCK 3D and 2D culture (Desclozeaux et al., 2008).

In addition to Rab11, several other Rab family proteins also function in regulating intracellular trafficking during epithelial morphogenesis (Homma et al., 2019). In summary, intracellular trafficking through different mechanisms not only directs *de novo* lumen formation but also contributes to the maintenance of cell polarity during epithelial morphogenesis.

Cell division and proliferation in MDCKII epithelial morphogenesis

During epithelial morphogenesis, cell proliferation and cell polarity are regulated and orchestrated to maintain the epithelial architecture. There are two key elements of cell division and proliferation during MDCKII epithelial morphogenesis. First, the proper orientation of cell division ensures the formation of monolayers of epithelial cells. Second, the tightly regulated cell proliferation rate also contributes to the normal lumenogenesis (Cerruti et al., 2013).

With correctly oriented mitotic spindle during cell division, MDCKII cells, like other epithelial cells, can proliferate and form a monolayer enclosing the central lumen during normal cystogenesis. Cdc42 regulates spindle orientation in MDCK cysts as indicated by that Cdc42 knockdown cells could disrupt spindle orientation and result in a multi-lumen phenotype (Qin et al., 2010). Further analysis using rescue and knockdown experiments with Tuba, a guanine exchange factor (GEF) of Cdc42, revealed that this protein activates Cdc42 and regulates spindle orientation (Qin et al., 2010). Using a similar experimental design, another GEF of Cdc42, ITSN2 was identified as a regulator for spindle orientation during MDCK cystogenesis (Rodriguez-Fraticelli et al., 2010). Unsurprisingly, normal lumen formation is correlated with properly oriented mitosis during cystogenesis, whereas abnormal formation could be a consequence of aberrant

regulation of spindle orientation. This was revealed by the previously mentioned studies as well as the study of cysts with knockdown of LGL, a component of machinery for controlling mitotic spindle orientation (Rodriguez-Fraticelli et al., 2010).

Cell proliferation rate decreases during MDCKII cystogenesis. A delicate *in silico* study reported that a decrease of cell proliferation occurs after lumen formation during MDCKII cystogenesis, at the stage where the cysts become "stabilized" (Engelberg et al., 2011). Quantification of EdU positive cells indicated that cell proliferation rate reaches the maximum after the cells are cultured for one day in 3D condition, which then drops significantly after three days of culture (unpublished data). The regulation of cell proliferation rate is important in normal cystogenesis as indicated by a study using mathematical models and experimental validation (Cerruti et al., 2013). Slowing down cell proliferation was shown to rescue to some extent the abnormal phenotypes of MDCK cysts (Cerruti et al., 2013).

To summarize, the regulation of cell division and proliferation are indispensable in the regulation of normal lumen formation and cell polarity during MDCKII epithelial morphogenesis.

TRANSCRIPTOME PROFILING OF MDCKII EPITHELIAL MORPHOGENESIS

Previous studies on the transcriptome of MDCKII cell models

It is intriguing to understand the mechanism by which a single lumen can be formed and maintained during MDCKII cystogenesis, where intracellular trafficking, cell polarity, and cell division are tightly regulated. The understanding of how transcriptome

changes during epithelial morphogenesis is vital in identifying key regulators and pathways that coordinate cell division and cell polarity establishment. The transcriptome data of MDCKII cell culture is accumulating in the past decade, and the related studies are summarized, including details about the culture conditions used and the technology utilized for transcriptome profiling (Table 1.1).

To date, there are several studies of MDCKII epithelial morphogenesis, in which mRNA profiles have been generated for MDCKII 3D or 2D cell culture (Galvez-Santisteban et al., 2012; Wells et al., 2013). Both studies aimed to utilize 3D culture to identify key regulators that participate in controlling epithelial morphogenesis and cystogenesis.

In the first study, the authors took the use of the fact that cells undergo polarization in both 3D and 2D cell cultures, but cells form cyst structures with central lumen only in 3D cell culture. By analyzing differentially expressed genes between 3D and 2D cell cultures using microarray data, followed by a comprehensive RNAi screening experiment and analysis of gene down-regulation in matched cancer versus normal samples, they identified a candidate key regulator Slp2-a/4-a as required for apical transport during lumen formation (Galvez-Santisteban et al., 2012).

The other study focused on mRNA profiles of mature cysts cultured for 8 days, from which interleukin-8 (IL-8) has been found essential for cystogenesis (Wells et al., 2013). Wells and colleagues performed a comparative analysis of expression profiles for fully polarized MDCKII cells cultured in 3D and 2D conditions. It is interesting that differentially expressed genes between 3D and 2D cultures are significantly enriched in pathways related to the immune and inflammatory response (Wells et al., 2013).

Furthermore, the study validated the up-regulation of IL-8, an important cytokine involved in immune response, at both the mRNA and protein levels, and revealed that IL-8 is required for cystogenesis in 3D culture by IL-8 inhibition experiments and rescue experiments (Wells et al., 2013).

HGF treatment is known to trigger partial EMT and subsequent tubulogenesis through cell migration and repolarization in MDCKII cells (O'Brien et al., 2004). HGF-induced tubule formation in MDCKII cells mimics the *in vivo* process and therefore used as a model to study tubulogenesis. Using this model, several studies investigated transcriptome changes during tubulogenesis and identified key regulators for the process (Chacon-Heszele et al., 2014; Kwon et al., 2011).

In one study, by comparing HGF treated and untreated 3D or 2D culture, it was shown that 3D and 2D culture differ in the response to HGF (Chacon-Heszele et al., 2014). Based on the data of 3D culture, this study revealed the significance of MAPK signaling pathway in controlling tubulogenesis. Importantly, a novel tubulogene, *MMP1*, was identified, which is up-regulated uniquely in 3D culture but not in 2D culture upon HGF treatment. These findings were validated using gene knockdown experiments in MDCKII cells (Chacon-Heszele et al., 2014).

Another study utilized a 2.5D co-culture system with MDCK cells grown on top of fibroblasts embedded in the collagen matrix to study time-course transcriptome changes during tubulogenesis (Kwon et al., 2011). The authors identified genes with expression changed at extension and tubule stage, respectively, compared with the initial monolayer stage. Notably, it was observed that *TNS4* is up-regulated during the

extension stage. The antibody blocking and RNAi experiments further showed the importance of *TNS4* in tubulogenesis.

Limitations of the previous studies

The transcriptome profiles of MDCKII cell models from the above studies are useful resources and the associated studies provide insightful information for studying epithelial morphogenesis. However, these transcriptome profiling studies are subject to several limitations. First, transcriptome data was collected at a single time point from the cell culture in the aforementioned studies, from which significant transcriptome changes likely remain largely uncaptured. Second, the previous studies mainly focused on a specific gene or pathway rather than providing a landscape of transcriptome changes during this process. Third, microarray experiments were performed to obtain mRNA expression data, which may result in inaccurate results due to the disadvantages of the technology, especially when applied to canine samples. For instance, microarray data have limited dynamic range, low gene coverage due to probes without accurate canine gene annotation, and batch effects that arise from different experiments, which makes integrative analysis nearly infeasible. Lastly, the models used in the analyses are subject to several limitations. Comparative analysis of 3D and 2D cell cultures could reveal the transcriptome changes unique to 3D cell culture. However, since 3D and 2D cell cultures somewhat share similar mechanisms underlying epithelial morphogenesis, the results from such analysis may be misleading. HGF-induced tubulogenesis of MDCKII cells recapitulates a complicated process that involves partial EMT, followed by a dedifferentiation process. It is not as suitable as MDCKII 3D cell culture for studying epithelial morphogenesis, which is a differentiation process.

OUR STUDY

Biological questions

In our study, we wanted to address the following questions.

- During MDCKII 3D cystogenesis, does the transcriptome gradually change over the course or switch at a certain time point?
- What could be the mechanisms underlying transcriptome changes?
- What are the similarities and differences between 3D and 2D cell culture?

We performed the following experiments and analyses to address these questions.

Transcriptome changes during MDCKII epithelial morphogenesis

Considering the limitations of previous transcriptome studies on MDCKII epithelial morphogenesis (see the previous section), we set out to collect RNA-seq samples from a full-time course of MDCKII cystogenesis using 3D culture, as well as from polarized cells using 2D culture as a comparison. With the RNA-seq samples, we revealed a qualitative transcriptome change that occurs after the first cell division and in parallel with lumen formation.

With the known functions of intracellular trafficking during MDCKII epithelial morphogenesis, especially the critical roles of Rab11-directed trafficking (see the previous sections), we wanted to investigate whether intracellular trafficking is linked to transcriptome changes during the process. Hence, we compared transcriptome changes in *DENND5A*- and *AVL9*-knockdown cells with those in wildtype cells. Our data indeed supported that intracellular trafficking likely contributes to transcriptome changes.

In addition, our study also revealed that 3D and 2D epithelial morphogenesis have largely shared transcriptome changes, but they differ in genes involved in mitochondrial functions.

Changes of open chromatin during MDCKII epithelial morphogenesis

Chromatin structure imposes obstacles on transcription mediated by RNA polymerase II and is thus linked to regulation of transcription activity (Li, Carey & Workman, 2007). Therefore, transcriptional change can be the consequence of the change in chromatin accessibility.

Chromatin accessibility assays are powerful to study the mechanisms of transcription regulation. In recent years, several methods have been developed, such as DNAnase-seq, FAIRE-seq, and ATAC-seq that directly isolate genomic locations of open chromatin regions and MNase-seq that indirectly assays for open chromatin (Tsompana and Buck, 2014). Among these methods, ATAC-seq becomes increasingly popular, with its high sensitivity and specificity, as well as a small number of cells required for probing open chromatin (Tsompana & Buck, 2014). ATAC-seq utilizes Tn5 transposase that inserts sequencing adapters into accessible chromatin regions which can be inferred by mapping of sequencing reads from the amplified libraries (Figure 1.4) (Buenrostro et al., 2015). A pilot study employed ATAC-seq to reveal genomic locations of open chromatin, nucleosome, and chromatin compaction (Buenrostro et al., 2013; Tsompana & Buck, 2014), following which ATAC-seq were routinely used for studying open chromatin and gene expression regulation in many contexts, including disease development, cell differentiation, and tissue development.

In our study, we designed ATAC-seq experiments to investigate the changes of chromatin accessibility during MDCKII epithelial morphogenesis, with samples of non-polarized cells, collected at 0h, and fully polarized cells grown in 3D or 2D culture (Figure 1.4). With the ATAC-seq data, we identified changes in open chromatin regions during epithelial morphogenesis, which could be linked to the changes observed in the transcriptome. Also, we found chromatin accessibility decreases during 3D cystogenesis.

The study will be described in more detail in the next chapters, including introduction, results, and discussions in chapter 2, supplemental information and methods in chapter 3, as well as conclusions and future directions in chapter 4.

FIGURES AND TABLES

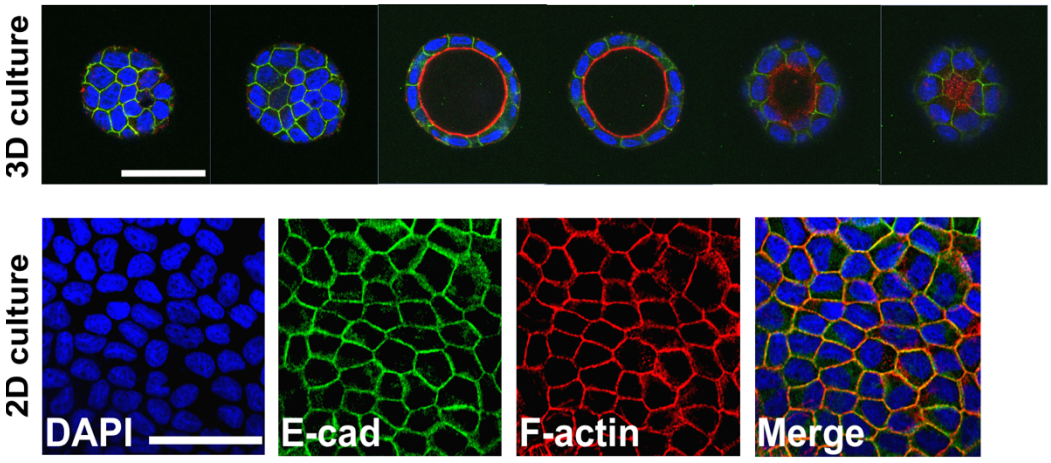


Figure 1.1. Confocal images showing representative fully polarized MDCKII cells in 3D and 2D culture. The cells were stained for nuclei, E-cadherin, and F-actin. Scale bar, 20 μm.

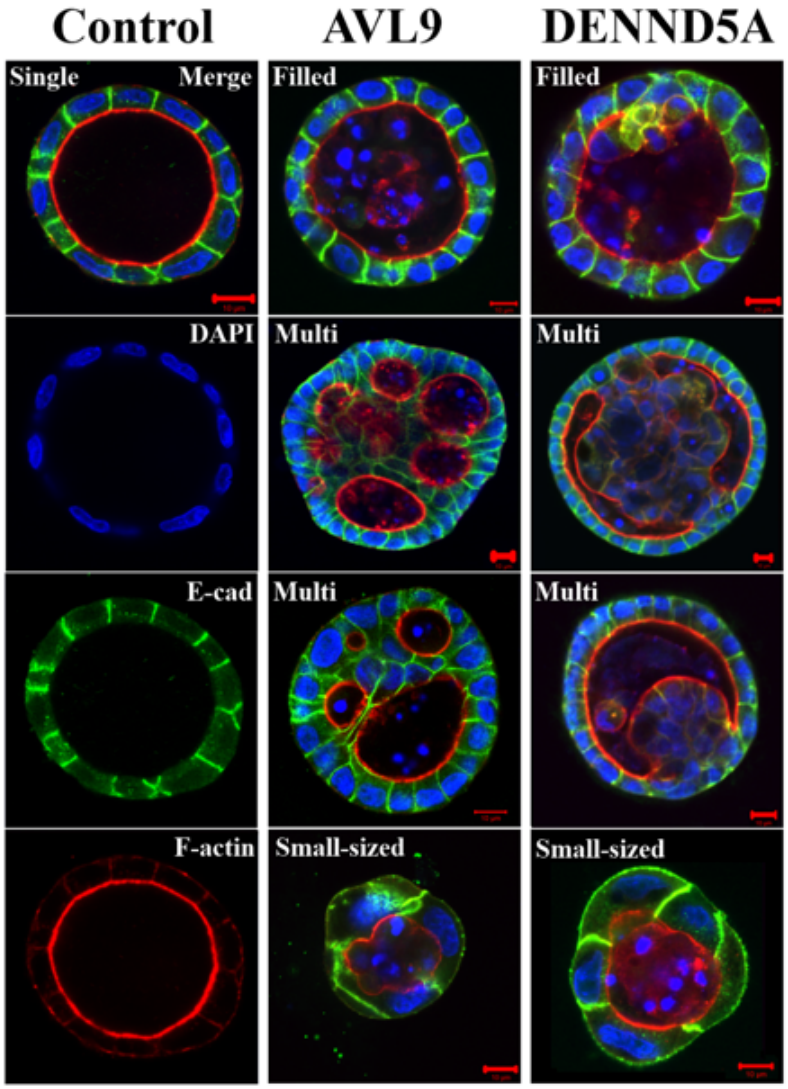


Figure 1.2. Knockdown of individual tumor suppressor significantly disrupts cell polarity and cystogenesis in MDCKII 3D culture. Representative confocal images of representative cysts of the control, and *AVL9*- or *DENND5A*- knockdown clones. Adapted from (Li et al., 2014)

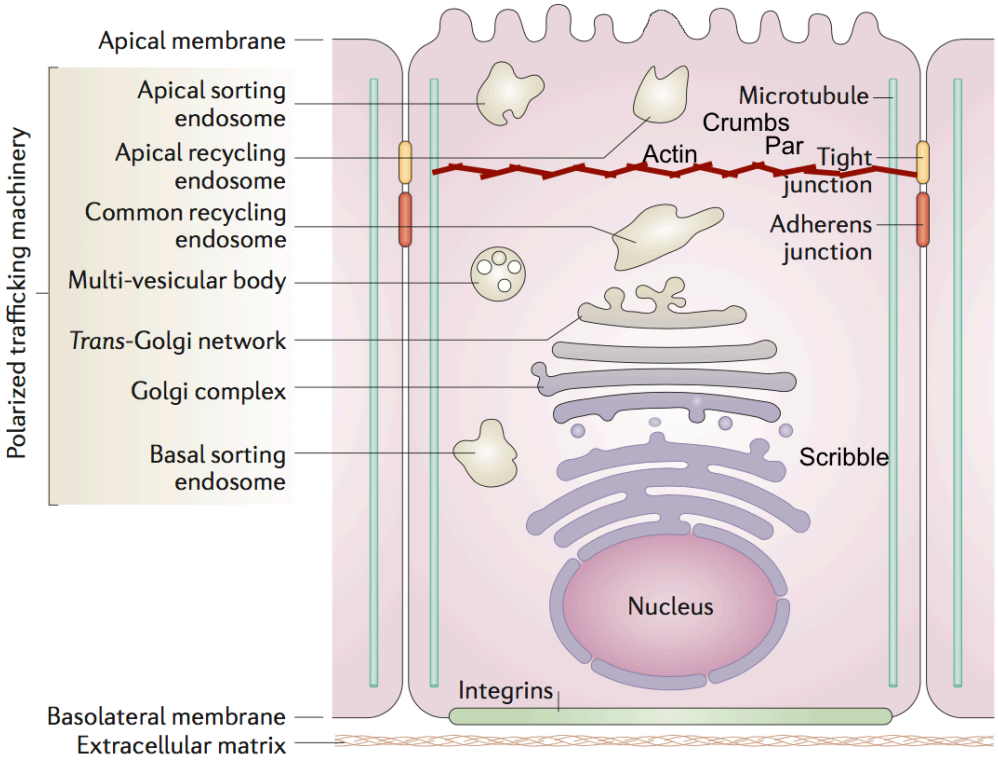


Figure 1.3. Cell polarity in typical epithelial cells. Cell polarity consists of intracellular trafficking through polarized trafficking machinery, core polarity complexes, and cell-cell junctions that interact with the cytoskeleton. The organelles are also well-organized with polarity.

