Article:
Lu, H., Goh, K.J., Tan, E.K. et al. (2025) Nuclear-cytoplasmic translocation of MCIDAS couples transcription with massive de novo centriole biogenesis in multiciliated cells. Cell Reports, 44 (10). 116321. ISSN: 2211-1247

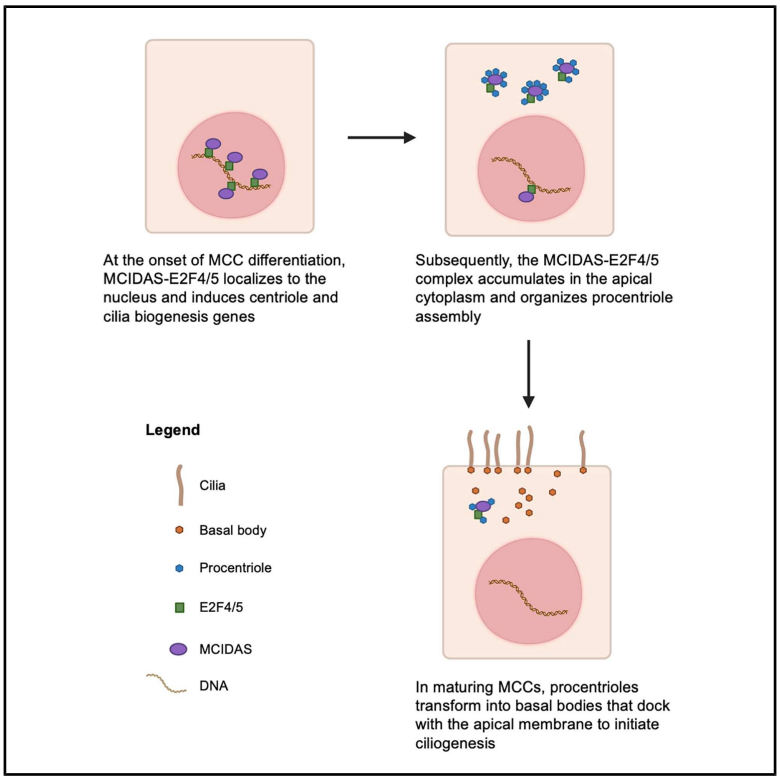
https://doi.org/10.1016/j.celrep.2025.116321

Cell Reports

Article

Nuclear-cytoplasmic translocation of MCIDAS couples transcription with massive *de novo* centriole biogenesis in multiciliated cells

Graphical abstract



Authors

Hao Lu, Kim Jee Goh, Ee Kim Tan, ..., Alice Meunier, N. Ray Dunn, Sudipto Roy

Correspondence

sudipto_roy@a-star.edu.sg

In brief

How multiciliated cells (MCCs) generate hundreds of centrioles for multiciliation has remained an unresolved problem in cell biology. Lu et al. show that the MCC "master" transcriptional regulator, MCIDAS, not only functions in the nucleus to induce genes for multiciliation but also undergoes cytoplasmic translocation to organize explosive centriole biogenesis.

Highlights

- MCIDAS undergoes nucleus-to-cytoplasmic translocation during MCC differentiation
- MCIDAS loss blocks E2F4/5 cytoplasmic accumulation, inhibiting centriole assembly
- Cytoplasmic MCIDAS-E2F4/5 organizes massive *de novo* centriole biogenesis in MCCs
- MCIDAS targets encode MCC-specific proteins, with many linked to human MCC disorders

Lu et al., 2025, Cell Reports *44*, 116321
October 28, 2025
https://doi.org/10.1016/j.celrep.2025.116321

CellPress

Cell Reports

CellPress
OPEN ACCESS

Article

Nuclear-cytoplasmic translocation of MCIDAS couples transcription with massive *de novo* centriole biogenesis in multiciliated cells

Hao Lu,[1,12] Kim Jee Goh,[2,12,13] Ee Kim Tan,[3,12] Cameron T. James,[1,4,12,14] Arnab Ghosh,[5,6,12] Amélie-Rose Boudjema,[7] Paolo Alberto Lorenzini,[8,9] Colin D. Bingle,[4] Sebastian Maurer-Stroh,[9,10] Nidhan K. Biswas,[5] Alice Meunier,[7] N. Ray Dunn,[2,3,15] and Sudipto Roy[1,11,16,*]

[1]Institute of Molecular and Cell Biology (IMCB), Agency for Science, Technology and Research (A*STAR), Proteos, 61 Biopolis Drive, Singapore 138673, Singapore
[2]Skin Research Institute of Singapore, 11 Mandalay Road #17-01 Clinical Sciences Building, Singapore 308232, Singapore
[3]Lee Kong Chian School of Medicine, Nanyang Technological University, Clinical Sciences Building, 11 Mandalay Road, Singapore 308232, Singapore
[4]Division of Clinical Medicine, School of Medicine and Public Health, University of Sheffield, Sheffield S10 2JF, UK
[5]Biotechnology Research and Innovation Council - National Institute of Biomedical Genomics (BRIC-NIBMG), Kalyani, India
[6]Biotechnology Research and Innovation Council - Regional Centre for Biotechnology (BRIC-RCB), Faridabad, India
[7]Institut de Biologie de l'École Normale Supérieure (IBENS), CNRS, UMR 8197, INSERM, U1024, Paris Sciences et Lettres (PSL), Research University, Paris, France
[8]Infectious Disease Labs (IDLabs), Agency for Science, Technology and Research (A*STAR), 8a Biomedical Grove, #05-13, Immunos, Singapore 1386348, Singapore
[9]Bioinformatics Institute (BII), Agency for Science, Technology and Research (A*STAR), 30 Biopolis Street, #07-01, Matrix, Singapore 138671, Singapore
[10]Yong Loo Lin School of Medicine and Department of Biological Sciences, National University of Singapore, Singapore 117597, Singapore
[11]Department of Pediatrics, Yong Loo Lin School of Medicine, National University of Singapore, 1E Kent Ridge Road, Singapore 119288, Singapore
[12]These authors contributed equally
[13]Present address: Hoxton Farms, HYLO Building, 105 Bunhill Row, London EC1Y 8LZ, UK
[14]Present address: Australian Institute for Bioengineering and Nanotechnology (AIBN), Level 3W, Building #75, Old Cooper Rd, QLD 4067, Australia
[15]Present address: Vertex Pharmaceuticals, 50 Northern Avenue, Boston, MA 02210, USA
[16]Lead contact
*Correspondence: sudipto_roy@a-star.edu.sg
https://doi.org/10.1016/j.celrep.2025.116321

SUMMARY

Multiciliated cells (MCCs) bear numerous motile cilia that drive fluid flow, but how numerous centrioles for multiciliation are generated has remained unresolved. Here, we report that the "master" MCC transcriptional regulator, MCIDAS, moonlights in the cytoplasm to organize massive centriole biogenesis. Like MCIDAS, its co-transcriptional factors, E2F4 and E2F5, also undergo cytoplasmic accumulation, colocalizing with MCIDAS and forming procentrioles. MCIDAS loss inhibited E2F4/5 cytoplasmic accumulation and blocked centriole assembly. Furthermore, we show that the cytoplasmic accumulation of MCIDAS is mediated by CRM1-dependant nuclear export, and its inhibition specifically compromised centriole biogenesis. By contrast, on loss of parental centrioles and deuterosomes, which does not impair centriole formation, E2F4 cytoplasmic localization remained unaffected, establishing that the MCIDAS-E2F4/5 cytoplasmic complex represents the *de novo* centriole biogenesis pathway. We have also assembled a comprehensive list of MCIDAS targets, a resource which will enable further exposition of MCC biology and pathological mechanisms of motile ciliopathies.