Adapted from (Rodriguez-Boulan and Macara, 2014)

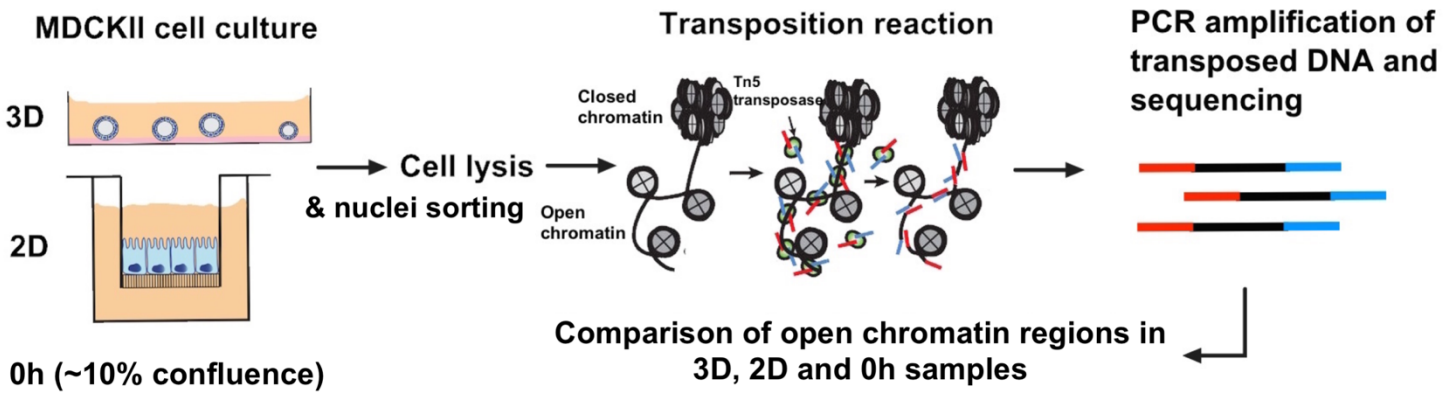


Figure 1.4. The experimental design of ATAC-seq experiments for MDCKII cells.

Table 1.1. Published transcriptome studies of MDCKII cells. Cell culture conditions and technology used for transcriptome profiling are indicated.

GEO accession no.	Cell type	Cell culture	Culture conditions	Microarray platform
GSE49518 (Chacon-Heszele et al., 2014)	MDCKII	3D	type I collagen matrix; 10 days; 0 or 24 h of HGF (100 ng/ml)	Affymetrix Canine Genome 2.0 Array
	MDCKII	2D	24-mm Transwell filter (pore size: 0.4 μm); 6 days; 0 or 24 h of HGF (100 ng/ml)	Affymetrix Canine Genome 2.0 Array
GSE32495 (Galvez-Santisteban et al., 2012)	MDCKII	3D	Matrigel; 36 hours	Affymetrix Canine Genome 2.0 Array
	MDCKII	2D	Regular P100 dish; 36 hours	Affymetrix Canine Genome 2.0 Array
GSE28381 (Wells et al., 2013)	MDCKII	3D	type I collagen matrix (embedded); 8 days	Affymetrix Canine Genome 2.0 Array
	MDCKII	2D	type I collagen matrix (on-top); 8 days	Affymetrix Canine Genome 2.0 Array

CHAPTER 2

A QUALITATIVE CHANGE IN THE TRANSCRIPTOME OCCURS AFTER THE FIRST CELL CYCLE AND COINCIDES WITH LUMEN ESTABLISHMENT DURING MDCKII CYSTOGENESIS[1]

---

[1] Wang, T., Kwon, S.-H., Peng, X., Urdy, S., Lu, Z., Schmitz, R.J., Dalton, S., Mostov, K.E., and Zhao, S. (2020). A qualitative change in the transcriptome occurs after the first cell cycle and coincides with lumen establishment during MDCKII cystogenesis. iScience, 101629.

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.

Keywords: MDCKII, epithelial morphogenesis, trafficking, transcriptome, chromatin, mitotic bookmarking, β-catenin, HNF1B, time-course RNA-seq, ATAC-seq

INTRODUCTION

Madin-Darby canine kidney (MDCK) cells represent a well-established cell line model for epithelial morphogenesis research (Datta et al., 2011; Dukes et al., 2011; Elia and Lippincott-Schwartz, 2009; McRae et al., 2017; Pillman et al., 2018; Sottocornola et al., 2010). 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 (Li et al., 2014; Tang 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.

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 hours (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 hepatocyte growth factor (HGF) in 3D, 2D or 2.5D culture conditions (Balkovetz et al., 2004; Chacon-Heszele et al., 2014; Kwon et al., 2011). 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-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 hours, and 3, 5, 8 and 14 days after seeding (Figures 2.1A and 3.1). 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 hours 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 2.1A). 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 2.1B) until they reached over-confluence, where and when apico-basolateral polarity is well-established (Figure 2.1B).

We included three biological repeats, which consist of three separate experiments starting with cells cultured at different times, 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 (Martin-Belmonte et al., 2007; Muthuswamy et al., 2001) (Figure 2.1C). Samples with expected expression patterns (e.g., higher *MYC* expression but lower *CTNNB1* expression for early time course samples, while 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 (Figure 3.1B), 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, 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 2.2A). The "early" cluster consists of all the early-stage samples, including the cells at seeding (0h) and 3D culture at 24 hours afterwards (3D24h). All other 3D samples make up the "3D" cluster, and all 2D samples constitute the "2D" cluster (Figure 2.2A). 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 2.2A). Second, unsupervised hierarchical clustering using the top-most variably expressed genes among the samples consistently identifies the same three clusters with high stability (Figures 2.2A and 3.2).

The results indicate a biphasic, rather than a linear, change among the 3D time-course samples (Figure 2.2A). 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 2.2B), namely the same early and 3D clusters as described above (Figure 2.2A).

To determine when the observed biphasic event takes place (Figure 2.2B), we identified genes that are differentially expressed (DE) between any two adjacent time points during 3D cystogenesis. Remarkably, 10 times more DE genes (>1000 each for up- and down regulation) were found between the 24h and the day 3 samples, compared to other time intervals (Figure 2.2C). Moreover, the transcriptome is largely stable before 24h and after day 3 (Figure 2.2C). As the 24h to day 3 period is when the lumen is becoming established (Figures 2.1A and 3.1), 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 2.2, we performed single sample gene set enrichment analysis (ssGSEA) with known cell polarity gene sets compiled from published studies (Assemat et al., 2008; Bryant et al., 2010; Bryant and Mostov, 2008; McCaffrey and Macara,

2012). We noted that many are significantly up-regulated in 2D cells, including genes of cell-cell junctions and cell polarity complex (Figure 2.3A). 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 and publications (Benita et al., 2009; Fevr et al., 2007; Koinuma et al., 2009; Kosinski et al., 2007; Mjelle et al., 2015; Van der Flier et al., 2007). These include cell cycle, intracellular trafficking, metabolism, redox regulation and signaling. The analysis indicates that each cluster has distinct molecular features (Figure 2.3B). 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 2.3B). 2D samples are characterized by up-regulation of trafficking-related functions and down-regulation of mitochondria-associated genes (Figure 2.3B). 3D samples feature more down-regulated functions, including glycolysis, hypoxia, TGF-β signaling, chromatin remodeling and others (Figure 2.3B). However, oxidative phosphorylation and redox-related functions are up-regulated in 3D samples.

Cell cycle genes are transcriptionally down-regulated while 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, the most prominent feature is down-regulation of cell cycle genes in both 3D and 2D samples, compared to early samples (Figure 2.4A). Specifically, ~63.5% of down-regulated genes overlap between 3D and 2D clusters (Figure 2.4B). Among them, ~50% are associated with cell cycle, many being targets of E2F and/or MYC (Figure 2.4B). These observations are consistent with the molecular characteristics of early samples (Figure 2.3B).

We identified ~1,000 up-regulated genes in each of the 3D and 2D groups (Figure 2.4A), among which 71% overlap (Figure 2.4B). 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 2.4B).

We validated several up- or down-regulated genes at the protein level. These include *MYC* (Figure 2.5A), *VIM* (encoding the mesenchymal marker vimentin) (Figure 3.4E), and *DENND5A* (encoding a Rab11-interacting intracellular trafficking protein) (Figure 3.4F).

Lastly, down-regulated genes are more likely to be located in the nucleus and to be MYC targets, compared to up-regulated genes. However, we found no differences in the fraction of genes with miRNA or transcriptional factor-binding sites (Figure 2.4C).

Our analysis indicates that during MDCKII epithelial morphogenesis, cell cycle genes are down-regulated while trafficking-related genes are up-regulated (Figures 2.4A and 2.4B). To validate this finding, we investigated relevant microarray data published previously (Chacon-Heszele et al., 2014; Galvez-Santisteban et al., 2012; Wells et al., 2013). First, we compared up- or down-regulated genes during 3D or 2D epithelial

morphogenesis from our study, indicated in Figure 2.4A, with DE genes identified between 36h (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 3.6A). Moreover, like our study (Figure 2.4A), down-regulation of cell cycle genes is also a prominent feature found for these microarray studies (Figure 3.6A). Up-regulated genes have much smaller overlaps (Figure 3.6A), but one consistency is up-regulation of trafficking-related genes in 2D condition (Figures 2.4A and 3.6A). Second, we compared our genes (Figure 2.4A) to the DE genes before and after HGF treatment from another microarray study (Chacon-Heszele et al., 2014) (Figure 3.6B). 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 3.6B). 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 2.2C), including down-regulation of cell cycle genes and up-regulation of trafficking-related genes (Figures 2.3 and 2.4), occurs in the same time frame (Figures 2.1 and 2.2) 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 2.4D), and accelerated cell proliferation after the first cell cycle when compared to 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 3.3, 3.4B-D). The results are consistent with their phenotype described above. First, a direct comparison between KD cells and wild types indicates that in KD cells, intracellular trafficking genes are significantly down-regulated, while cell proliferation genes are significantly up-regulated (Figures 3.3 and 3.4D). Second, we compared the up- and down-regulated genes identified during epithelial morphogenesis of each cell type (Figure 2.4D). Down-regulated genes unique to wild type cells are at least 3 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 2.4D). 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 2.4A) 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 2.3 and 2.4). Hence, we investigated β-catenin.

Resembling its transcript level change (Figure 3.5A), the amount of the MYC protein is significantly decreased during MDCKII epithelial morphogenesis (Figure 2.5A). Meanwhile, the amount of β-catenin remains relatively constant throughout the process (Figure 2.5A), despite differences in the levels of transcripts (Figure 2.5A). 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 2.5B), but depleted from the nucleus after the first cell cycle (Figure 2.5B). β-catenin is mostly concentrated at the cell-cell junctions after the first cell division (Figure 2.5B). This differential distribution is statistically significant after examining >40 cells in each condition (Figure 2.5B).

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 to 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 2.5B). Second, during 3D cystogenesis of *AVL9*-KD cells, the total amount of β-catenin increases (Figure 2.5A), which is accompanied by mostly constant levels of the *MYC* transcript (Figure 3.5A) and protein (Figure 2.5A). Indeed, MYC target genes are significantly up-regulated in *AVL9*-KD cells, compared to wild type cells (Figure 3.5B). Similar observations were made for *DENND5A*-KD clones (Figure 3.5C).

In an attempt to assess if WNT signaling contributes to the observed depletion of nuclear β-catenin in normal morphogenesis (Figure 2.5B), we divided WNT signaling genes into three groups: 1) WNT ligands and receptors; 2) β-catenin degradation complex; and 3) β-catenin targets. 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 2.5C). This differs from 2D epithelial morphogenesis, where WNT signaling appears to be up-regulated (Figure 2.5C). As expected, β-catenin targets (and MYC targets) are down-regulated for both 3D and 2D (Figure 2.5C). Hence, it is possible that no significant change occurs to WNT signaling during 3D cystogenesis.

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 2.6A). These are consistent with up-regulated genes unique to 3D or 2D epithelial morphogenesis (Figure 3.4A). 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 2.6B).

We also analyzed published MDCKII microarray data (Chacon-Heszele et al., 2014; Wells et al., 2013) 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 3.6C). 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 2.2C and 2.4A) 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.

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 to 0h samples (Figure 3.7A). As a comparison, the corresponding number is only ~1,000 for 2D samples (Figure 3.7A). Supporting this, differential peak analysis reveals that OCRs have shrunken or disappeared in 3D samples by ~3-fold, compared to 0h samples (Figure 2.7A). The corresponding number is significantly smaller for 2D samples (Figure 2.7A). Moreover, 3D samples harbor more *de novo* changes (i.e., newly-emerged or disappeared OCRs) than 2D samples (Figure 2.7A), 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% to 16% of OCRs are located within the promoter regions (Figure 3.7B), defined as 2kb upstream or downstream of each transcription start site (TSS) and with many validated via H3K4me3 enrichment analysis (Villar et al., 2015) (Figure 3.7B). The remaining OCRs are located in exons (15-16%), introns (34-35%) and intergenic regions (34-36%) (Figure 3.7B), with the majority within 20kb upstream or downstream of the TSS (Figure 3.7B). These numbers are similar to those reported in human ATAC-seq studies (Ackermann et al., 2016; Corces et al., 2018; Pastor et al., 2018).

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 to OCRs in 0h samples (Figure 3.7B). No such difference was observed for 2D samples (Figure 3.7B). 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 2.7A and 3.8).

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 5kb upstream or downstream of the TSS (Figure 3.7B), 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 3.7B). For OCRs that are changed during epithelial morphogenesis and within non-promoter regions, those expanded or newly-emerged contain more putative enhancers, compared to those that are shrunken or disappeared, for both 3D and 2D culture (Figures 2.7A and 3.8).

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 2.7B). 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 2.7C). The opposite was observed for SLC12A2 (Figure 2.7C), which encodes a basolateral Na^+-K^+-$2Cl^-$ cotransporter participating in cation reabsorption in kidney (Manohar and Leung, 2018). Other examples are provided in

Figure 3.9A. These observations support that chromatin remodeling underlies the observed transcript abundance changes for a significant fraction of genes (Figures 2.2C, 2.4A and 2.7B).

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, all of which have proliferative function. Among expanded OCRs, the HNF1B-binding motif is the most significantly enriched (Figures 2.7D and 3.9B). 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 3.9B). 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 2.7D and 3.9B). Importantly, a significant overlap was observed between genes with either type of enrichment (Figures 2.7D and 3.9B). *HNF1B* expression also increases during the course of epithelial morphogenesis (Figures 2.7D and 3.7E). 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).

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 24h 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 3.10). During MDCK lumenogenesis, β-catenin is likely transported to the newly formed cell-cell junctions via a Rab11a-directed network (Bryant et al., 2010; Chen et al., 1999; Lock and Stow, 2005). 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 3.10) 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 (Kosinski et al., 2007; Merlos-Suarez et al., 2011; Wang et al., 2018a; Wang et al., 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 (Di Cecilia et al., 2016; Schepers and Clevers, 2012)?

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 (Li et al., 2014; Royer and Lu, 2011; Tang et al., 2014). 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 to 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 2.4A and 2.4B). 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 2.5 and 3.7F). However, we lack experimental data to support this. Future studies, such as those that disrupts 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), indicate that mitotic bookmarking of HNF1B during the first cell division of MDCKII epithelial morphogenesis may lead to up-regulation of HNF1B target genes (Figures 2.7 and 3.9). 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 2.6). However, the mechanism of this finding remains unknown, and future studies are required.

FIGURES

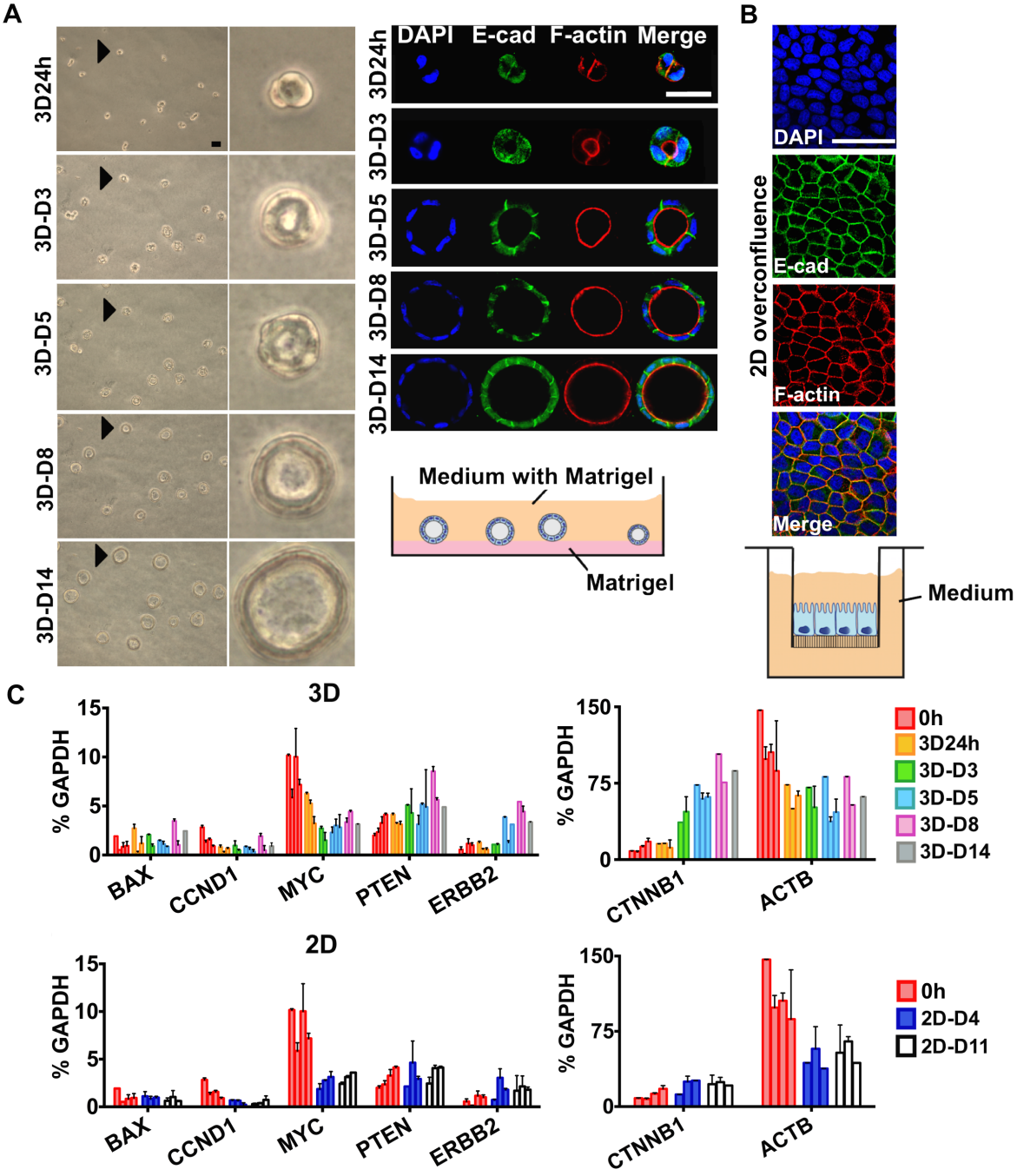

A
3D24h
3D-D3
3D-D5
3D-D8
3D-D14
DAPI
E-cad
F-actin
Merge
Medium with Matrigel
Matrigel
B
2D overconfluence
DAPI
E-cad
F-actin
Merge
Medium
C
3D
% GAPDH
BAX
CCND1
MYC
PTEN
ERBB2
CTNNB1
ACTB
0h
3D24h
3D-D3
3D-D5
3D-D8
3D-D14
2D
2D-D4
2D-D11

Figure 2.1. 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 (SD: standard deviation) with three technical replicates. "0h" represents cells at seeding, "3D24h" represents 3D culture harvested at 24 hours after seeding and so on.