INTRODUCTION

Multiciliated cells (MCCs) are specialized ciliated cells found in the airways, brain ventricles, and reproductive organs, where metachronal beating of their dense ciliary arrays promotes mucus clearance, circulation of cerebrospinal fluid (CSF), and gamete transport, respectively.[1] The importance of MCC function is best exemplified by patients with the motile ciliopathy primary ciliary dyskinesia (PCD) and its variant subtype, reduced generation of multiple motile cilia (RGMC), who are afflicted with severe respiratory symptoms like bronchiectasis, are typically infertile, and can also develop ventriculomegaly and hydrocephalus.[2]

The mechanistic basis of explosive centriole biogenesis, which produces basal bodies for multiciliation during MCC differentiation, has remained an enduring problem in cell biology. Historically, two pathways have been postulated and experimentally explored over the past several decades. The vast majority of centrioles were thought to be assembled on doughnut-shaped organelles, the deuterosomes, that are peculiar to MCC precursors.[3–5] In addition, the parental centrioles have also been implicated in templating a minor population of centrioles,[6] akin to centriole duplication that occurs in cycling cells. Surprisingly, recent studies have shown that the deuterosomes as well as parental centrioles are dispensable for centriole amplification in MCCs, leading to the view that the process could be organized entirely *de novo*.[7–10] However, the details of this *de novo* pathway have remained elusive.

The paralogous GEMININ family proteins, GMNC and MCIDAS, act as "master" regulators of MCC development, with GMNC required to specify MCC precursors and MCIDAS acting downstream to promote MCC differentiation by activating centriole biogenesis and multiciliation genes.[11] Since GMNC and MCIDAS are coiled-coil-containing proteins, they are unable to directly bind DNA. Instead, they regulate transcription by associating with the cell-cycle transcriptional regulators, E2F4 and E2F5.[12–15] Besides functioning in MCC transcription, E2F4 has recently been shown to also exhibit a striking nuclear-cytoplasmic shift during MCC differentiation.[16] In the cytoplasm, it associates with the core deuterosomal protein, DEUP1,[5] and SAS6, an essential centriolar cartwheel component,[17] suggesting that it could represent an additional factor in centriole assembly.[16,18] Because deuterosomes and parental centrioles are dispensable for centriole amplification, the significance of the cytoplasmic accumulation of E2F4 has remained unresolved.

We now establish that in addition to its pivotal role in MCC-specific transcription, intriguingly, MCIDAS also undergoes cytoplasmic localization and is required for the cytoplasmic accumulation of E2F4 and E2F5. We show that the cytoplasmic localization of MCIDAS is mediated by CRM1-dependant nuclear export, and the cytoplasmic MCIDAS-E2F4/5 complex is as an important mediator of massive *de novo* centriole biogenesis in MCCs. Thus, MCC differentiation is coordinated by intimately coupled transcriptional and centriole biogenesis programs, through the nuclear-cytoplasmic translocation of the key transcriptional regulator, MCIDAS.

RESULTS

E2F5 accumulates in cytoplasmic puncta in MCCs, in the vicinity of forming procentrioles

Like E2F4, E2F5 can interact with DEUP1 and SAS6,[18] implying a role in MCC centriole biogenesis. To investigate this, we used antibodies against human E2F5 and analyzed its subcellular localization in differentiating airway MCCs derived from cultured human embryonic stem cell (hESC) line H9. Briefly, hESCs were differentiated sequentially into definitive endoderm and anterior foregut endoderm, followed by sorting for lung progenitors. Lung progenitors were grown into lung organoids and sorted for basal cells, which were then grown in submerged transwells to confluency, followed by culture at air-liquid interface (ALI) to stimulate airway epithelium differentiation[19,20] (Figures 1A–1C and Video S1). We found that during initiation of MCC differentiation, there was prominent E2F5 accumulation in the apical cytoplasm, similar to E2F4 reported previously[16] (Figure 1D). Moreover, the E2F5 cytoplasmic puncta were observed in the immediate vicinity of the centriolar component SAS6 as well as E2F4 (Figures 1E and 1F). We validated the E2F5 antibodies using MCC cultures derived from *E2F5* mutant hESCs (Figures S1A–S1F), where we failed to observe cytoplasmic E2F5 (Figures 1G and S1G). Thus, like E2F4, E2F5 undergoes cytoplasmic accumulation in differentiating MCCs, distributed juxtaposed to forming procentrioles (marked by SAS6) and E2F4, suggesting that it is indeed a part of the MCC centriole biogenesis pathway.

MCIDAS is required for cytoplasmic accumulation of E2F proteins in MCCs

We cultured MCCs from mouse tracheal epithelial cells (mTECs) of *Mcidas* mutant mice as described before[15] and monitored the cytoplasmic accumulation of E2F4/5. Unlike cultures from wild-type mice tracheae, where cytoplasmic E2F4/5 foci could be readily detected, we failed to observe such accumulation in *Mcidas* mutant mTECs (Figures 2A–2D). Using gene editing in hESCs, we also generated a mutation in human *MCIDAS*, an 8 bp homozygous deletion predicted to produce a highly truncated protein (Figures S2A–S2D). While wild-type cells readily produced MCCs, *MCIDAS* mutant hESCs failed to do so, mimicking the phenotype of MCIDAS loss reported from mice and PCD patients[15,21] (Figures S2E and S2F). Ciliary transcription factors like FOXJ1, RFX2, and RFX3 were expressed but were observed in fewer cells, and their levels appeared lower (Figures S2E–S2K). We again failed to detect the cytoplasmic accumulation of E2F4 and E2F5 (Figures 2E–2J). This defect could arise from several possibilities: (1) *E2F* gene transcription and/or expression and stability of the E2F proteins could be impaired on MCIDAS loss, (2) MCIDAS could induce expression of one or more proteins that facilitates E2F4/5 cytoplasmic accumulation, or (3) MCIDAS could more directly regulate E2F4/5 cytoplasmic accumulation since it associates with them for transcriptional regulation. Because E2F4/5 expression is not affected by MCIDAS loss (see Figures S3A–S3D for nuclear E2F4/5 expression in *Mcidas* mutant mTECs at ALI day 0; Figures S3E and S3G for E2F4 protein analysis in *Mcidas* mutant mTECs; Table S1 for *E2f4/5* mRNA expression analysis by RNA sequencing [RNA-seq] of *Mcidas* mutant mTECs as well as the discussion section) and the identification of any MCIDAS target that could facilitate their cytoplasmic accumulation will require extensive functional interrogation of MCIDAS target genes, we investigated whether MCIDAS itself undergoes nuclear-cytoplasmic translocation during MCC differentiation and thereby facilitates the cytoplasmic accumulation of the E2F proteins.