See also Figure 3.1.

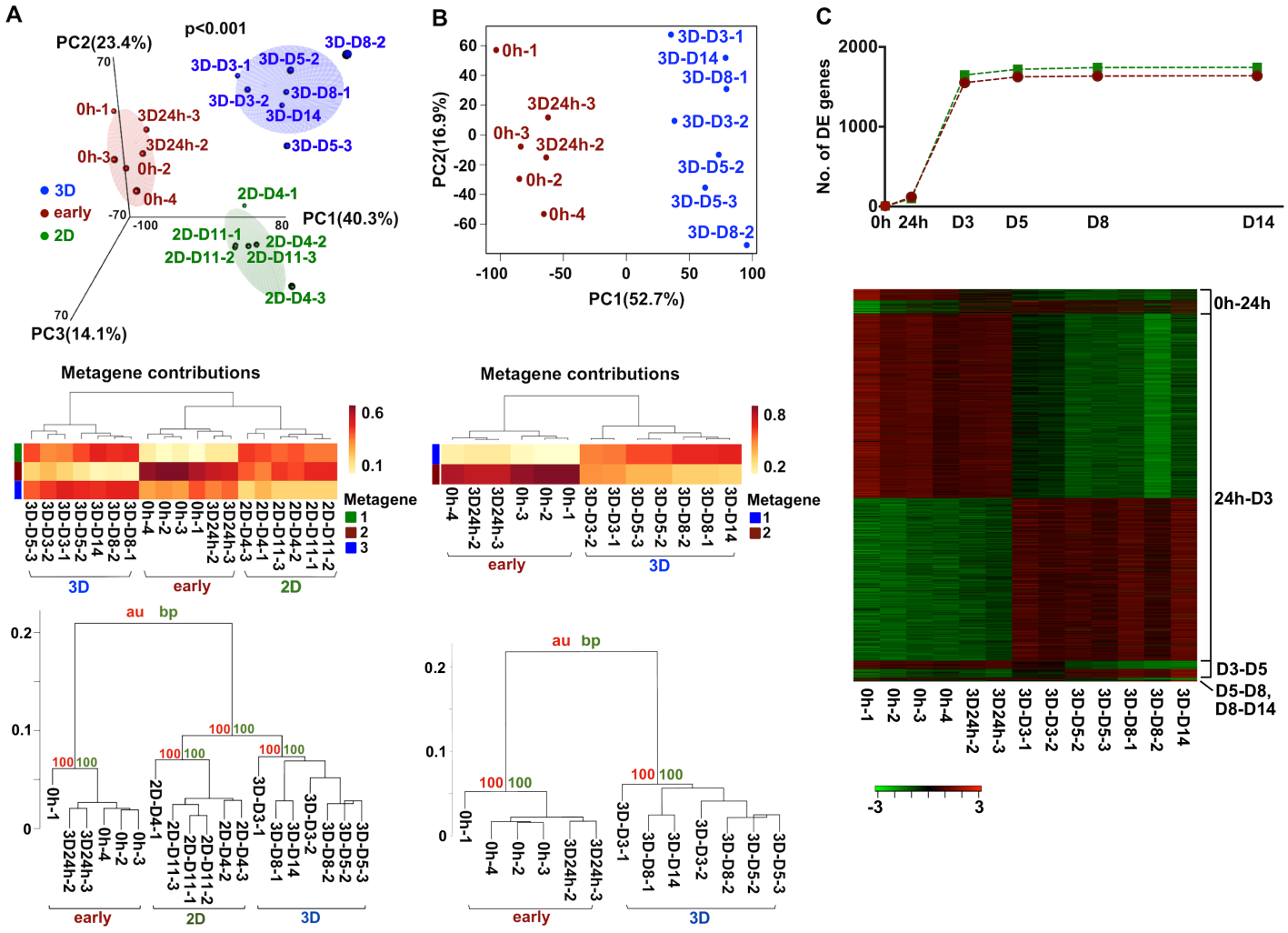

A
p<0.001
PC2(23.4%)
PC1(40.3%)
PC3(14.1%)
3D
early
2D
Metagene contributions
Metagene
1
2
3
au bp
B
PC2(16.9%)
PC1(52.7%)
C
No. of DE genes
0h 24h D3 D5 D8 D14
0h-24h
24h-D3
D3-D5
D5-D8, D8-D14

Figure 2.2. 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.

(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 hour 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, while 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 also Figure 3.2.

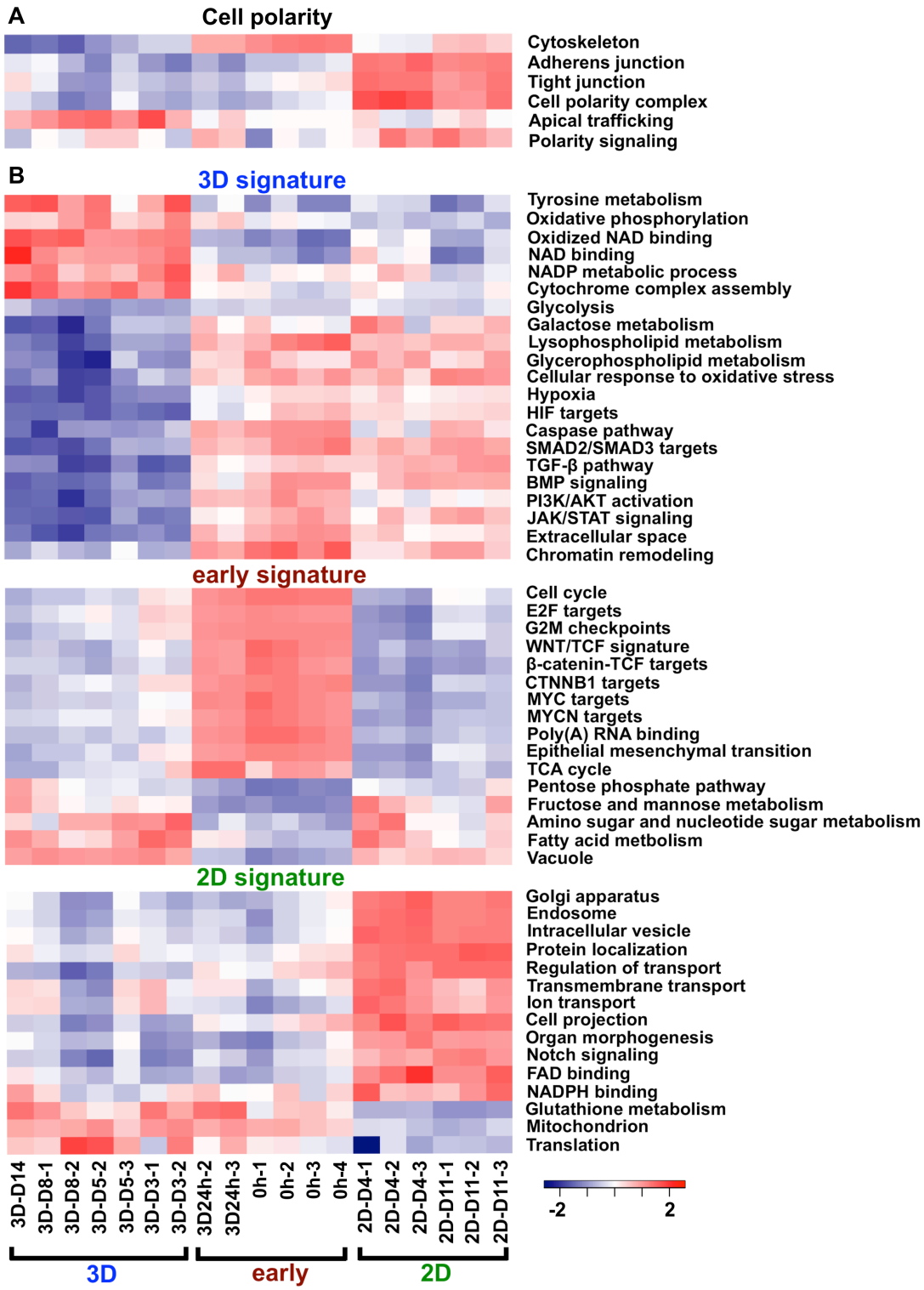

A
Cell polarity
Cytoskeleton
Adherens junction
Tight junction
Cell polarity complex
Apical trafficking
Polarity signaling
B
3D signature
Tyrosine metabolism
Oxidative phosphorylation
Oxidized NAD binding
NAD binding
NADP metabolic process
Cytochrome complex assembly
Glycolysis
Galactose metabolism
Lysophospholipid 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
early signature
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
2D signature
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
3D-D14
3D-D8-1
3D-D8-2
3D-D5-2
3D-D5-3
3D-D3-1
3D-D3-2
3D24h-2
3D24h-3
0h-1
0h-2
0h-3
0h-4
2D-D4-1
2D-D4-2
2D-D4-3
2D-D11-1
2D-D11-2
2D-D11-3
-2
2
3D
early
2D

Figure 2.3. 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 to 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 $p \leq 0.05$ by Wilcoxon tests.

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

See also Figure 3.3.

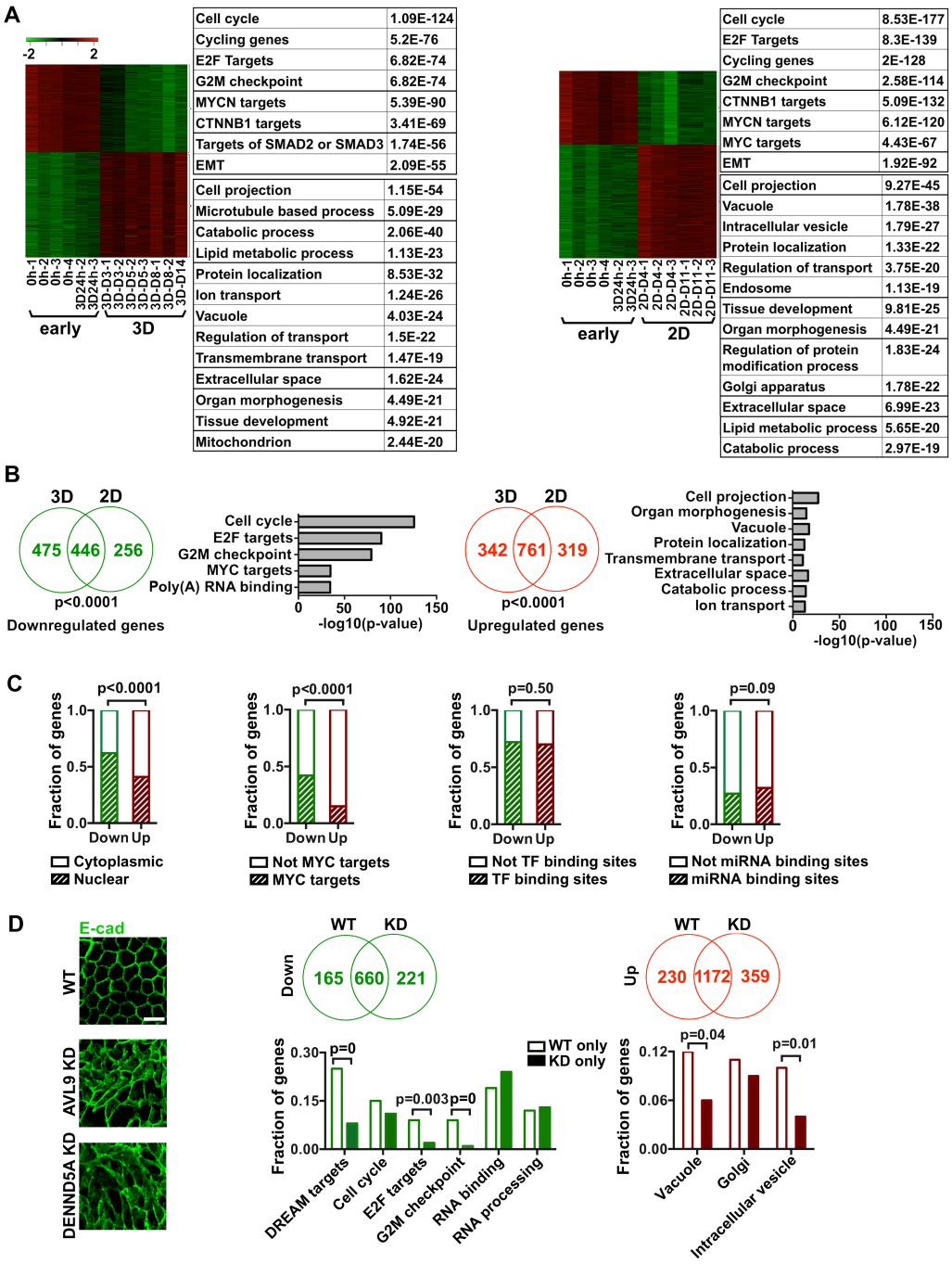

A
-2
2
0h-1
0h-2
0h-3
0h-4
3D24h-2
3D24h-3
3D-D3-1
3D-D3-2
3D-D5-2
3D-D5-3
3D-D8-1
3D-D8-2
3D-D14
early
3D
Cell cycle 1.09E-124
Cycling genes 5.2E-76
E2F Targets 6.82E-74
G2M checkpoint 6.82E-74
MYCN targets 5.39E-90
CTNNB1 targets 3.41E-69
Targets of SMAD2 or SMAD3 1.74E-56
EMT 2.09E-55
Cell projection 1.15E-54
Microtubule based process 5.09E-29
Catabolic process 2.06E-40
Lipid metabolic process 1.13E-23
Protein localization 8.53E-32
Ion transport 1.24E-26
Vacuole 4.03E-24
Regulation of transport 1.5E-22
Transmembrane transport 1.47E-19
Extracellular space 1.62E-24
Organ morphogenesis 4.49E-21
Tissue development 4.92E-21
Mitochondrion 2.44E-20
0h-1
0h-2
0h-3
0h-4
3D24h-2
3D24h-3
2D-D4-1
2D-D4-2
2D-D4-3
2D-D11-1
2D-D11-2
2D-D11-3
early
2D
Cell cycle 8.53E-177
E2F Targets 8.3E-139
Cycling genes 2E-128
G2M checkpoint 2.58E-114
CTNNB1 targets 5.09E-132
MYCN targets 6.12E-120
MYC targets 4.43E-67
EMT 1.92E-92
Cell projection 9.27E-45
Vacuole 1.78E-38
Intracellular vesicle 1.79E-27
Protein localization 1.33E-22
Regulation of transport 3.75E-20
Endosome 1.13E-19
Tissue development 9.81E-25
Organ morphogenesis 4.49E-21
Regulation of protein modification process 1.83E-24
Golgi apparatus 1.78E-22
Extracellular space 6.99E-23
Lipid metabolic process 5.65E-20
Catabolic process 2.97E-19
B
3D
2D
475
446
256
p<0.0001
Downregulated genes
Cell cycle
E2F targets
G2M checkpoint
MYC targets
Poly(A) RNA binding
0
50
100
150
-log10(p-value)
3D
2D
342
761
319
p<0.0001
Upregulated genes
Cell projection
Organ morphogenesis
Vacuole
Protein localization
Transmembrane transport
Extracellular space
Catabolic process
Ion transport
0
50
100
150
-log10(p-value)
C
p<0.0001
Fraction of genes
1.0
0.5
0.0
Down Up
Cytoplasmic
Nuclear
p<0.0001
Fraction of genes
1.0
0.5
0.0
Down Up
Not MYC targets
MYC targets
p=0.50
Fraction of genes
1.0
0.5
0.0
Down Up
Not TF binding sites
TF binding sites
p=0.09
Fraction of genes
1.0
0.5
0.0
Down Up
Not miRNA binding sites
miRNA binding sites
D
E-cad
WT
AVL9 KD
DENND5A KD
WT
KD
Down
165
660
221
WT
KD
Up
230
1172
359
WT only
KD only
Fraction of genes
0.30
0.15
0.00
p=0
p=0.003
p=0
DREAM targets
Cell cycle
E2F targets
G2M checkpoint
RNA binding
RNA processing
Fraction of genes
0.12
0.06
0.00
p=0.04
p=0.01
Vacuole
Golgi
Intracellular vesicle

Figure 2.4. Cell cycle genes are transcriptionally down-regulated while 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, 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 panel B) are significantly more enriched in nuclear location and MYC targets, but are similarly enriched in miRNA and transcriptional factor (TF) binding sites, compared to up-regulated genes. P-values were calculated with Fisher's exact tests.

(D) During epithelial morphogenesis, wild type (WT) cells have more cell cycle genes down-regulated and more intracellular trafficking genes up-regulated, compared to *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 also Figure 3.4.