MCIDAS accumulates in the cytoplasm during MCC differentiation

Although previous studies have reported the use of two different antibodies against MCIDAS, the subcellular expression pattern and dynamics of the endogenous protein during MCC development have not been analyzed. The antibody generated in-house

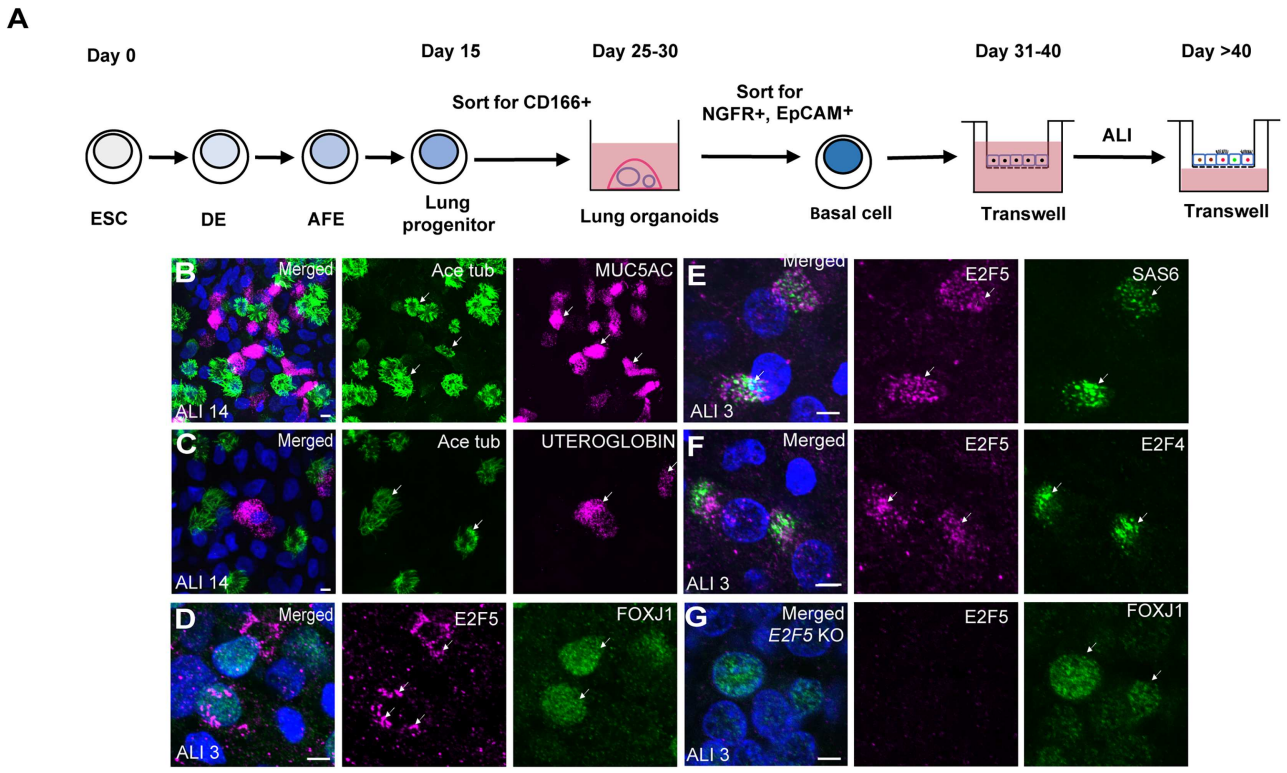


Figure 1. E2F5 accumulates in cytoplasmic puncta in differentiating MCCs, closely associated with centriolar proteins

(A) H9 hESC differentiation protocol into airway epithelial cells. ESC: embryonic stem cell; DE: definitive endoderm; AFE: anterior foregut endoderm; ALI: air-liquid interface.

(B) Differentiated hESC cultures contained MCCs (acetylated tubulin [Ace-tub], arrows) and mucus producing cells (MUC5AC, arrows). Scale bar: 5 μm.

(C) Differentiated hESC cultures contained MCCs (Ace-tub, arrows) and secretory club cells (UTEROGLOBIN [aka SCGB1A1], arrows). Scale bar: 5 μm.

(D) E2F5 cytoplasmic puncta (arrows) in MCCs. MCCs were identified using FOXJ1 staining (arrows). Scale bar: 5 μm.

(E) E2F5 cytoplasmic puncta (arrows) in the vicinity of SAS6 (arrows). Colocalization (white signal) is indicated in merged panel (arrow). Scale bar: 5 μm.

(F) E2F5 (arrows) and E2F4 (arrows) cytoplasmic puncta in close proximity. Scale bar: 5 μm.

(G) E2F5 cytoplasmic puncta were absent from MCCs derived from *E2F5* mutant (*E2F5* KO) hESCs. MCCs were identified using FOXJ1 staining (arrows). Scale bar: 5 μm.

4′,6-diamidino-2-phenylindole (DAPI) was used to highlight nuclei (blue) in (B)–(G).

by the Lygerou group[22] was not made available to us. A commercially produced antibody reported by Boon et al.[21] is likely to be non-specific: while Boon et al. have reported migration of MCIDAS at around 38 kDa using this antibody (the predicted molecular weight of MCIDAS), we and others have consistently found that overexpressed as well as endogenous MCIDAS migrates at around 70 kDa (possibly due to post-translational modifications, such as phosphorylation, of the protein [see Ma et al.,[12] Lu et al.,[15] Arbi et al.,[23] and this study]). Not surprisingly, we failed to detect endogenous MCIDAS with this antibody in our hESC-derived MCC cultures. To overcome this limitation, we used CRISPR/Cas9-mediated knockin to introduce triple hemagglutinin (3×HA) epitope tag in frame with the N terminus of MCIDAS in H9 hESCs (immediately after the start codon; hereafter referred to as HA-MCIDAS) (Figures S4A–S4E). We engineered the HA tag at the N terminus because the C terminus of GMNC and MCIDAS is vital for interaction with the E2F proteins[12,13,15] (see Figure S4F for validation of HA-tagged MCIDAS protein). It is unlikely that the tag disrupted MCIDAS activity, as a single HA tag placed at a similar position does not affect interaction with E2F4/5 or the ability to rescue MCC formation in *Mcidas* mutant mTECs.[15] At the start of MCC differentiation, we found an exclusively nuclear localized HA-MCIDAS signal (Figures 3A and 3A′). However, subsequently, when procentriole biogenesis began, prominent puncta of HA-MCIDAS accumulated in the cytoplasm, with a concomitant decline in nuclear intensity (Figures 3B, 3B′, and 3G). The cytoplasmic and nuclear HA signals were found in two discrete subcellular locations: nuclei were located basally, while the cloud of cytoplasmic HA was more apically distributed (Figures 3A, 3A′, 3B, and 3B′), similar to E2F5 (see above) and E2F4.[16] With further progress in MCC differentiation, the cytoplasmic staining disappeared and was not detectable in maturing MCCs (Figures 3C, 3C′, and 3G).

GMNC localizes exclusively to nuclei of differentiating MCCs

As introduced before, GMNC and MCIDAS are paralogous proteins that are hierarchically deployed in MCC-specific transcriptional regulation.[13,24,25] However, like MCIDAS, endogenous localization of GMNC during MCC differentiation has also not

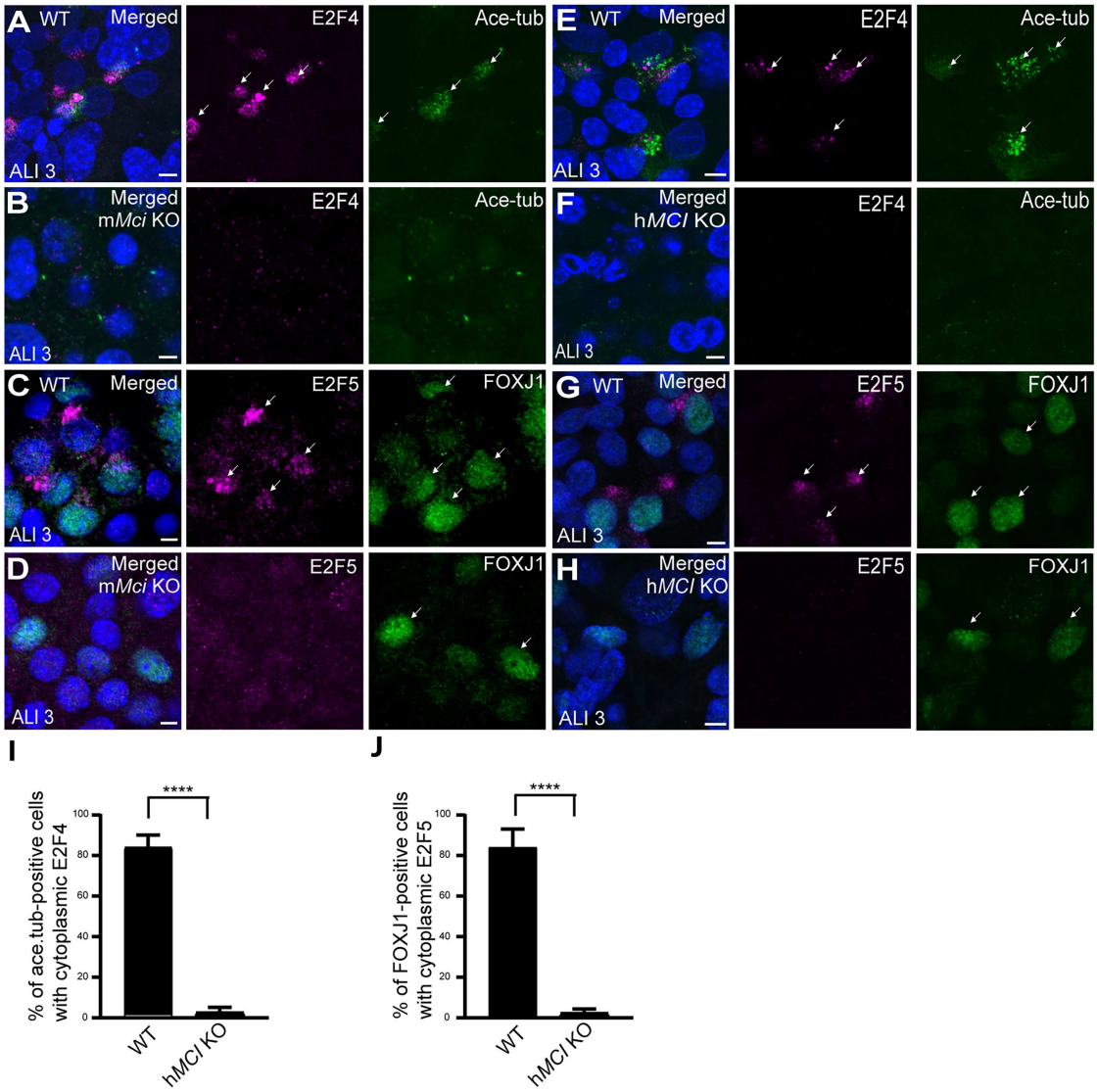


Figure 2. MCIDAS is required for cytoplasmic accumulation of E2F4/5

(A) Cytoplasmic E2F4 puncta (arrows) in differentiating MCCs derived from wild-type mTECs. MCCs were identified using acetylated tubulin staining (Ace-tub, arrows). Scale bar: 5 μm.

(B) Absence of cytoplasmic E2F4 in MCCs derived from *Mcidas* mutant (*mMci* KO) mTECs. Acetylated tubulin (Ace-tub) staining revealed absence of multicilia. Scale bar: 5 μm.

(C) Cytoplasmic E2F5 puncta (arrows) in MCCs derived from wild-type mTECs. MCCs were identified using FOXJ1 staining (arrows). Scale bar: 5 μm.

(D) Absence of cytoplasmic E2F5 in MCCs derived from *Mcidas* mutant (*mMci* KO) mTECs. MCCs were identified using FOXJ1 staining (arrows). Scale bar: 5 μm.

(E) Cytoplasmic E2F4 puncta (arrows) in MCCs derived from wild-type hESCs. MCCs were identified using acetylated tubulin (Ace-tub, arrows) staining. Scale bar: 5 μm.

(F) Absence of E2F4 cytoplasmic puncta in MCCs derived from *MCIDAS* mutant (*hMCI* KO) hESCs. Acetylated tubulin (Ace-tub) staining revealed absence of multicilia formation. Scale bar: 5 μm.

(G) Cytoplasmic E2F5 puncta (arrows) in MCCs derived from wild-type hESCs. MCCs were identified using FOXJ1 staining (arrows). Scale bar: 5 μm.

(H) Absence of cytoplasmic E2F5 puncta in MCCs derived from *MCIDAS* mutant hESCs. MCCs were identified using FOXJ1 staining (arrows). Scale bar: 5 μm. DAPI was used to highlight nuclei (blue) in (A)–(H).

(I) Percentage of cells showing cytoplasmic E2F4 in wild type and *MCIDAS* mutant (h*MCI* KO) MCCs. 3 biological replicates; 8 random microscope fields from each replicate (wild type: 50/62, 39/43, 50/64; *MCIDAS* mutant:1/19, 1/21, 0/29); ****$p < 0.0001$.