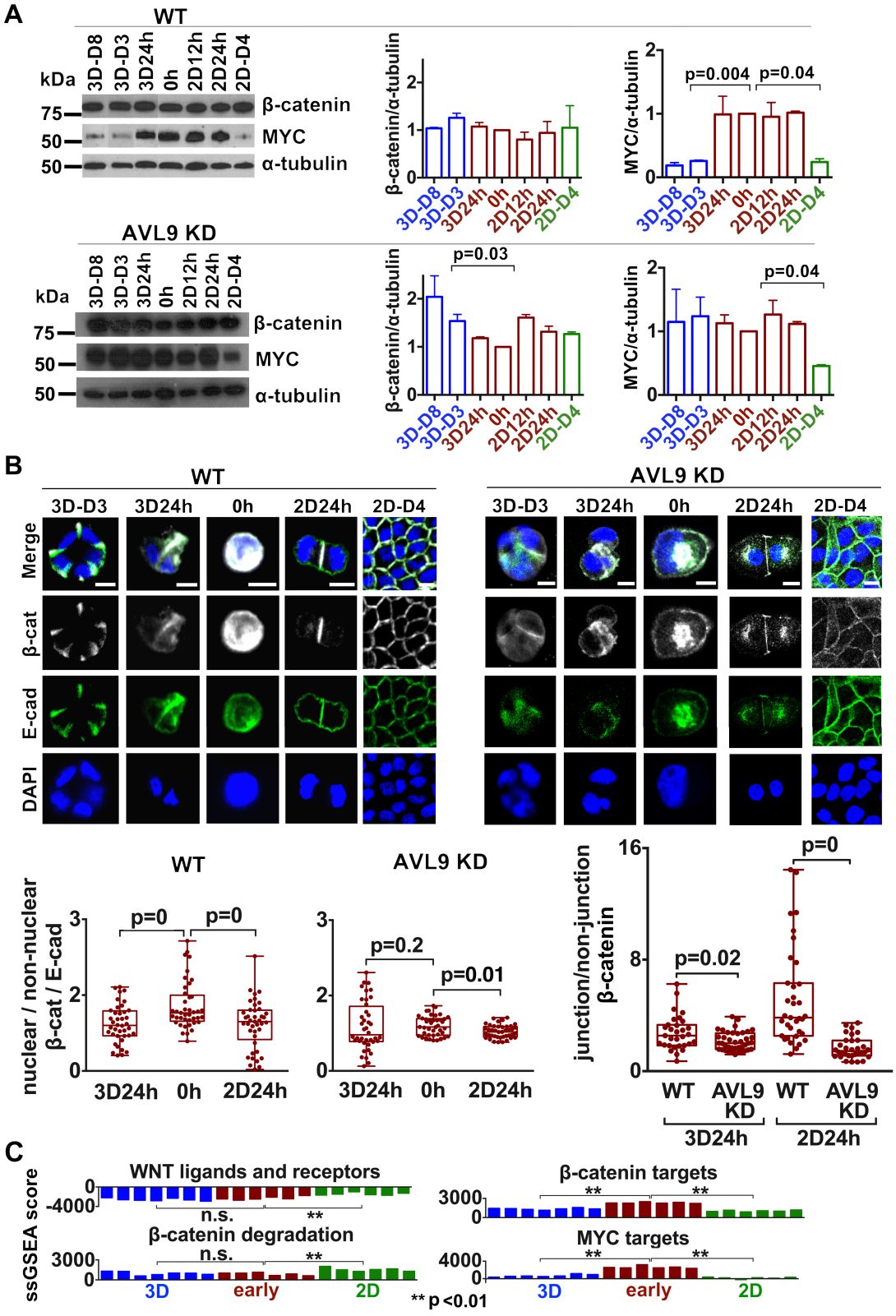

A
WT
kDa
3D-D8
3D-D3
3D24h
0h
2D12h
2D24h
2D-D4
75
50
50
β-catenin
MYC
α-tubulin
β-catenin/α-tubulin
MYC/α-tubulin
p=0.004
p=0.04
AVL9 KD
p=0.03
p=0.04
B
WT
AVL9 KD
3D-D3
3D24h
0h
2D24h
2D-D4
Merge
β-cat
E-cad
DAPI
nuclear / non-nuclear
β-cat / E-cad
p=0
p=0
p=0.2
p=0.01
junction/non-junction
β-catenin
p=0.02
p=0
WT
AVL9 KD
3D24h
2D24h
C
ssGSEA score
WNT ligands and receptors
β-catenin degradation
β-catenin targets
MYC targets
n.s.
**
3D
early
2D
** p <0.01

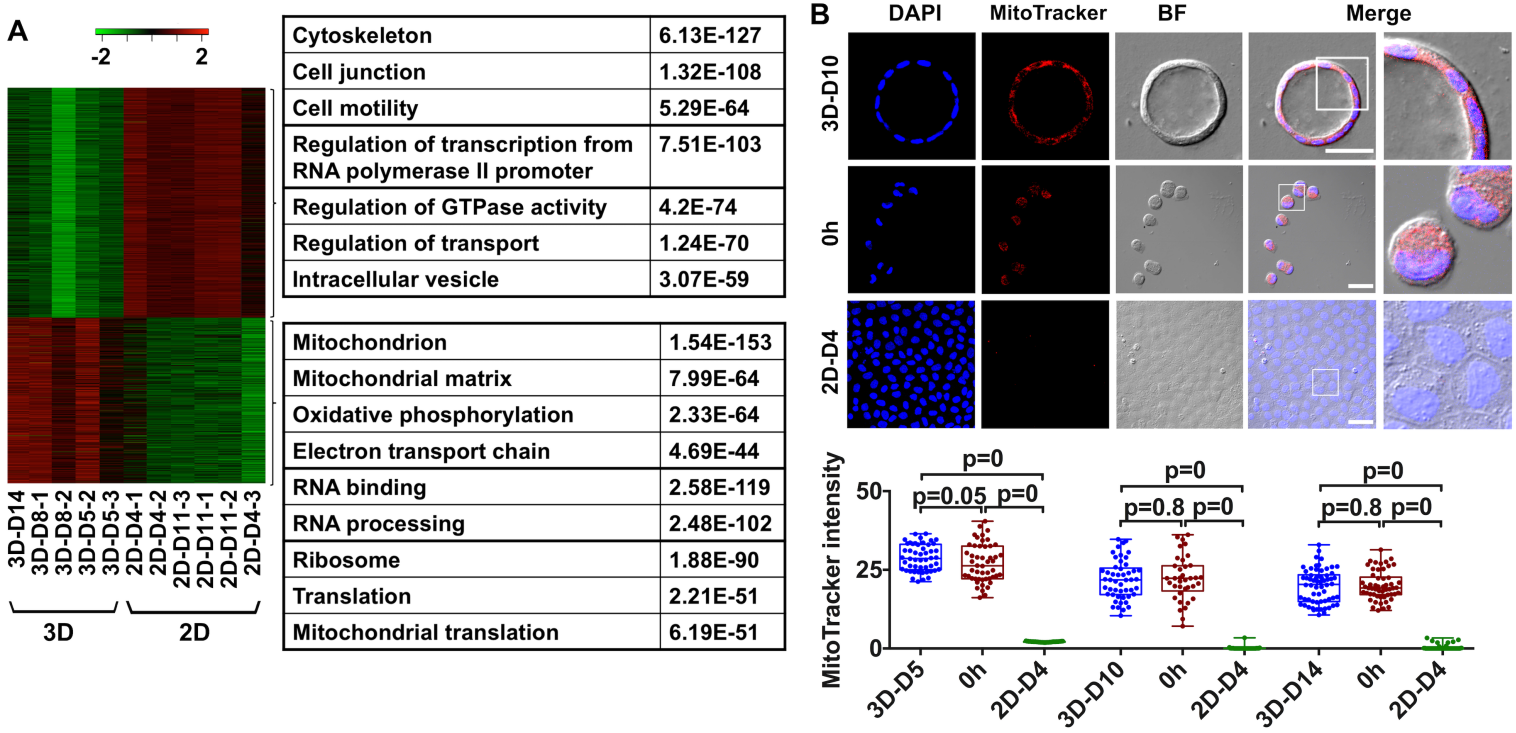
Figure 2.5. β-catenin is depleted in the nucleus and enriched at the cell-cell junctions, while MYC is down-regulated, upon the completion of the first cell cycle.

(A) In WT cells, β-catenin stays constant while MYC is significantly reduced in both 3D and 2D samples, compared to the early group. In *AVL9*-KD cells, β-catenin increases significantly while 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 3.5.



Cytoskeleton	6.13E-127
Cell junction	1.32E-108
Cell motility	5.29E-64
Regulation of transcription from RNA polymerase II promoter	7.51E-103
Regulation of GTPase activity	4.2E-74
Regulation of transport	1.24E-70
Intracellular vesicle	3.07E-59

Mitochondrion	1.54E-153
Mitochondrial matrix	7.99E-64
Oxidative phosphorylation	2.33E-64
Electron transport chain	4.69E-44
RNA binding	2.58E-119
RNA processing	2.48E-102
Ribosome	1.88E-90
Translation	2.21E-51
Mitochondrial translation	6.19E-51

Figure 2.6. 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 3.6.

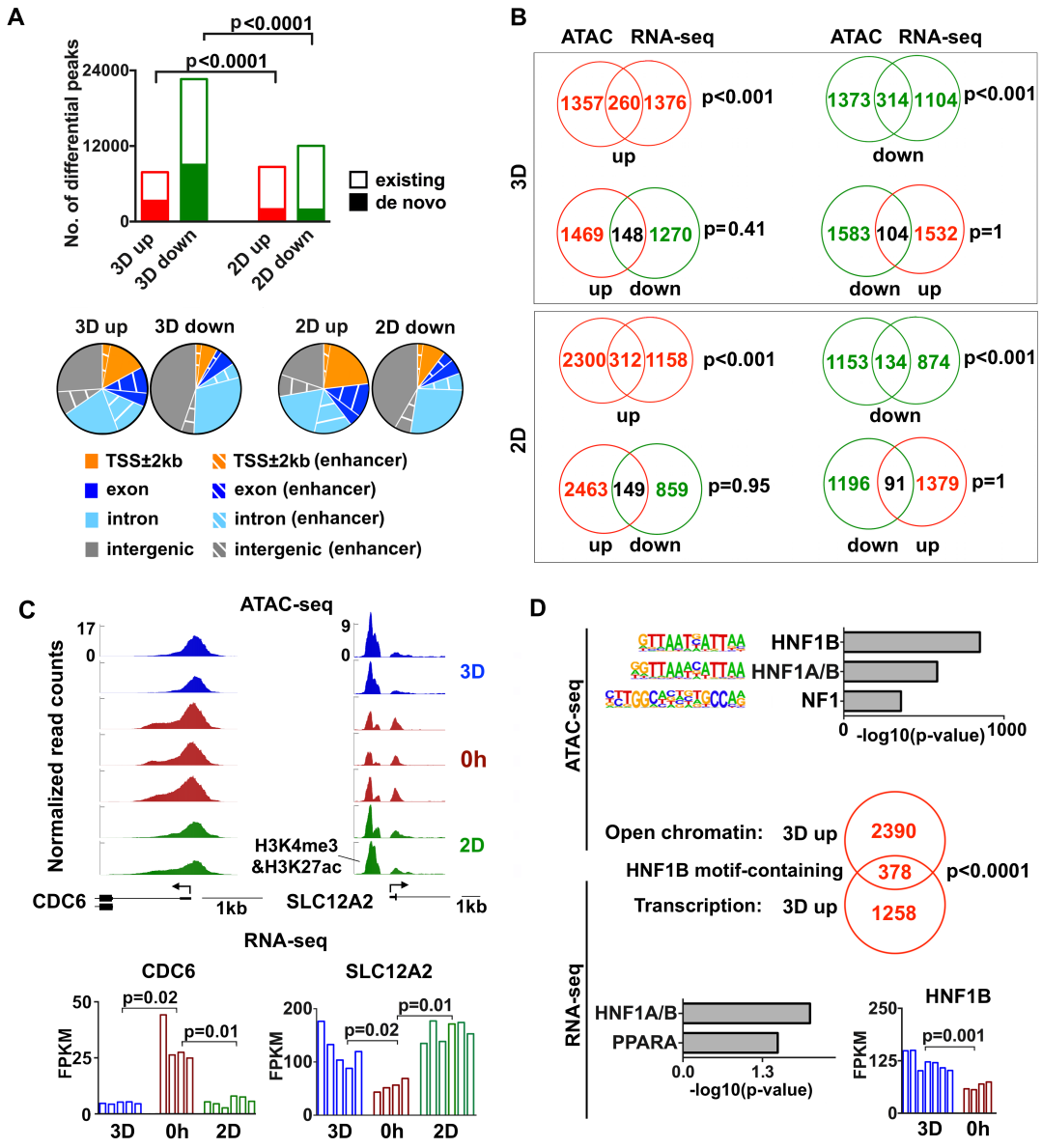


Figure 2.7. Transcriptomic changes during MDCKII epithelial morphogenesis are associated with chromatin remodeling.

(A) Decrease in chromatin accessibility is more extensive during 3D epithelial morphogenesis, compared to 2D. The bar plot (top) indicates increased/decreased (existing) or emerged/disappeared (*de novo*) ATAC-seq peaks in 3D and 2D samples, when compared to 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 to those 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, while SLC12A2 is a target of HNF1B.

(D) HNF1B binding sites are among the top-most 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-regulatedgenes and genes that harbor expanded/emerged and HNF1B motif-containing ATAC-seq peaks in their gene body and 20kb flanking regions. P-values were calculated with permutation tests. Bar plot shows up-regulation of HNF1B.

See also Figures 3.7-3.10.

CHAPTER 3

SUPPLEMENTAL FIGURES AND METHODS

SUPPLEMENTAL FIGURES

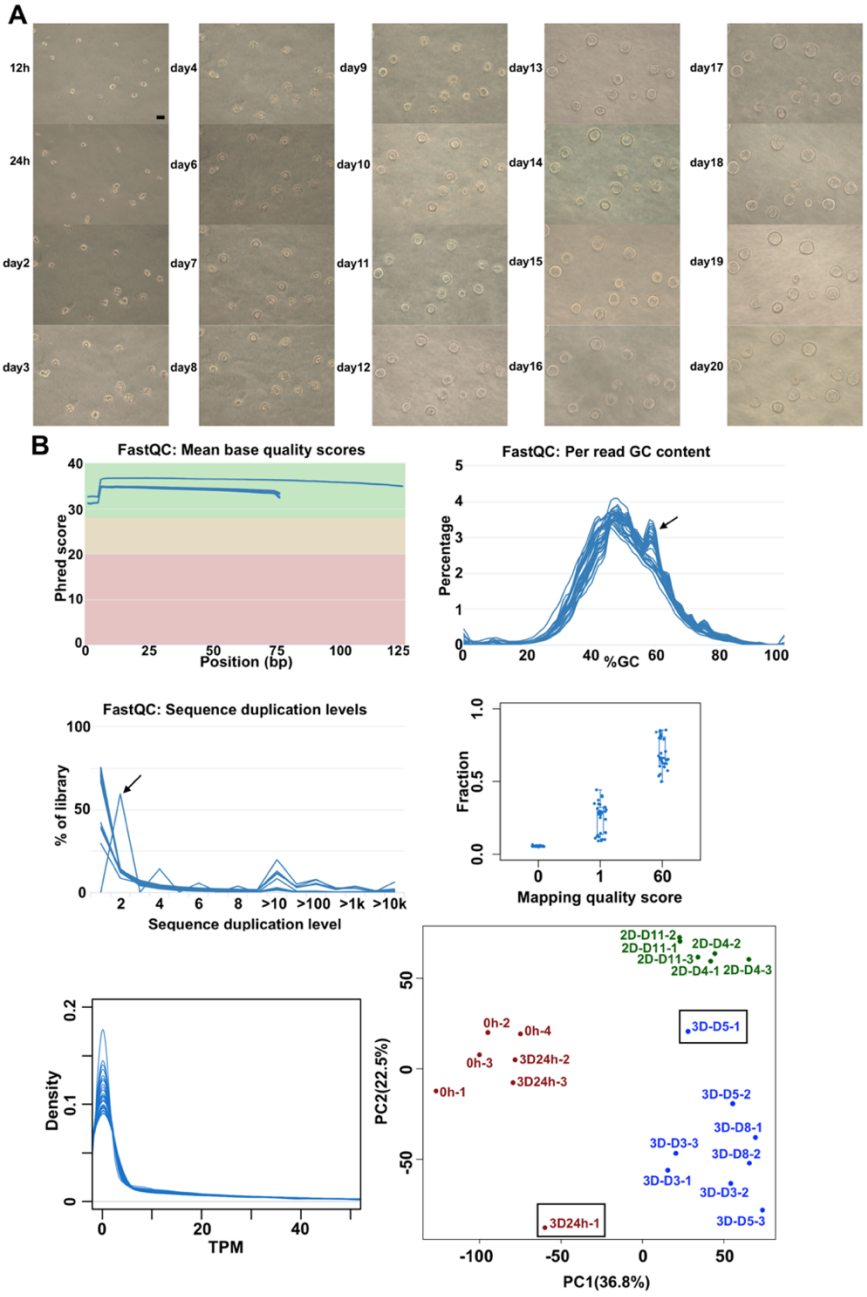


Figure 3.1. Time-course RNA-seq experiments. Related to Figure 2.1.

(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.