(J) Percentage of cells showing cytoplasmic E2F5 in wild type and *MCIDAS* mutant (h*MCI* KO) MCCs. 3 biological replicates; 9 random microscope fields (wild type: 43/51, 41/44, 48/64; *MCIDAS* knockout: 1/31, 0/27, 2/57); ***$p = 0.0001$. Unpaired, two-tailed Student's *t*-test was used to compare quantitative analyses. ***$p < 0.001$.

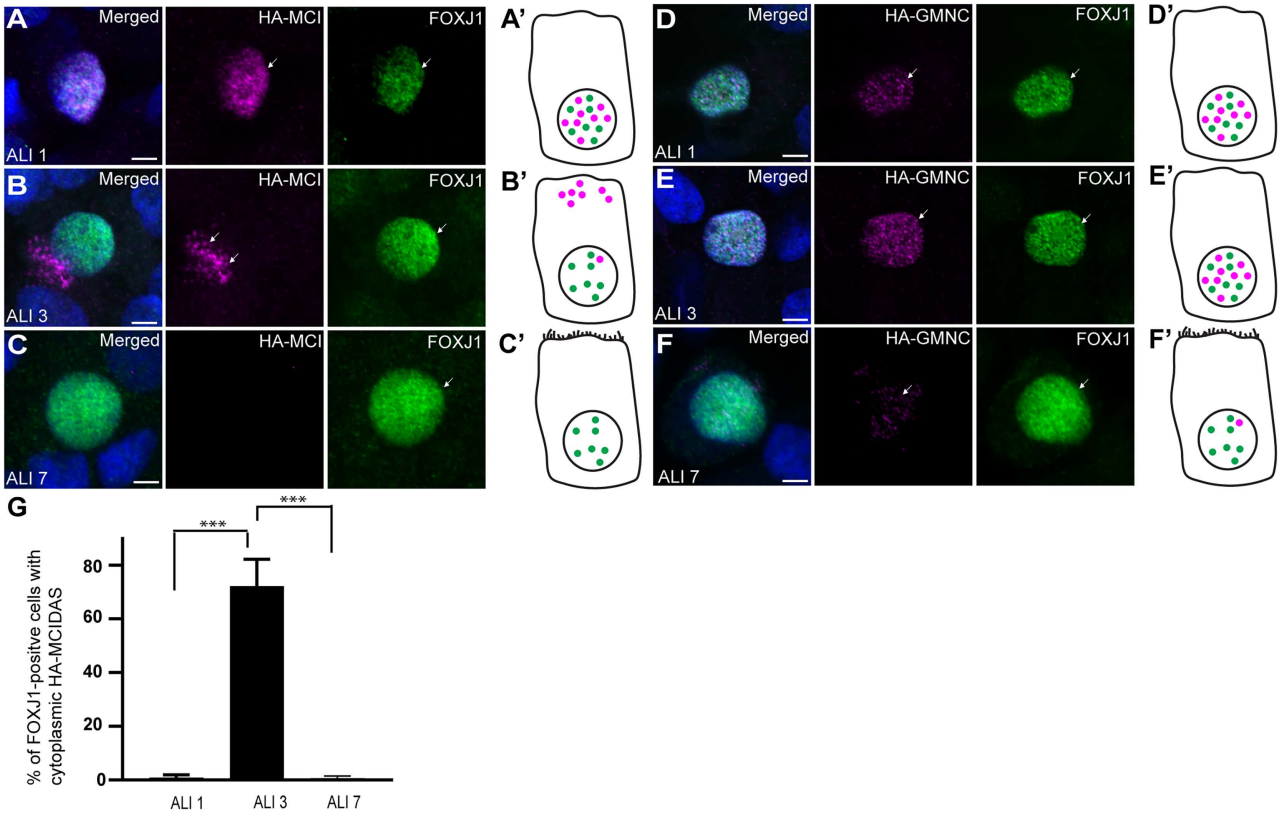


Figure 3. MCIDAS accumulates in the cytoplasm during MCC differentiation

(A) Exclusively nuclear-localized HA signal (HA-MCI, arrow) in a differentiating MCC derived from HA-MCIDAS knockin hESCs. MCCs were identified using FOXJ1 staining (arrow). Scale bar: 5 μm. (A′) Diagram illustrating nuclear HA-MCIDAS (magenta) and FOXJ1 (green).

(B) Prominent cytoplasmic HA puncta (HA-MCI, arrows) in an MCC derived from HA-MCIDAS knockin hESCs. MCCs were identified using FOXJ1 staining (arrow). Scale bar: 5 μm. (B′) Diagram illustrating cytoplasmic HA-MCIDAS (magenta) and nuclear FOXJ1 (green).

(C) Cytoplasmic HA staining was undetectable in MCC derived from HA-MCIDAS knockin hESCs. MCCs were identified using FOXJ1 staining (arrow). Scale bar: 5 μm. (C′) Diagram illustrating nuclear FOXJ1 (green).

(D and E) Exclusively nuclear-localized HA signal (HA-GMNC, arrows) in MCCs derived from HA-GMNC knockin hESCs. MCCs were identified using FOXJ1 staining (arrows). Scale bars: 5 μm. (D′ and E′) Diagrams illustrating nuclear HA-GMNC (magenta) and FOXJ (green).

(F) Nuclear HA staining was strongly reduced (arrow) in an MCC derived from HA-GMNC knockin hESCs. MCCs were identified using FOXJ1 staining (arrow). Scale bar: 5 μm.

(F′) Diagram illustrating remnant nuclear HA-GMNC (magenta) and FOXJ1 (green).

DAPI was used to highlight nuclei (blue) in (A)–(F).

(G) Percentage of cells showing cytoplasmic HA-MCIDAS in ALI day 1, 3, and 7. 3 biological replicates; 10 random microscope fields from each replicate (ALI day 1: 0/21, 0/34, 1/46; ALI day 3: 27/43, 47/67, 25/30; ALI day 7: 0/57, 1/63, 0/48). ALI day 1 vs. ALI day 3 ***$p = 0.0003$, ALI 3 vs. ALI day 7 ***$p = 0.0003$. Unpaired, two-tailed Student's *t*-test was used to compare quantitative analyses. ***$p < 0.001$.

been documented. We also engineered the 3×HA tag at the N terminus of GMNC in H9 cells (after the first 3 amino acids; hereafter referred to as HA-GMNC; Figures S5A–S5F) and monitored localization of the protein during MCC differentiation (like MCIDAS, we have shown earlier that such an N-terminal tag does not compromise GMNC function[15]). Unlike HA-MCIDAS, HA-GMNC remained exclusively nuclear localized (Figures 3D, 3D′, 3E, and 3E′) and disappeared from maturing MCCs (Figures 3F and 3F′).

Bioinformatic analysis of mouse and human MCIDAS revealed putative nuclear export sequences (NESs) (two in mouse and three in the human protein; Figures S6A–S6C), as well as a nucleolar localization sequence (NoLS) in human MCIDAS (Figure S6D; there seems to be a NoLS in mouse MCIDAS as well, but it did not cross the significance level in our analysis). The presence of a NoLS in place of a nuclear localization sequence (NLS) suggests that the former may possess a nuclear localization function.[26] In contrast to MCIDAS, for GMNC, we could only identify an NLS in both the human and the mouse protein, but no NES (Figure S6E). Machine learning classification by NESmapper[27] and 3D structure modeling of MCIDAS NES top motifs suggest that they are helix leucine-rich motifs of the class 1a/b types that can bind to the exportin CRM1 for nuclear export (Figures S6B and S7A–S7D).

Cytoplasmic MCIDAS clusters in the vicinity of and colocalizes with E2F4, E2F5, and centriolar proteins

We next analyzed the distribution of centriolar biogenesis proteins and core centriolar components vis-a-vis cytoplasmic

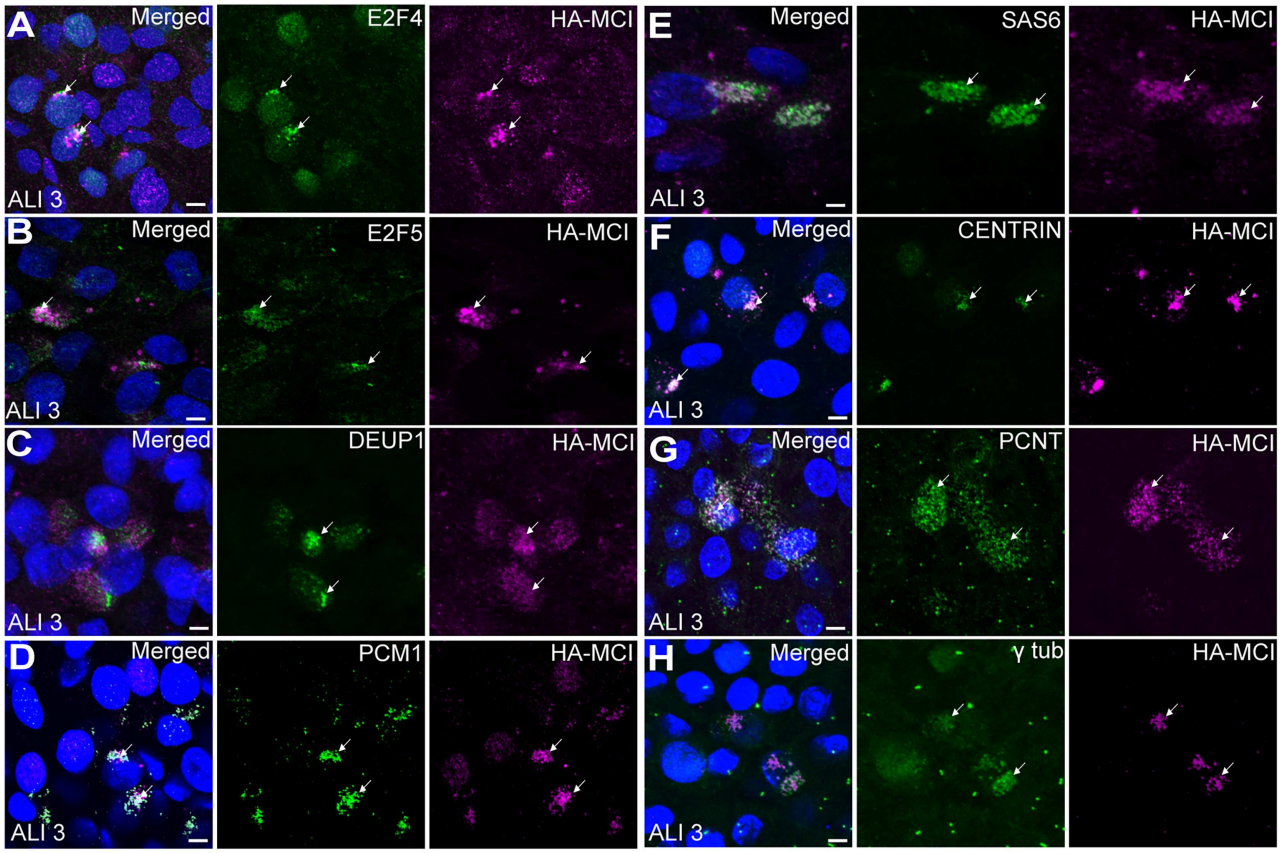


Figure 4. Cytoplasmic MCIDAS accumulates in the vicinity of and colocalizes with E2F proteins and other centriolar components

(A) E2F4 (arrows) colocalized (white signal in merged panel, arrows) with cytoplasmic HA-MCIDAS (HA-MCI, arrows) in differentiating MCCs derived from HA-MCIDAS knockin hESCs. Scale bar: 5 μm.

(B) E2F5 (arrows) colocalized (white signal in merged panel, arrow) with cytoplasmic HA-MCIDAS (HA-MCI, arrows) in MCCs derived from HA-MCIDAS knockin hESCs. Scale bar: 5 μm.

(C) DEUP1 (arrows) clustered in the vicinity of cytoplasmic HA-MCIDAS (arrows) in MCCs derived from HA-MCIDAS knockin hESCs. Scale bar: 5 μm.

(D) PCM1 (arrows) colocalized (white signal in merged panel, arrows) with cytoplasmic HA-MCIDAS (arrows) in MCCs derived from HA-MCIDAS knockin hESCs. Scale bar: 5 μm.

(E) SAS6 (arrows) clustered in the vicinity of cytoplasmic HA-MCIDAS (arrows) in MCCs derived from HA-MCIDAS knockin hESCs. Scale bar: 5 μm.

(F) CENTRIN (arrows) colocalized (white signal in merged panel, arrows) with cytoplasmic HA-MCIDAS (arrows) in MCCs derived from HA-MCIDAS knockin hESCs. Scale bar: 5 μm.

(G) PCNT (arrows) colocalized (white signal in merged panel, arrow) with cytoplasmic HA-MCIDAS (arrows) in MCCs derived from HA-MCIDAS knockin hESCs. Scale bar: 5 μm.

(H) γ-tub (arrows) clustered in the vicinity of cytoplasmic HA-MCIDAS (arrows) in MCCs derived from HA-MCIDAS knockin hESCs. Scale bar: 5 μm.

DAPI was used to highlight nuclei (blue) in (A)–(H).

HA-MCIDAS. We first examined E2F4 and E2F5 and found a significant proximity, including overlap, in the distribution between these two E2F proteins and cytoplasmic HA-MCIDAS puncta (Figures 4A and 4B). In addition, we analyzed DEUP1 and PCM1 (a component of the centriolar satellites) and also other centriolar and pericentriolar material proteins like SAS6, CENTRIN, PCNT, and γ-tubulin. Since many of these proteins have already been shown to accumulate close to or to colocalize with cytoplasmic E2F4[16] and E2F5 (this study), we expected to see a similar pattern with HA-MCIDAS as well. Accordingly, we found that all these proteins accumulated in the vicinity or colocalized with HA-MCIDAS (Figures 4C–4H), supporting the notion that cytoplasmic MCIDAS could facilitate centriole biogenesis in MCCs.