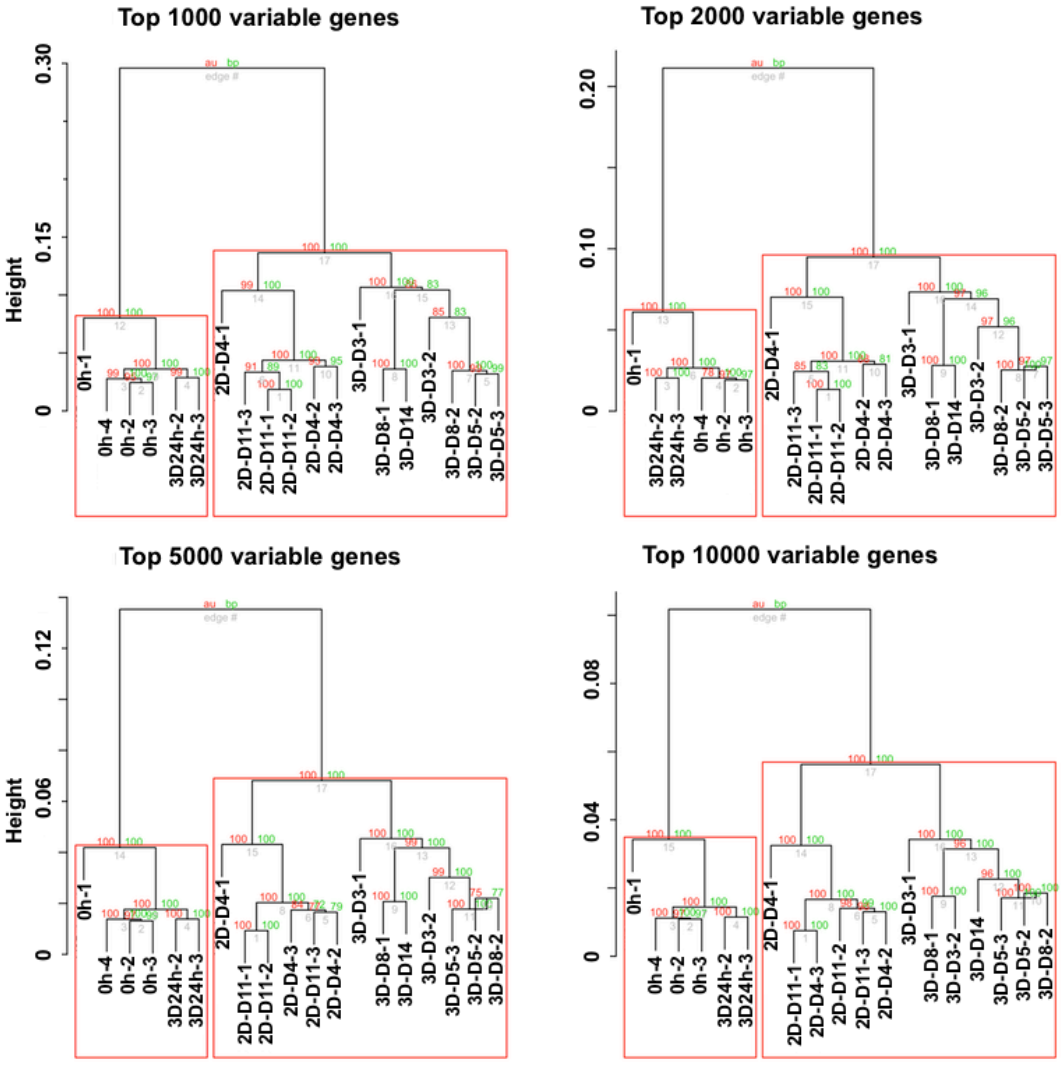


Figure 3.2. Hierarchical clustering with top variably expressed genes across samples supports three sample clusters (3D, early and 2D). Related to Figure 2.2. Bootstrapping was performed with permutations. Approximately unbiased (au) and bootstrap probability (bp) values are shown in red and green, respectively.

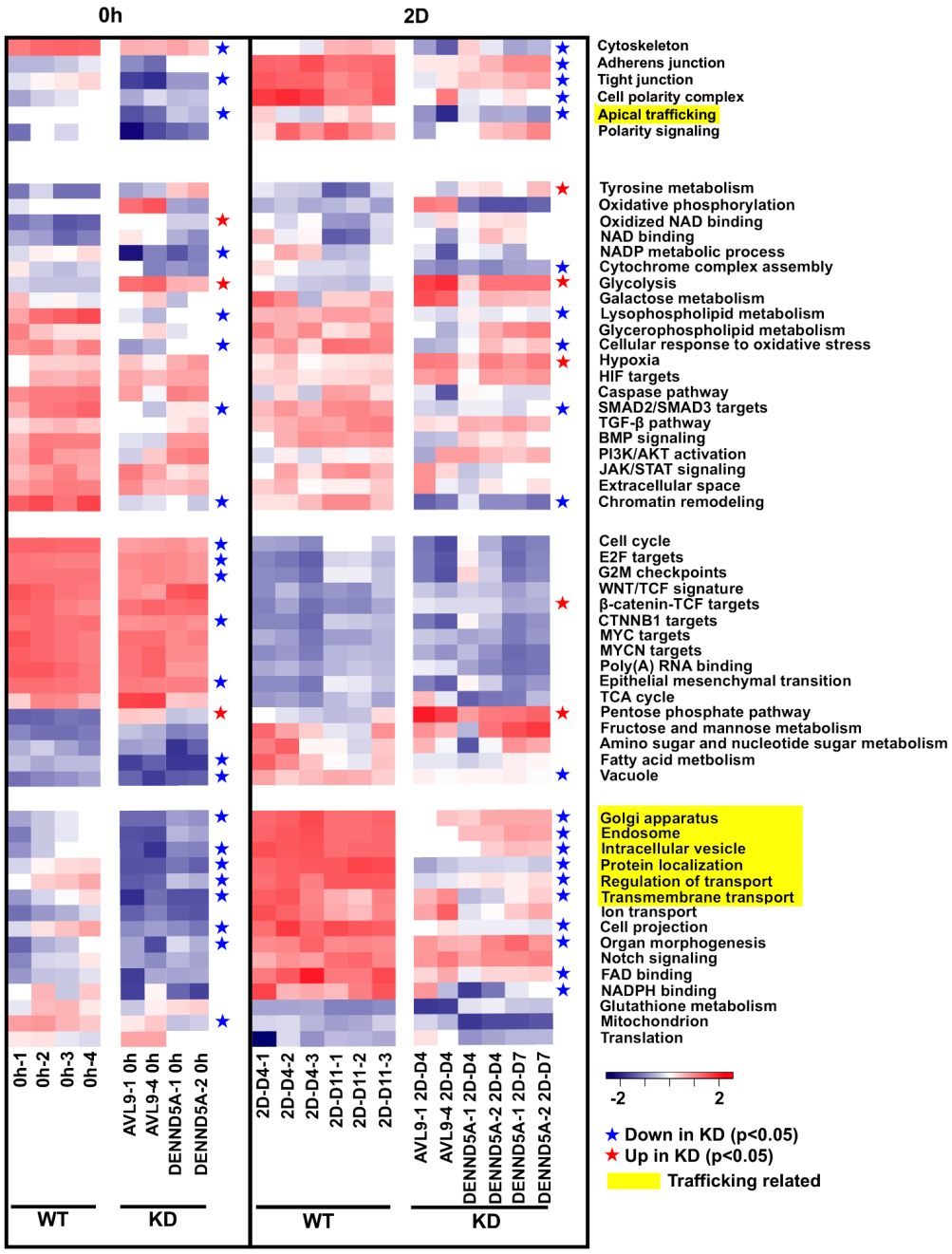


Figure 3.3. The molecular characterization of MDCKII WT, *AVL9*- and *DENND5A*-KD cells with RNA-seq data. Related to Figures 2.3 and 2.6.

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 same gene sets are used as in Figure 2.3.

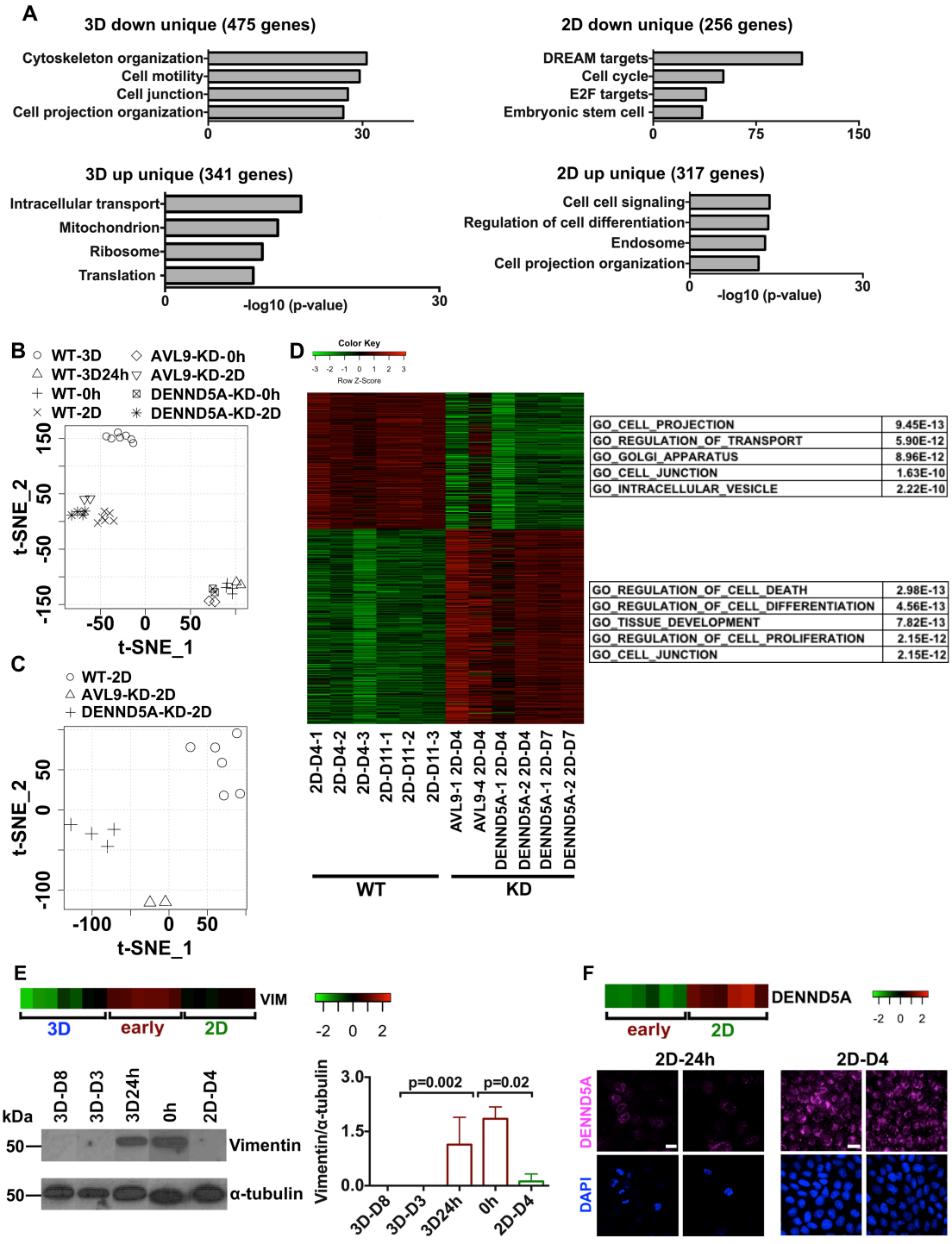

A
3D down unique (475 genes)
Cytoskeleton organization
Cell motility
Cell junction
Cell projection organization
0
30
2D down unique (256 genes)
DREAM targets
Cell cycle
E2F targets
Embryonic stem cell
0
75
150
3D up unique (341 genes)
Intracellular transport
Mitochondrion
Ribosome
Translation
0
30
-log10 (p-value)
2D up unique (317 genes)
Cell cell signaling
Regulation of cell differentiation
Endosome
Cell projection organization
0
30
-log10 (p-value)
B
WT-3D
WT-3D24h
WT-0h
WT-2D
AVL9-KD-0h
AVL9-KD-2D
DENND5A-KD-0h
DENND5A-KD-2D
t-SNE_2
t-SNE_1
C
WT-2D
AVL9-KD-2D
DENND5A-KD-2D
t-SNE_2
t-SNE_1
D
Color Key
Row Z-Score
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
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
E
VIM
3D
early
2D
kDa
3D-D8
3D-D3
3D24h
0h
2D-D4
50
Vimentin
50
α-tubulin
Vimentin/α-tubulin
p=0.002
p=0.02
F
DENND5A
early
2D
2D-24h
2D-D4
DENND5A
DAPI

Figure 3.4. Comparison between MDCKII WT and KD cells indicates the down-regulation of intracellular trafficking genes and up-regulation of cell proliferation genes in KD cells. Related to Figure 2.4.

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

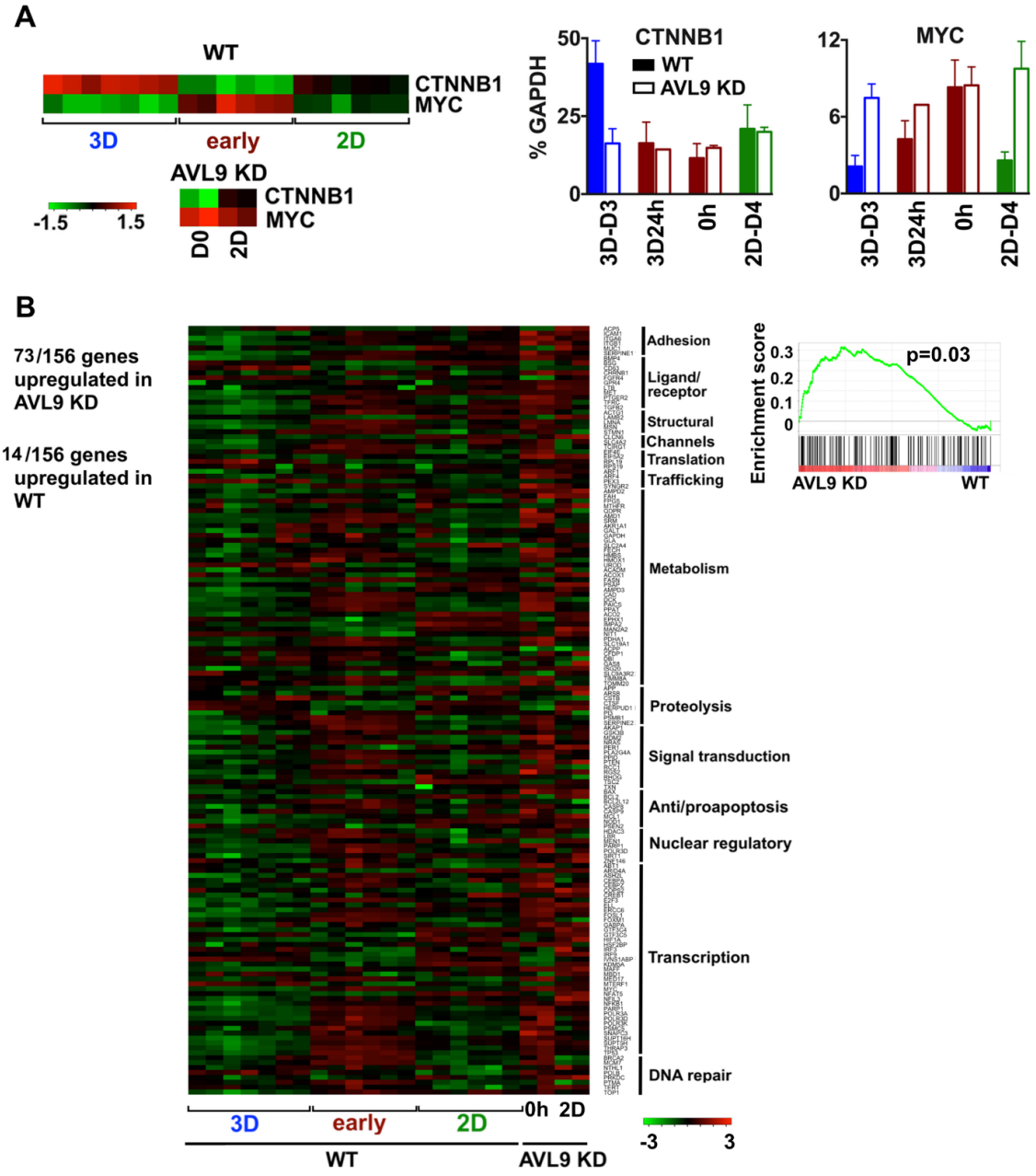
(B) t-SNE analysis (11,090 genes) of MDCKII WT and KD samples shows three clusters: 3D, 2D and early.

(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 down-regulation 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 up-regulation 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
AVL9 KD
CTNNB1
MYC
-1.5
1.5
D0
2D
CTNNB1
WT
AVL9 KD
% GAPDH
50
25
0
3D-D3
3D24h
0h
2D-D4
MYC
12
6
0
3D-D3
3D24h
0h
2D-D4
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
Enrichment score
0.3
0.2
0.1
0
p=0.03
AVL9 KD
WT
3D
early
2D
0h 2D
WT
AVL9 KD
-3
3

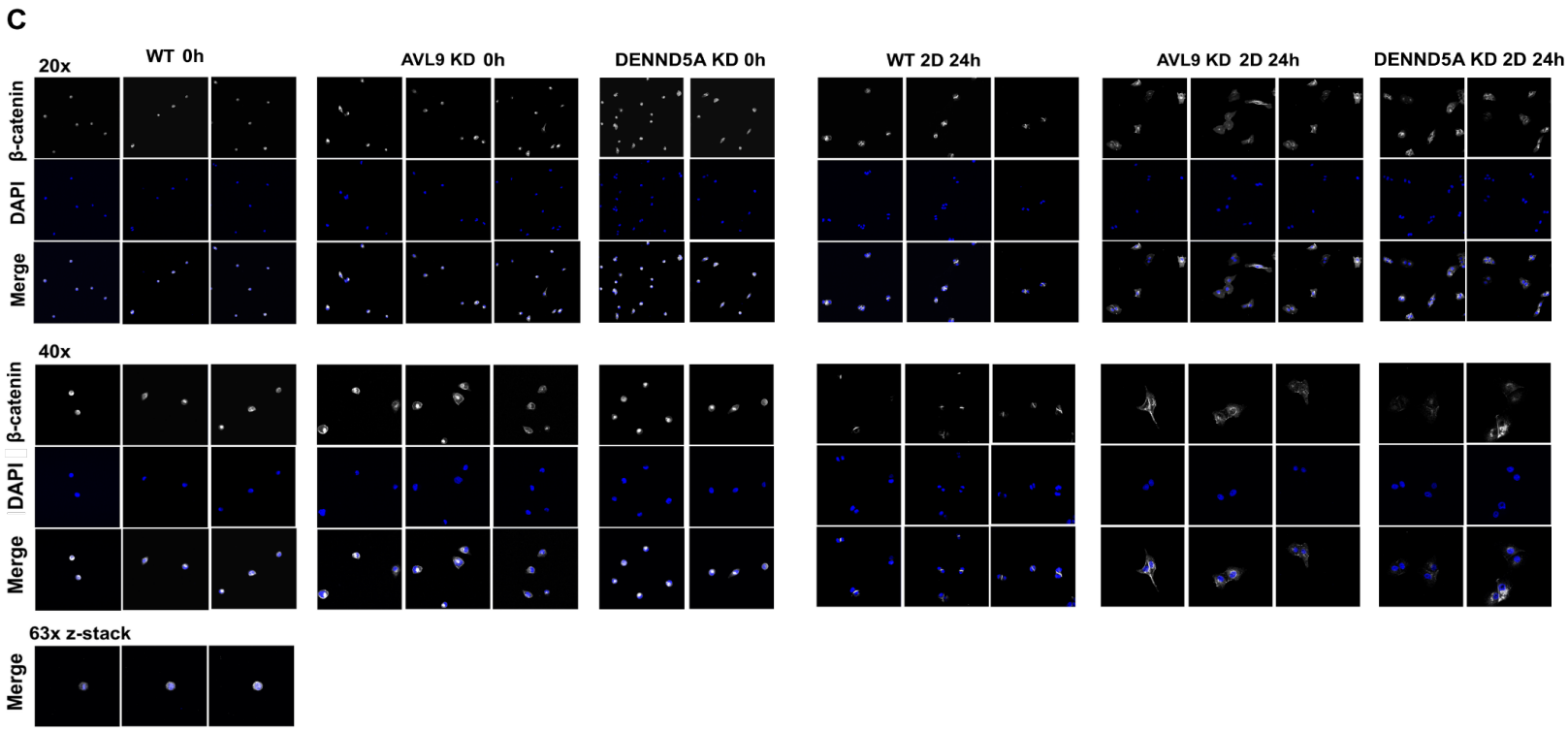
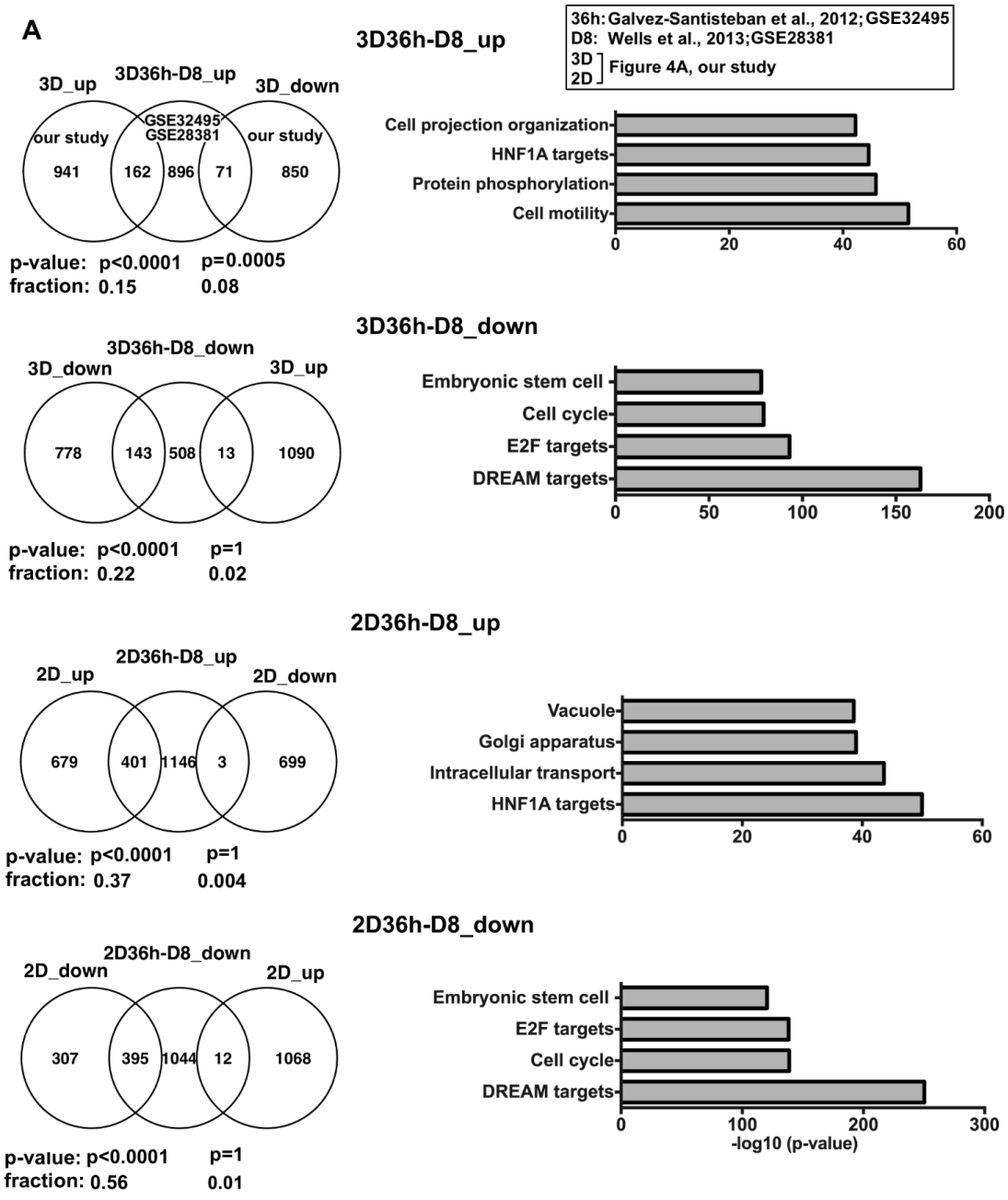

C
20x
WT 0h
AVL9 KD 0h
DENND5A KD 0h
WT 2D 24h
AVL9 KD 2D 24h
DENND5A KD 2D 24h
β-catenin
DAPI
Merge
40x
β-catenin
DAPI
Merge
63x z-stack
Merge

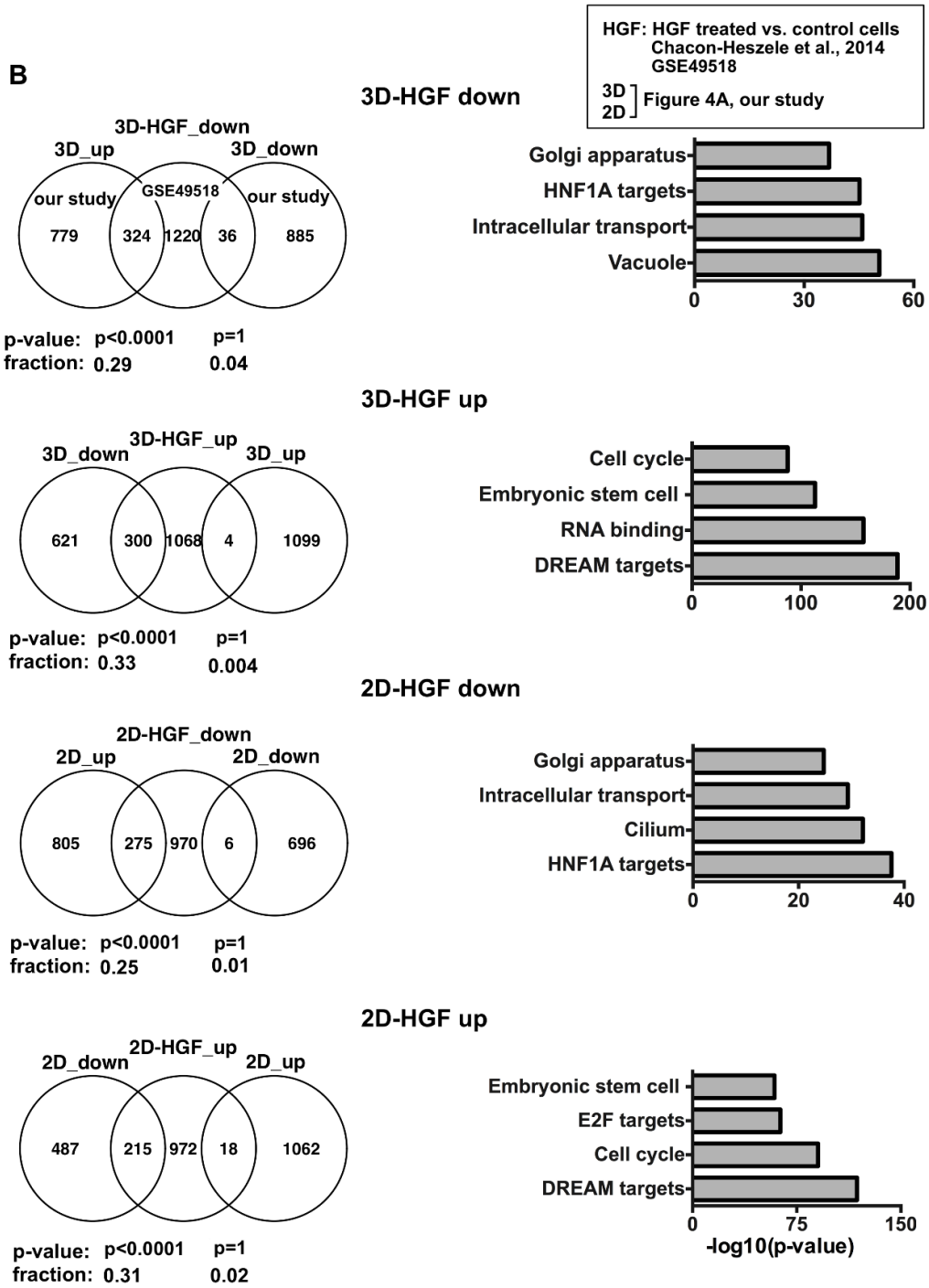
Figure 3.5. *MYC* target genes are more abundantly expressed during epithelial morphogenesis of *AVL9*-KD cells, with β-catenin detected in the nucleus and cytoplasm after the first cell cycle, unlike WT cells. Related to Figure 2.5.

(A) At the transcript level, CTNNB1 is up-regulated while MYC is down-regulated 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 of *AVL9*-KD cells compared to WT cells, calculated from GSEA.

(G) 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
3D_up
3D36h-D8_up
3D_down
our study
GSE32495
GSE28381
our study
941
162
896
71
850
p-value: p<0.0001 p=0.0005
fraction: 0.15 0.08
3D36h-D8_up
36h: Galvez-Santisteban et al., 2012; GSE32495
D8: Wells et al., 2013;GSE28381
3D
2D
Figure 4A, our study
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
3D36h-D8_down
Embryonic stem cell
Cell cycle
E2F targets
DREAM targets
0
50
100
150
200
2D_up
2D36h-D8_up
2D_down
679
401
1146
3
699
p-value: p<0.0001 p=1
fraction: 0.37 0.004
2D36h-D8_up
Vacuole
Golgi apparatus
Intracellular transport
HNF1A targets
0
20
40
60
2D_down
2D36h-D8_down
2D_up
307
395
1044
12
1068
p-value: p<0.0001 p=1
fraction: 0.56 0.01
2D36h-D8_down
Embryonic stem cell
E2F targets
Cell cycle
DREAM targets
0
100
200
300
-log10 (p-value)


B
3D-HGF_down
3D_up
3D_down
our study
GSE49518
our study
779
324
1220
36
885
p-value: p<0.0001 p=1
fraction: 0.29 0.04
3D-HGF down
HGF: HGF treated vs. control cells
Chacon-Heszele et al., 2014
GSE49518
3D
2D
Figure 4A, our study
Golgi apparatus
HNF1A targets
Intracellular transport
Vacuole
0
30
60
3D-HGF_up
3D_down
3D_up
621
300
1068
4
1099
p-value: p<0.0001 p=1
fraction: 0.33 0.004
3D-HGF up
Cell cycle
Embryonic stem cell
RNA binding
DREAM targets
0
100
200
2D-HGF_down
2D_up
2D_down
805
275
970
6
696
p-value: p<0.0001 p=1
fraction: 0.25 0.01
2D-HGF down
Golgi apparatus
Intracellular transport
Cilium
HNF1A targets
0
20
40
2D-HGF_up
2D_down
2D_up
487
215
972
18
1062
p-value: p<0.0001 p=1
fraction: 0.31 0.02
2D-HGF up
Embryonic stem cell
E2F targets
Cell cycle
DREAM targets
0
75
150
-log10(p-value)

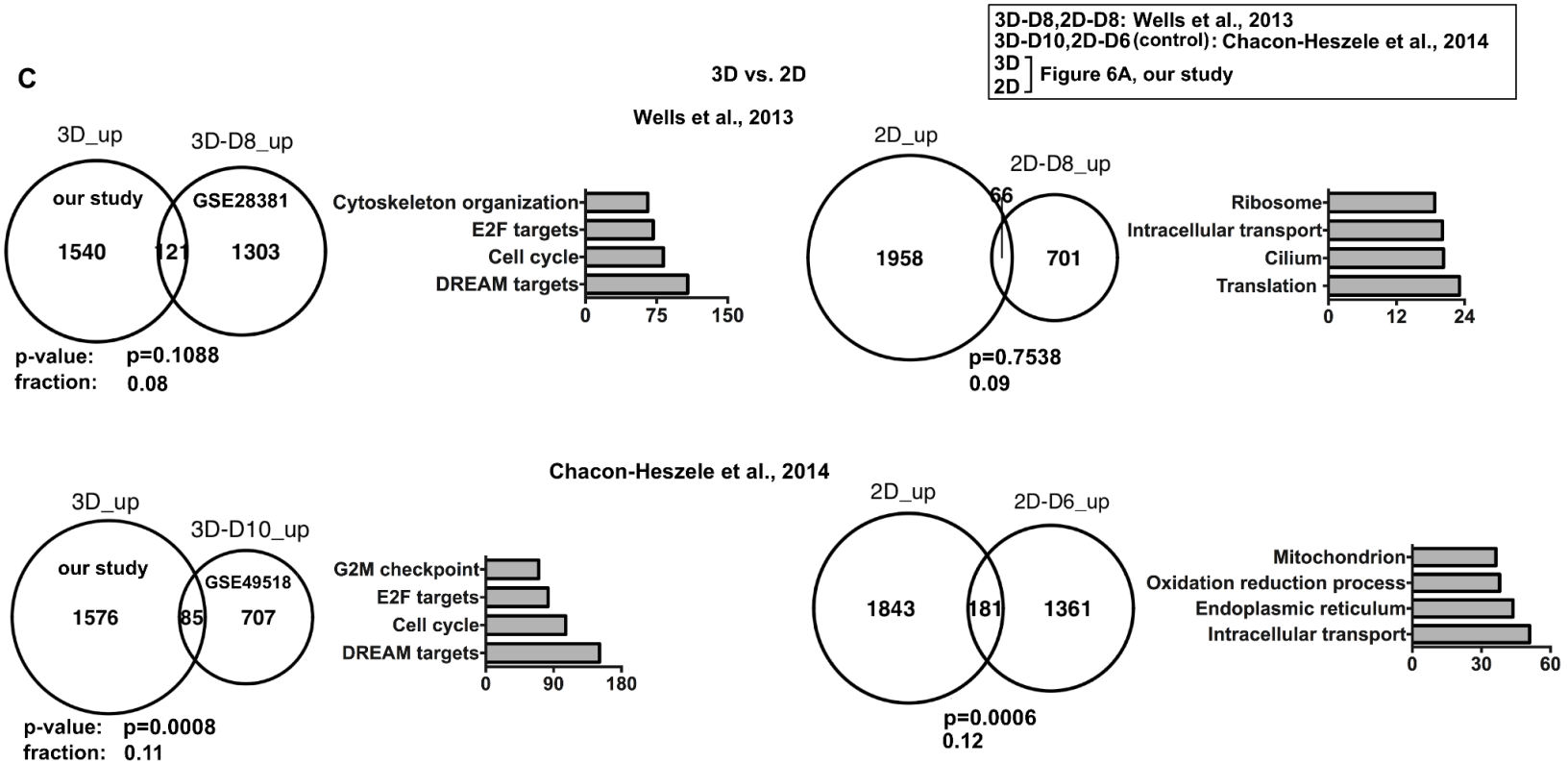

C
3D vs. 2D
3D-D8,2D-D8: Wells et al., 2013
3D-D10,2D-D6 (control): Chacon-Heszele et al., 2014
3D
2D
Figure 6A, our study
Wells et al., 2013
3D_up
3D-D8_up
our study
GSE28381
1540
121
1303
p-value: p=0.1088
fraction: 0.08
Cytoskeleton organization
E2F targets
Cell cycle
DREAM targets
0
75
150
2D_up
2D-D8_up
66
1958
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
GSE49518
1576
85
707
p-value: p=0.0008
fraction: 0.11
G2M checkpoint
E2F targets
Cell cycle
DREAM targets
0
90
180
2D_up
2D-D6_up
1843
181
1361
p=0.0006
0.12
Mitochondrion
Oxidation reduction process
Endoplasmic reticulum
Intracellular transport
0
30
60

Figure 3.6. 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 2.4 and 2.6.

(A) Comparison with DE genes (FDR< 0.05 and >2-fold change) between 36h and day8 samples from two microarray studies (Galvez-Santisteban et al., 2012; Wells et al., 2013). 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.

(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 findings are presented as described in panel A.

(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). DE genes for our study are indicated Figure 6A. Venn diagrams indicate the overlap, while bar plots indicate enriched functions of the microarray DE genes.

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.