MCIDAS loss blocks initiation of centriole biogenesis

We have previously reported that in the absence of MICDAS, MCC precursors are specified, but they fail to generate multiple basal bodies.[15] Consistent with this, we had demonstrated that *Deup1* expression was markedly reduced and DEUP1-positive deuterosome numbers were also significantly less.[15] To further define the precise defect in centriole biogenesis, we investigated the status of PCM1 and SAS6. PCM1 is a centriolar satellites protein, which is also a component of the fibro-granular material but is not itself essential for MCC centriole biogenesis,[28–30] while SAS6 is the essential seeding component for procentriole formation and constitutes the cartwheel structure around which the 9 triplet microtubule blades of the procentriole are assembled.[17,31]

We observed that the cytoplasmic accumulation of both these proteins was almost undetectable on MCIDAS loss (Figures 5A–5J), indicating an impairment at the inception of centriole assembly. While *Pcm1* expression was unaffected by the loss of MCIDAS (Table S1), there was a slight reduction in *Sass6* transcription, but protein levels appeared unaffected (Table S2; Figures S3F and S3G). Thus, as argued above for the E2F proteins, our inability to detect PCM1 and SAS6 in MCIDAS-deficient cells by immunofluorescence is unlikely to arise from an effect on their expression, but rather from their failure to undergo MCIDAS-dependent cytoplasmic accumulation.

Nuclear export of MCIDAS is necessary for centriole amplification

To establish that the inability of MCIDAS mutant MCC precursors to amplify centrioles arises not only from dampening of the MCC transcriptional program but also from a failure to nucleate centriole assembly by cytoplasmic MCIDAS, we next manipulated its nuclear export. We overexpressed wild-type human MCIDAS in HEK293T cells and found both nuclear and cytoplasmic localization (Figure S8A). However, on treatment with leptomycin B, a specific inhibitor of CRM1-dependent nuclear export,[32] overexpressed MCIDAS was exclusively nuclear (Figure S8B). Leptomycin B also inhibited the cytoplasmic accumulation of HA-tagged endogenous MCIDAS in differentiating hESC-derived MCCs (Figures 6A and 6B).

To establish the functionality of the MCIDAS NES motifs, we mutated the second of the three leucines (L_{238}) in the second NES to aspartic acid (D_{238}) and examined the MICDAS localization in 293T cells. We selected the second NES since the first and the second NESs are conserved in mouse as well as human MCIDAS, and the second NES was associated with a higher score in our NES prediction analysis (Figures S6A–S6C). Similar to the effect of leptomycin B, we found that the L > D mutant version was exclusively nuclear localized (Figure S8C). Overexpression of GMNC and MCIDAS in HEK293T cells can induce the MCC gene expression signature.[13,15] Hence, using Quantitative reverse-transcription PCR (RT-qPCR), we analyzed the induction of well-established MCIDAS target genes—*FOXJ1*, *CCNO*, *CDC20B*, and *DEUP1*—and found that on leptomycin B treatment as well as mutation of the NES, MCIDAS remained transcriptionally competent (Figure S8D). We then investigated if inhibition of nuclear export could impair centriole biogenesis by treating wild-type mTEC cultures with leptomycin B. Leptomycin B was administered at the end of ALI day 1 and maintained till the end of ALI day 2.5, when the cells were fixed for immunostaining. In controls, we noted a large number of RFX2-positive MCC precursors with a cloud of cytoplasmic SAS6 staining signifying centriole amplification (Figures 6C and 6G). By contrast, in leptomycin B-treated cultures, most RFX2-positive cells were devoid of cytoplasmic SAS6 (Figures 6D and 6G). We noted, however, that RFX2-positive MCC precursors were also reduced on leptomycin B treatment (Figure 6D). Since leptomycin B is a generic inhibitor of nuclear export, nuclear-cytoplasmic translocation of many other proteins is likely affected, which could result in this effect.

To more incisively dissect the cytoplasmic function of MCIDAS, we used lentivirus-mediated overexpression of wild-type and the L > D mutant in *Mcidas* mutant mTECs to assess their abilities to promote centriole amplification and rescue MCC formation (Figures S8E and S8F). Using this strategy, we have previously shown that overexpression of wild-type MCIDAS can induce supernumerary MCCs in wild-type mTEC cultures and effectively rescue MCC formation in the *Mcidas* mutant.[15] Wild-type MCIDAS overexpression induced high levels of RFX2, and these cells were associated with a cloud of cytoplasmic E2F4, signifying centriole amplification (Figures 6E, 6H, and S8E). By contrast, even though L > D mutant MCIDAS induced high levels of RFX2 expression, cytoplasmic E2F4 was never observed (Figures 6F, 6H, and S8F). Furthermore, in line with our previous work,[15] while wild-type MCIDAS restored MCC formation, L > D mutant MCIDAS failed to do so (Figures S8G and S8H). In sum, these findings provide strong support for the requirement of cytoplasmic MCIDAS in centriole biogenesis during MCC differentiation, independent of its role in programming gene expression.

Cytoplasmic MCIDAS-E2F4/5 mediates *de novo* centriole biogenesis in MCCs

To discern whether MCIDAS-E2F4/E2F5 complex is part of the deuterosomal or parental centriole-dependent pathways or represents the *de novo* mechanism of centriole biogenesis, we finally examined cytoplasmic E2F4 localization in differentiating mouse ependymal MCCs devoid of deuterosomes and the parental centrioles. As described previously,[7] this was achieved through the treatment of ependymal MCC precursors from *Deup1* mutant mice with centrinone, an inhibitor of PLK4, the key kinase that initiates centriole biogenesis in proliferating cells as well as in post-mitotic MCC precursors.[33,34] We stained for SAS6 to identify cells undergoing centriole amplification from the earliest stages. We also stained for CEP164, a distal appendage protein that is a marker of mature centrioles,[35] to identify cells bereft of parental centrioles. Consistent with our observations with airway MCCs, we found that in ependymal MCCs as well a cloud of E2F4 formed around the parental centrioles, and SAS6-positive procentrioles localized within this region (Figure 7A), indicating that cytoplasmic accumulation of the E2F proteins is a shared feature of MCC differentiation in distinct epithelial lineages. In cells deprived of parental centrioles and/or deuterosomes, where centriole amplification is unaffected, procentrioles have been observed to emerge from a PCNT-rich fibro-granular region.[7] In centrinone-treated *Deup1* mutant ependymal MCCs, we analyzed the following three situations: cells devoid of parental centrioles but containing deuterosomes, cells retaining parental centrioles but deprived of deuterosomes, and cells in which neither parental centrioles nor deuterosomes were present. The E2F4 staining pattern was indistinguishable among these three experimental conditions and controls: in each instance, a cytoplasmic cloud of E2F4 was apparent and found to cradle the emerging procentrioles (Figures 7B–7D). We conclude that while cytoplasmic accumulation of E2F4 requires MCIDAS, it is independent of the parental centrioles and deuterosomes. Importantly, because the cytoplasmic localization of E2F4 as well as MCIDAS is essential for centriole amplification in MCCs (Mori et al.[16] and this study), these findings incriminate the cytoplasmic MCIDAS-E2F4/5