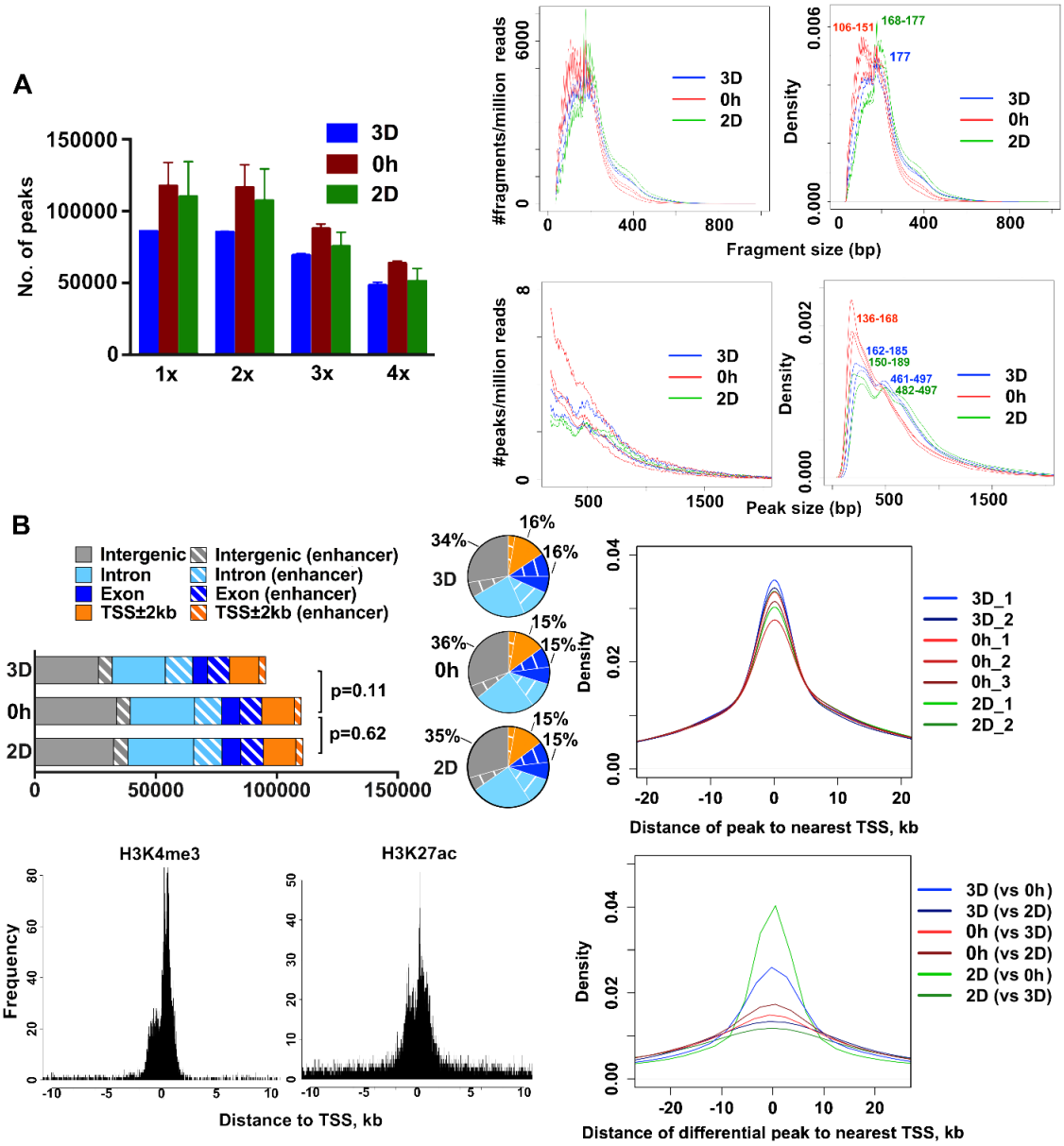

A
No. of peaks
150000
100000
50000
0
1x
2x
3x
4x
3D
0h
2D
#fragments/million reads
6000
0
0
400
800
Fragment size (bp)
Density
0.006
0.000
106-151
168-177
177
#peaks/million reads
8
0
500
1500
Peak size (bp)
0.002
136-168
162-185
150-189
461-497
482-497
B
Intergenic
Intron
Exon
TSS±2kb
Intergenic (enhancer)
Intron (enhancer)
Exon (enhancer)
TSS±2kb (enhancer)
p=0.11
p=0.62
50000
100000
150000
34%
16%
36%
15%
35%
Density
0.04
0.02
0.00
-20
-10
0
10
20
Distance of peak to nearest TSS, kb
3D_1
3D_2
0h_1
0h_2
0h_3
2D_1
2D_2
H3K4me3
H3K27ac
Frequency
80
60
40
20
50
30
10
-10
-5
5
10
Distance to TSS, kb
3D (vs 0h)
3D (vs 2D)
0h (vs 3D)
0h (vs 2D)
2D (vs 0h)
2D (vs 3D)
Distance of differential peak to nearest TSS, kb

Figure 3.7. Chromatin accessibility decreases during MDCKII 3D epithelial morphogenesis, revealed by ATAC-seq analysis. Related to Figure 2.7.

(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.

(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.

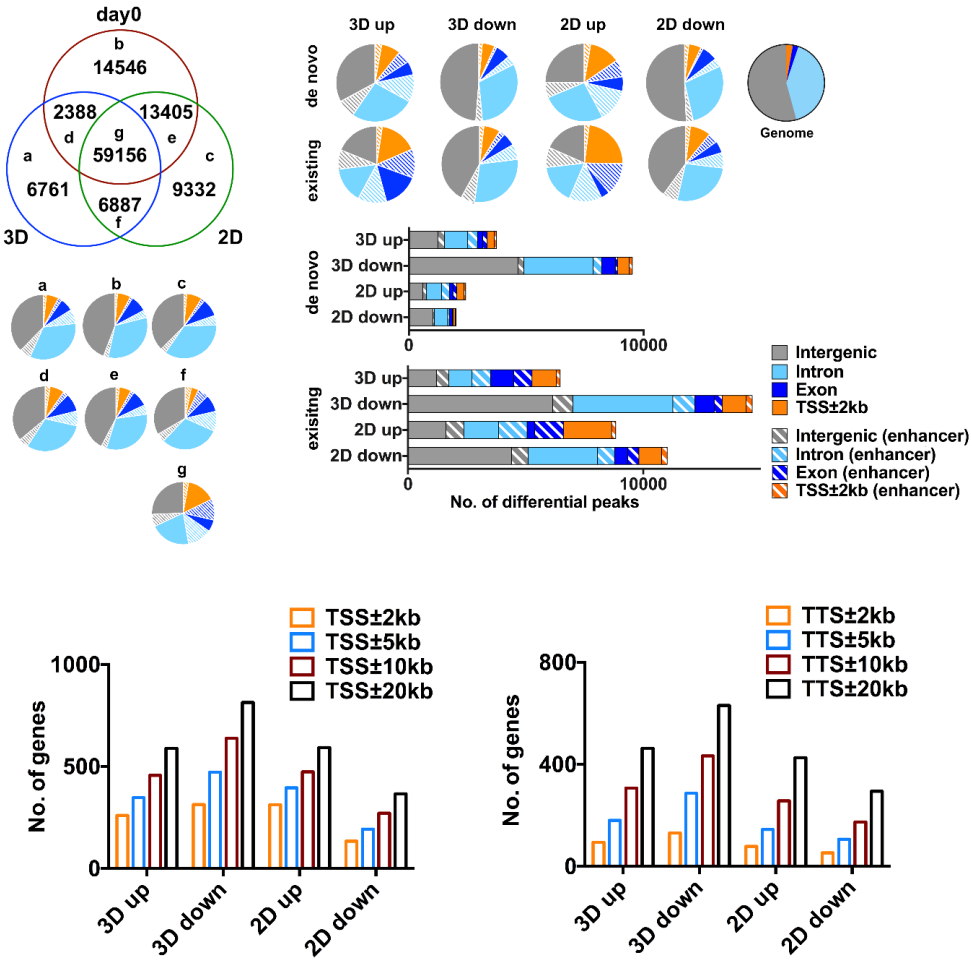


Figure 3.8. Genomic regions with increased chromatin accessibility during MDCKII epithelial morphogenesis are more likely to be found in promoters and exonic regions. Related to Figure 2.7.

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.

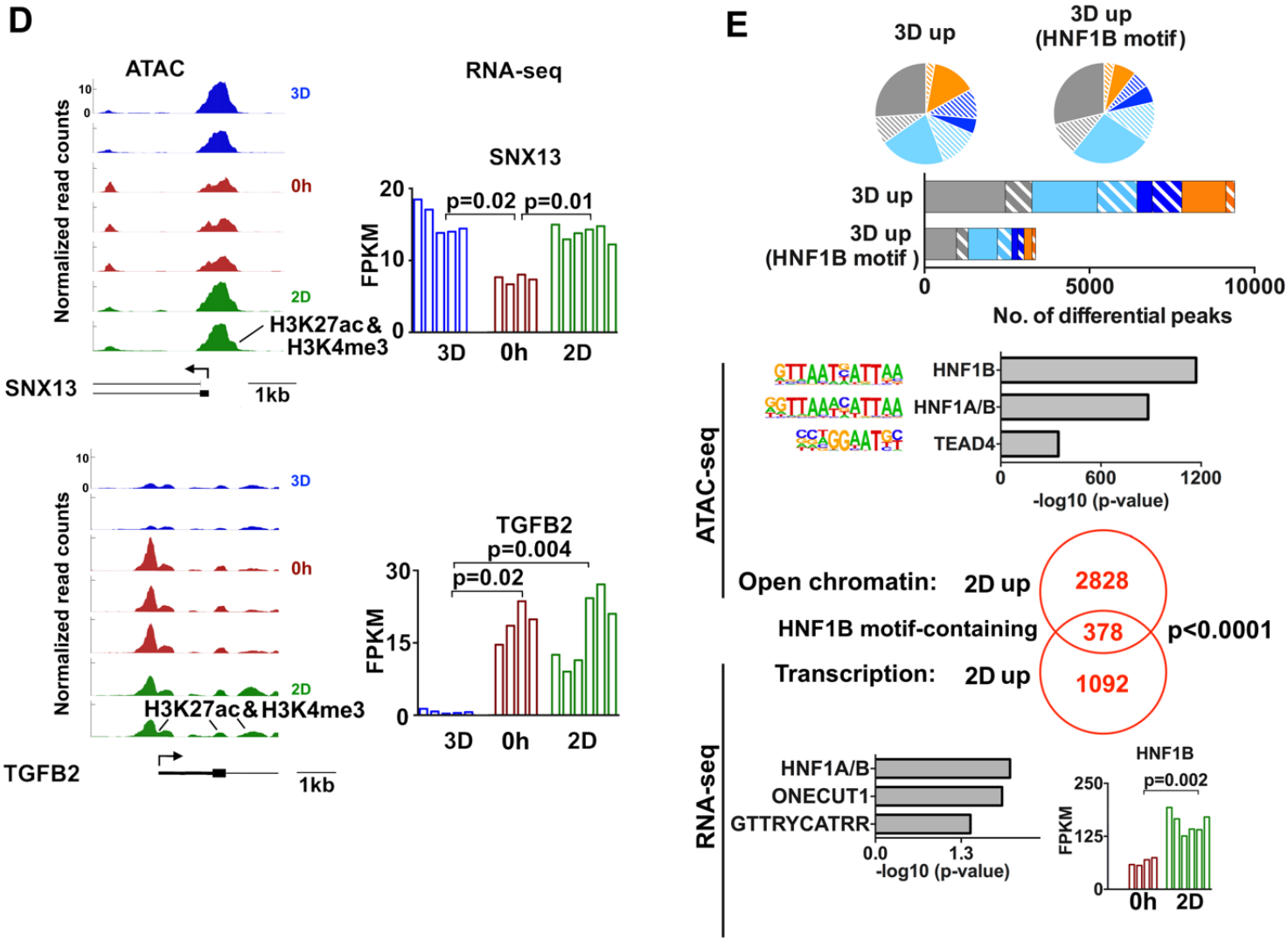

D
ATAC
RNA-seq
Normalized read counts
3D
0h
2D
H3K27ac & H3K4me3
SNX13
1kb
FPKM
p=0.02
p=0.01
TGFB2
p=0.004
E
3D up
3D up (HNF1B motif)
No. of differential peaks
ATAC-seq
HNF1B
HNF1A/B
TEAD4
-log10 (p-value)
Open chromatin:
2D up
2828
HNF1B motif-containing
378
p<0.0001
Transcription:
1092
RNA-seq
ONECUT1
GTTRYCATRR

Figure 3.9. Examples of chromatin accessibility correlating with gene expression, and up-regulation of HNF1B target genes during MDCKII epithelial morphogenesis. Related to Figure 2.7.

(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 up-regulated genes during 2D epithelial morphogenesis. Venn diagram indicates significant overlap between genes up-regulated 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 up-regulation of HNF1B, and the p-value of Wilcoxon test is indicated.

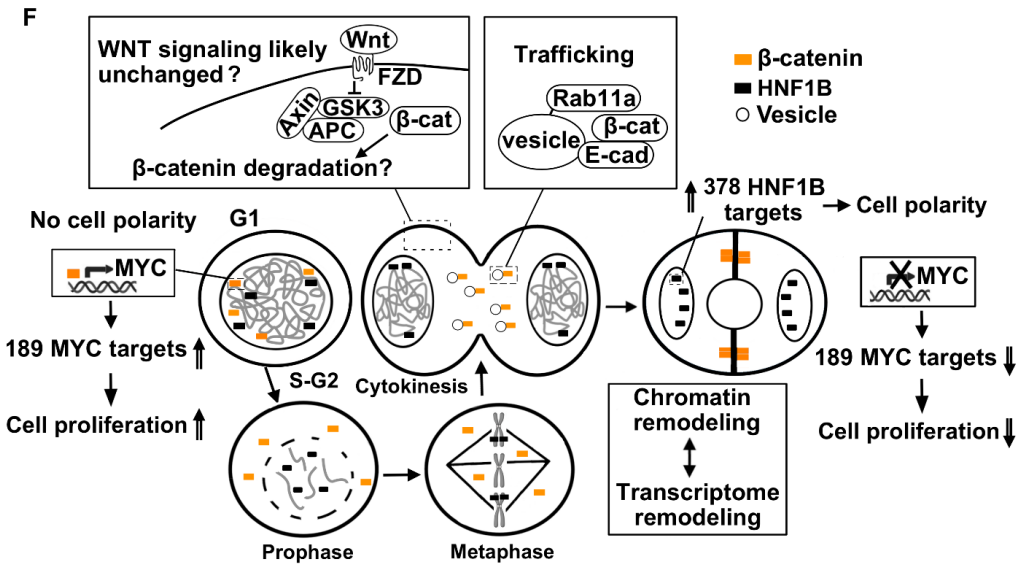


Figure 3.10. Our model proposes that the qualitative change in the transcriptome during MDCKII 3D cystogenesis is a result of the redistribution of transcription factors during the first cell cycle. Related to Figures 2.2, 2.4, 2.5 and 2.7. Our model is exemplified by β-catenin, which translocates to the cell-cell junctions, and HNF1B, which remains chromatin-bound via mitotic bookmarking, during the first cell cycle.

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, $1x10^6$ 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, Cat# 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. 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 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).

CHAPTER 4

CONCLUSIONS

MAIN FINDINGS

The findings from our study are significant in several aspects. First, our time-course RNA-seq study revealed a qualitative transcriptome change, which occurs after the first cell division and in parallel with lumen formation, during MDCKII cystogenesis. Second, we linked the intracellular trafficking to the observed transcriptome changes, using trafficking and cell polarity disrupted *DENND5A* and *AVL9*-knockdown cells. This is exemplified by the translocation of β-catenin from nucleus to cell-cell junctions which likely leads to the down-regulation of MYC and other cell cycle-related genes. Third, the integration of RNA-seq and ATAC-seq data revealed the association between transcriptome changes and changes in open chromatin regions underlying MDCKII epithelial morphogenesis. Importantly, we identified HNF1B which likely leads to the gene up-regulation that contributes to cell polarity establishment. Fourth, our ATAC-seq analysis indicated chromatin accessibility decreases during epithelial morphogenesis, with 3D cells having more drastic changes in open chromatin regions, compared with 2D cells. Fifth, although 3D and 2D epithelial morphogenesis share similar transcriptome reprogramming, which involves down-regulation of cell cycle-related genes, they show differences in metabolic phenotype. Specifically, 3D cells maintain active mitochondria and mitochondrial respiration whereas 2D cells likely reply greatly on glycolysis. Lastly, the RNA-seq and ATAC-seq data from our study would provide good resources for the

researchers utilizing MDCK cell models or a broader scientific community studying epithelial morphogenesis and differentiation.

We proposed a conceptual model for transcriptome remodeling as a consequence of the redistribution of transcription factors and chromatin remodeling, as exemplified, but not limited to β-catenin and HNF1B, during the first cell division (Figure 3.10). First, some DNA-associated transcription factors, such as β-catenin, are squeezed out of the nucleus and floated in the cytoplasm by trafficking vesicles, which occurs when the chromatin is condensed and the nuclear membrane is dismantled during the prophase of the first cell cycle. At the stage, other transcription factors, such as HNF1B, which have a bookmarking effect, remain associated with mitotic chromosomes even when the nuclear membrane disappears. Second, during cytokinesis and afterward, some transcription factors (e.g., β-catenin) in the cytoplasm are transported to the cell-cell junctions other locations, while some others may re-enter the nucleus when the chromatin relaxes and the nuclear membrane reforms. The other transcription factors (e.g., HNF1B) are still associated with the chromatin to activate or suppress their target genes. As a result, the chromatin is remodeled at a large scale, resulting in a qualitative change in the transcriptome. The model may be applied to other differentiation systems (e.g., differentiation from transit-amplifying cells to terminally differentiated cells during colon crypt development as described in chapter 2), which at least partially explains transcriptome changes during differentiation.

FUTURE DIRECTIONS

In our study, we hypothesized that β-catenin translocation likely results in the down-regulation of its target *MYC* and other cell cycle-related genes by a comparison between wildtype and *AVL9*-/*DENND5A*-knockdown cells during MDCKII epithelial morphogenesis. There are several limitations of our study. First, we may need further experiments to validate protein-level down-regulation of cell cycle genes, especially the important ones, such as TP53 and E2F1/3. Second, the disruption of β-catenin in the two knockdown cell lines is not specific, so there could be other possibilities that cause the relatively high expression of *MYC* and its target genes in knockdown cells. The experiments on MDCKII cells with specific disruption of β-catenin may provide more solid evidence to support our hypothesis. Previous studies show the functions of APC and Axin in exporting nuclear β-catenin and the roles of TCF4 and BCL9 in enriching β-catenin in the nucleus (Henderson and Fagotto, 2002; Krieghoff et al., 2006). We may be able to disrupt β-catenin translocation by manipulation of these proteins.

Future studies are required to dissect the mechanisms underlying the HNF1B regulated transcriptome program. First, our analysis of enhancers and promoters are solely based on previous published ChIP-seq data generated from canine liver tissues (Villar et al., 2015). In the future, H3K4me3, H3K4me1, and H3K27ac ChIP-seq data would be generated specific to MDCKII cells as resources for identifying promoters and enhancers that can be integrated into our ATAC-seq analysis. Second, additional experiments are required to validate our current findings on HNF1B. To validate that HNF1B activation contributes to the transcriptome switch involving the up-regulation of cell polarity and other genes, HNF1B ChIP-seq data can be generated and compared

between non-polarized 0h cells and fully polarized cells in 3D and 2D culture. Furthermore, Hi-ChIP experiments may be conducted for HNF1B, from which more accurate enhancer-promoter interactions associated with gene regulation by HNF1B can be identified. These data could be useful to validate our ATAC-seq results. Third, we have not tested the hypothesis that the bookmarking effect of HNF1B indeed contributes to the transcriptome switch. In the future, MDCKII cells with HNF1B P256S or V265L mutation, which leads to the disassociation between HNF1B and DNA, could be generated using CRISPR/Cas9 technology (Lerner et al., 2016). Transcriptome changes during epithelial morphogenesis of these mutant cells can be compared with those of wildtype cells. Fourth, we hypothesize that HNF1B immediately activates its target genes after the first cell division when β-catenin is translocated to the cell-cell junctions, based on previously mentioned findings from our study and the recent discovery that β-catenin and HNF1B compete for binding the same motif (Chan et al., 2019). Time-course ChIP-seq experiments for β-catenin and HNF1B of MDCKII epithelial morphogenesis including time points immediately before and after the first cell division, may be performed to test this hypothesis. It is expected that there is a significant overlap between β-catenin ChIP-seq peaks before the first cell division and HNF1B ChIP-seq peaks after that.

From our RNA-seq analysis and experimental validation, we revealed that 3D cells maintain active mitochondria and mitochondrial respiration but 2D cells do not. In line with our findings, a previous proteomic study showed MDCK cells grown in 2D culture have increased aerobic glycolysis compared to 3D culture (Pampaloni et al., 2010). To further compare metabolic phenotypes between MDCKII 3D and 2D culture,

Seahorse assay could be performed in the future to quantify the activity of mitochondrial respiration and glycolysis using the metrics of oxygen consumption rate and extracellular acidification rate, respectively. Previously, technical issues arise when we attempted to generate metabolic profiles for 3D culture using Seahorse assay, which may be resolved by using the assay plate specifically designed for 3D spheroids (Noel et al., 2017).

A metabolic switch has been linked to the down-regulation of MYC activity in cell fate determination during early ectodermal differentiation (Cliff et al., 2017). In our system, we observed the down-regulation of MYC activity in both 3D and 2D epithelial morphogenesis. Other reasons rather than MYC activity likely exist to explain the difference. To dissect the mechanisms underlying the up-regulation of mitochondria-related genes at the transcript level in 3D culture, a preliminary analysis was performed for mitochondrial biogenesis and functions related transcription factors in terms of their motif enrichment using ATAC-seq data. Not surprisingly, several key transcriptional regulators, such as NRF1, PPARα, PPARγ, and ERRα, how enriched binding motifs in expanded open chromatin regions in 3D cells compared to 2D cells. Interestingly, we observed an opposite trend for NRF2. In addition, a previously identified master regulator of mitochondrial biogenesis, PGC-1 (Lin et al., 2005), did not show a significant difference in motif enrichment between 3D and 2D culture, and its mRNA expression is constantly low during epithelial morphogenesis. It is possible that transcriptional activation of genes related to mitochondrial functions occurs through the action of the previously mentioned transcription factors in 3D culture. To validate this, ChIP-seq experiments for these transcription factors may be conducted to confirm the observations from ATAC-seq analysis. For a mechanistic study, the metabolic phenotype

can be obtained for MDCKII cells deficient in specific transcription factor(s) identified from ATAC-seq and ChIP-seq analysis and then compared with that of control cells.

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 biology *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.

Astrinidis, A., Cash, T.P., Hunter, D.S., Walker, C.L., Chernoff, J., and Henske, E.P. (2002). Tuberin, the tuberous sclerosis complex 2 tumor suppressor gene product, regulates Rho activation, cell adhesion and migration. Oncogene *21*, 8470-8476.

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. American Journal of Physiology-Renal Physiology *286*, F702-F710.

Baum, B., and Georgiou, M. (2011). Dynamics of adherens junctions in epithelial establishment, maintenance, and remodeling. The Journal of cell biology *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.

Bergmann, C., Guay-Woodford, L.M., Harris, P.C., Horie, S., Peters, D.J.M., and Torres, V.E. (2018). Polycystic kidney disease. Nat Rev Dis Primers *4*, 50.

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. Proceedings of the National Academy of Sciences of the United States of America *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.

Buenrostro, J.D., Giresi, P.G., Zaba, L.C., Chang, H.Y., and 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.

Carmona, F.J., Davalos, V., Vidal, E., Gomez, A., Heyn, H., Hashimoto, Y., Vizoso, M., Martinez-Cardus, A., Sayols, S., Ferreira, H.J., *et al*. (2014). A Comprehensive DNA Methylation Profile of Epithelial-to-Mesenchymal Transition. Cancer research *74*, 5608.

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. Journal of Cell Biology *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 Renal 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. Proceedings of the National Academy of Sciences of the United States of America *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. Journal of Cell Biology *144*, 687-699.

Chuman, L., Fine, L.G., Cohen, A.H., and Saier, M.H., Jr. (1982). Continuous growth of proximal tubular kidney epithelial cells in hormone-supplemented serum-free medium. J Cell Biol *94*, 506-510.

Cliff, T.S., Wu, T., Boward, B.R., Yin, A., Yin, H., Glushka, J.N., Prestegaard, J.H., and Dalton, S. (2017). MYC Controls Human Pluripotent Stem Cell Fate Decisions through Regulation of Metabolic Flux. Cell Stem Cell *21*, 502-516 e509.

Coradini, D., Casarsa, C., and Oriana, S. (2011). Epithelial cell polarity and tumorigenesis: new perspectives for cancer detection and treatment. Acta Pharmacol Sin *32*, 552-564.

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*.

Datta, A., Bryant, D.M., and Mostov, K.E. (2011). Molecular regulation of lumen morphogenesis. Current biology : CB *21*, R126-136.

Desclozeaux, M., Venturato, J., Wylie, F.G., Kay, J.G., Joseph, S.R., Le, H.T., and Stow, J.L. (2008). Active Rab11 and functional recycling endosome are required for E-cadherin trafficking and lumen formation during epithelial morphogenesis. Am J Physiol Cell Physiol *295*, C545-556.

Desrochers, T.M., Palma, E., and Kaplan, D.L. (2014). Tissue-engineered kidney disease models. Adv Drug Deliv Rev *69-70*, 67-80.

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.*, et al*. (2016). RBM5-AS1 Is Critical for Self-Renewal of Colon Cancer Stem-like Cells. Cancer research *76*, 5615-5627.

Domaschenz, R., Kurscheid, S., Nekrasov, M., Han, S., and Tremethick, D.J. (2017). The Histone Variant H2A.Z Is a Master Regulator of the Epithelial-Mesenchymal Transition. Cell Reports *21*, 943-952.

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

Elia, N., and Lippincott-Schwartz, J. (2009). Culturing MDCK cells in three dimensions for analyzing intracellular dynamics. Current protocols in cell biology / editorial board, Juan S Bonifacino [et al] *Chapter 4*, Unit 4 22.

Engelberg, J.A., Datta, A., Mostov, K.E., and Hunt, C.A. (2011). MDCK cystogenesis driven by cell stabilization within computational analogues. PLoS Comput Biol *7*, e1002030.

Feng, J., Liu, T., Qin, B., Zhang, Y., and Liu, X.S. (2012). Identifying ChIP-seq enrichment using MACS. Nature protocols *7*, 1728-1740.

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.

Hao, Y., Du, Q., Chen, X., Zheng, Z., Balsbaugh, J.L., Maitra, S., Shabanowitz, J., Hunt, D.F., and Macara, I.G. (2010). Par3 controls epithelial spindle orientation by aPKC-mediated phosphorylation of apical Pins. Curr Biol *20*, 1809-1818.

He, J., Zhou, H., Meng, J., Zhang, S., Li, X., Wang, S., Shao, G., Jin, W., Geng, X., Zhu, S., *et al*. (2020). Cardamonin retards progression of autosomal dominant polycystic kidney disease via inhibiting renal cyst growth and interstitial fibrosis. Pharmacol Res *155*, 104751.

Heinz, S., Benner, C., Spann, N., Bertolino, E., Lin, Y.C., Laslo, P., Cheng, J.X., Murre, C., Singh, H., and 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.

Henderson, B.R., and Fagotto, F. (2002). The ins and outs of APC and beta-catenin nuclear transport. EMBO Rep *3*, 834-839.

Homma, Y., Kinoshita, R., Kuchitsu, Y., Wawro, P.S., Marubashi, S., Oguchi, M.E., Ishida, M., Fujita, N., and Fukuda, M. (2019). Comprehensive knockout analysis of the Rab family GTPases in epithelial cells. J Cell Biol *218*, 2035-2050.

Jalal, F., Lemay, G., Zollinger, M., Berteloot, A., Boileau, G., and Crine, P. (1991). Neutral endopeptidase, a major brush border protein of the kidney proximal nephron, is directly targeted to the apical domain when expressed in Madin-Darby canine kidney cells. J Biol Chem *266*, 19826-19832.

Kobayashi, W., and Ozawa, M. (2013). The transcription factor LEF-1 induces an epithelial-mesenchymal transition in MDCK cells independent of beta-catenin. Biochem Biophys Res Commun *442*, 133-138.

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. Proceedings of the National Academy of Sciences of the United States of America *104*, 15418-15423.

Krieghoff, E., Behrens, J., and Mayr, B. (2006). Nucleo-cytoplasmic distribution of beta-catenin is regulated by retention. J Cell Sci *119*, 1453-1463.

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., *et al*. (2016). Enrichr: a comprehensive gene set enrichment analysis web server 2016 update. Nucleic acids research *44*, W90-97.

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

Lee, M., and Vasioukhin, V. (2008). Cell polarity and cancer--cell and tissue polarity as a non-canonical tumor suppressor. J Cell Sci *121*, 1141-1150.

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 research *44*, 8097-8111.

Lesko, A.C., Goss, K.H., Yang, F.F., Schwertner, A., Hulur, I., Onel, K., and Prosperi, J.R. (2015). The APC tumor suppressor is required for epithelial cell polarization and three-dimensional morphogenesis. Biochimica et Biophysica Acta (BBA) - Molecular Cell Research *1853*, 711-723.

Li, H., and 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., and Zhao, S. (2014). Cancer driver candidate genes AVL9, DENND5A and NUPL1 contribute to MDCK cystogenesis. Oncoscience *1*, 854-865.

Lin, J., Handschin, C., and Spiegelman, B.M. (2005). Metabolic control through the PGC-1 family of transcription coactivators. Cell Metab *1*, 361-370.

Liu, D., Xiong, H., Ellis, A.E., Northrup, N.C., Rodriguez, C.O., Jr., O'Regan, R.M., Dalton, S., and Zhao, S. (2014). Molecular homology and difference between spontaneous canine mammary cancer and human breast cancer. Cancer research *74*, 5045-5056.

Lock, J.G., and Stow, J.L. (2005). Rab11 in recycling endosomes regulates the sorting and basolateral transport of E-cadherin. Molecular biology of the cell *16*, 1744-1755.

Lu, Z., Hofmeister, B.T., Vollmers, C., DuBois, R.M., and Schmitz, R.J. (2017). Combining ATAC-seq with nuclei sorting for discovery of cis-regulatory regions in plant genomes. Nucleic Acids Res *45*, e41.

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.

Martin-Belmonte, F., and Perez-Moreno, M. (2011). Epithelial cell polarity, stem cells and cancer. Nat Rev Cancer *12*, 23-38.

Martin-Belmonte, F., Yu, W., Rodriguez-Fraticelli, A.E., Ewald, A.J., Werb, Z., Alonso, M.A., and Mostov, K. (2008). Cell-polarity dynamics controls the mechanism of lumen formation in epithelial morphogenesis. Curr Biol *18*, 507-513.

Masuda, M., Yageta, M., Fukuhara, H., Kuramochi, M., Maruyama, T., Nomoto, A., and Murakami, Y. (2002). The tumor suppressor protein TSLC1 is involved in cell-cell adhesion. J Biol Chem *277*, 31014-31019.

McCaffrey, L.M., and Macara, I.G. (2011). Epithelial organization, cell polarity and tumorigenesis. Trends Cell Biol *21*, 727-735.

McCaffrey, L.M., and Macara, I.G. (2012). Signaling pathways in cell polarity. Cold Spring Harbor perspectives in biology *4*.

McLaughlin, M.E., Kruger, G.M., Slocum, K.L., Crowley, D., Michaud, N.A., Huang, J., Magendantz, M., and Jacks, T. (2007). The Nf2 tumor suppressor regulates cell-cell adhesion during tissue fusion. Proc Natl Acad Sci U S A *104*, 3261-3266.

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.

Milani, P., Escalante-Chong, R., Shelley, B.C., Patel-Murray, N.L., Xin, X.F., Adam, M., Mandefro, B., Sareen, D., Svendsen, C.N., and Fraenkel, E. (2016). Cell freezing protocol suitable for ATAC-Seq on motor neurons derived from human induced pluripotent stem cells. Sci Rep-Uk *6*.

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 *30*, 53-67.

Moreno-Bueno, G., Portillo, F., and Cano, A. (2008). Transcriptional regulation of cell polarity in EMT and cancer. Oncogene *27*, 6958-6969.

Musch, A., Cohen, D., Yeaman, C., Nelson, W.J., Rodriguez-Boulan, E., and Brennwald, P.J. (2002). Mammalian homolog of Drosophila tumor suppressor lethal (2) giant larvae interacts with basolateral exocytic machinery in Madin-Darby canine kidney cells. Mol Biol Cell *13*, 158-168.

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. Nature cell biology *3*, 785-792.

Noel, P., Munoz, R., Rogers, G.W., Neilson, A., Von Hoff, D.D., and Han, H. (2017). Preparation and Metabolic Assay of 3-dimensional Spheroid Co-cultures of Pancreatic Cancer Cells and Fibroblasts. J Vis Exp.

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. Developmental 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.

Pampaloni, F., Stelzer, E.H., Leicht, S., and Marcello, M. (2010). Madin-Darby canine kidney cells are increased in aerobic glycolysis when cultured on flat and stiff collagen-coated surfaces rather than in physiological 3-D cultures. Proteomics *10*, 3394-3413.

Pastor, W.A., Liu, W., Chen, D., Ho, J., Kim, R., Hunt, T.J., Lukianchikov, A., Liu, X., Polo, J.M., Jacobsen, S.E., *et al*. (2018). TFAP2C regulates transcription in human naive pluripotency by opening enhancers. Nature cell biology *20*, 553-564.

Pertea, M., Kim, D., Pertea, G.M., Leek, J.T., and Salzberg, S.L. (2016). Transcript-level expression analysis of RNA-seq experiments with HISAT, StringTie and Ballgown. Nature Protocols *11*, 1650.

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. The EMBO journal *37*.

Prasad, H., Dang, D.K., Kondapalli, K.C., Natarajan, N., Cebotaru, V., and Rao, R. (2019). NHA2 promotes cyst development in an in vitro model of polycystic kidney disease. J Physiol *597*, 499-519.

Qin, Y., Meisen, W.H., Hao, Y., and Macara, I.G. (2010). Tuba, a Cdc42 GEF, is required for polarized spindle orientation during epithelial cyst formation. J Cell Biol *189*, 661-669.

Reich, M., Liefeld, T., Gould, J., Lerner, J., Tamayo, P., and Mesirov, J.P. (2006). GenePattern 2.0. Nature genetics *38*, 500-501.

Rodriguez-Boulan, E., and Macara, I.G. (2014). Organization and execution of the epithelial polarity programme. Nature Reviews Molecular Cell Biology *15*, 225-242.

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 Harbor perspectives in biology *4*, a007989.

Schluter, M.A., Pfarr, C.S., Pieczynski, J., Whiteman, E.L., Hurd, T.W., Fan, S., Liu, C.J., and Margolis, B. (2009). Trafficking of Crumbs3 during cytokinesis is crucial for lumen formation. Mol Biol Cell *20*, 4652-4663.

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 Harbor perspectives in biology *4*, a008284.

Shukla, P., Vogl, C., Wallner, B., Rigler, D., Muller, M., and Macho-Maschler, S. (2015). High-throughput mRNA and miRNA profiling of epithelial-mesenchymal transition in MDCK cells. BMC Genomics *16*, 944.

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., *et al*. (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. Developmental cell *19*, 126-137.

Takiar, V., Nishio, S., Seo-Mayer, P., King, J.D., Jr., Li, H., Zhang, L., Karihaloo, A., Hallows, K.R., Somlo, S., and Caplan, M.J. (2011). Activating AMP-activated protein kinase (AMPK) slows renal cystogenesis. Proc Natl Acad Sci U S A *108*, 2462-2467.

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.

Trapnell, C., Roberts, A., Goff, L., Pertea, G., Kim, D., Kelley, D.R., Pimentel, H., Salzberg, S.L., Rinn, J.L., and Pachter, L. (2012). Differential gene and transcript expression analysis of RNA-seq experiments with TopHat and Cufflinks. Nature Protocols *7*, 562-578.

Tsompana, M., and Buck, M.J. (2014). Chromatin accessibility: a window into the genome. Epigenetics & Chromatin *7*, 33.

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., *et al*. (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. Nature medicine *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*.

Wei, C., Bhattaram, V.K., Igwe, J.C., Fleming, E., and Tirnauer, J.S. (2012). The LKB1 tumor suppressor controls spindle orientation and localization of activated AMPK in mitotic epithelial cells. PLoS One *7*, e41118.

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*, C966-975.

Yamashita, K., Ide, M., Furukawa, K.T., Suzuki, A., Hirano, H., and Ohno, S. (2015). Tumor suppressor protein Lgl mediates G1 cell cycle arrest at high cell density by forming an Lgl-VprBP-DDB1 complex. Mol Biol Cell *26*, 2426-2438.

Yang, B., Sonawane, N.D., Zhao, D., Somlo, S., and Verkman, A.S. (2008). Small-molecule CFTR inhibitors slow cyst growth in polycystic kidney disease. J Am Soc Nephrol *19*, 1300-1310.

Zhang, Y., Liu, T., Meyer, C.A., Eeckhoute, J., Johnson, D.S., Bernstein, B.E., Nusbaum, C., Myers, R.M., Brown, M., Li, W., *et al*. (2008). Model-based Analysis of ChIP-Seq (MACS). Genome Biology *9*, R137.

Zhang, Y., Young, A., Zhang, J., and Chen, X. (2015). P73 tumor suppressor and its targets, p21 and PUMA, are required for madin-darby canine kidney cell morphogenesis by maintaining an appropriate level of epithelial to mesenchymal transition. Oncotarget *6*, 13994-14004.

Zhou, H., Gao, J., Zhou, L., Li, X., Li, W., Li, X., Xia, Y., and Yang, B. (2012). Ginkgolide B inhibits renal cyst development in in vitro and in vivo cyst models. Am J Physiol Renal Physiol *302*, F1234-1242.