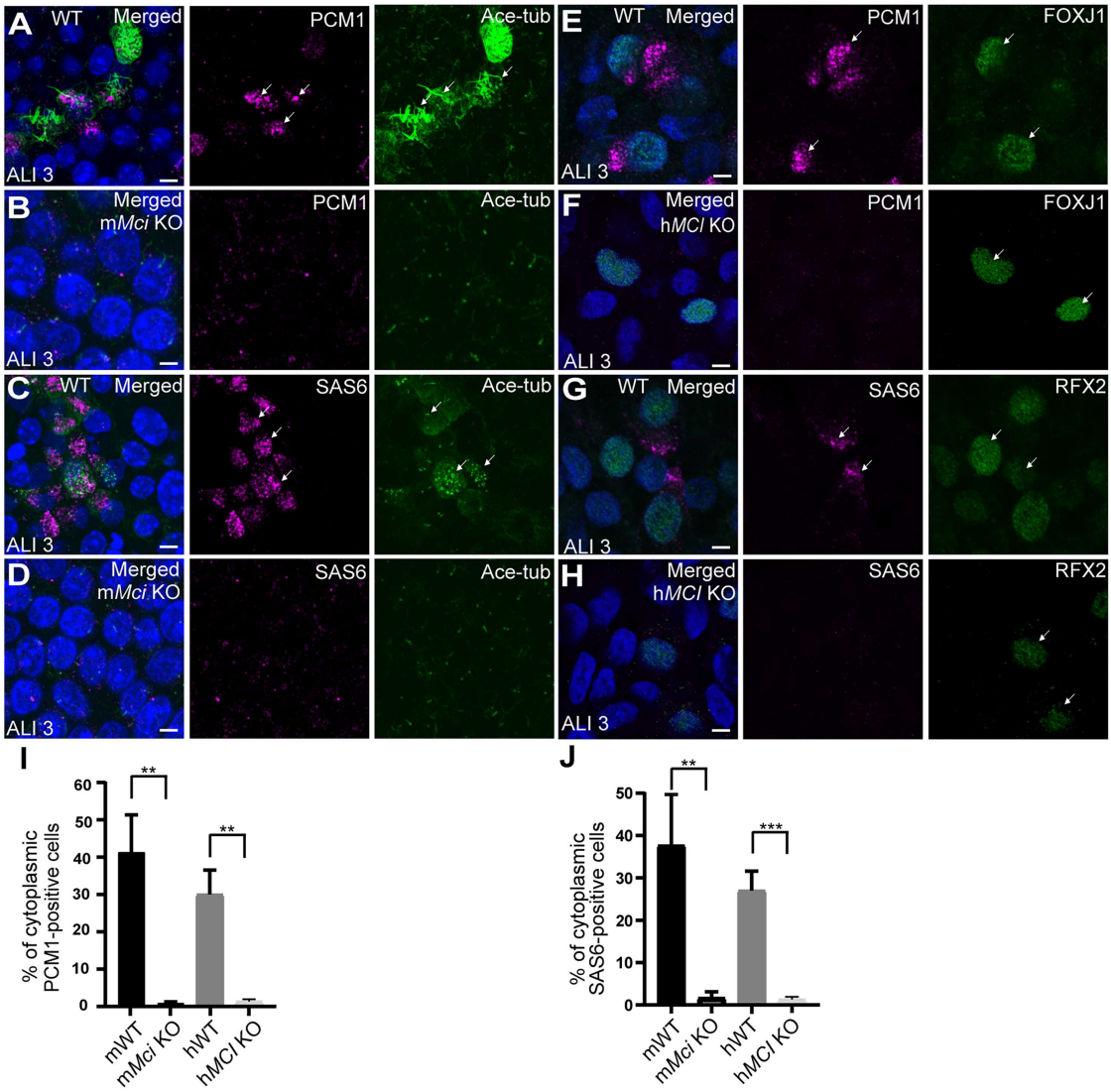


Figure 5. Earliest stage of centriole biogenesis is inhibited in MCIDAS deficient MCC precursors

(A) PCM1 accumulation in the cytoplasm (arrows) of differentiating MCCs derived from wild-type mTECs. MCCs were identified using acetylated tubulin staining (Ace-tub, arrows). Scale bar: 5 μm.

(B) PCM1 puncta were undetectable in MCCs derived from *Mcidas* mutant (*mMci* KO) mTECs. Acetylated tubulin (Ace-tub) staining revealed absence of multicilia. Scale bar: 5 μm.

(C) Cytoplasmic SAS6 (arrows) in MCCs derived from wild-type mTECs. MCCs were identified using acetylated tubulin staining (Ace-tub, arrows). Scale bar: 5 μm.

(D) Cytoplasmic SAS6 was undetectable in MCCs derived from *Mcidas* mutant mTECs. Acetylated tubulin (Ace-tub) staining revealed absence of multicilia. Scale bar: 5 μm.

(E) Cytoplasmic PCM1 accumulation (arrows) in MCCs derived from wild-type hESCs. MCCs were identified using FOXJ1 staining (arrows). Scale bar: 5 μm.

(F) PCM1 was undetectable in MCCs derived from *MCIDAS* mutant hESCs. MCCs were identified using FOXJ1 staining (arrows). Scale bar: 5 μm.

(G) Cytoplasmic SAS6 (arrows) in MCCs derived from wild-type hESCs. MCCs were identified using RFX2 staining (arrows). Scale bar: 5 μm.

(H) SAS6 was undetectable in MCCs derived from *MCIDAS* mutant hESCs. MCCs were identified using RFX2 staining (arrows). Scale bar: 5 μm.

DAPI was used to highlight nuclei (blue) in (A)–(H).

(I) Percentage of MCCs showing cytoplasmic PCM1 in wild type (mWT) and *Mcidas* mutant (m*Mci* KO) mTECs and wild type (hWT) and *MCIDAS* mutant (h*MCI* KO) hESCs. 3 biological replicates; 8 random microscope fields from each replicate (mWT: 31/99, 43/121, 36/71; m*Mci* KO: 1/134, 2/161, 1/112; $^{**}p = 0.0017$; hWT: 28/121, 36/101, 30/97; h*MCI* KO: 1/110, 2/124, 1/91; $^{**}p = 0.0015$). Unpaired, two-tailed Student's *t*-test was used to compare quantitative analyses. $^{**}p < 0.01$.

(J) Percentage of MCCs showing cytoplasmic SAS6 in wild-type (mWT) and *Mcidas* mutant (m*Mci* KO) mTECs and wild-type (hWT) and *MCIDAS* mutant (h*MCI* KO) hESCs. 3 biological replicates; 8 random microscope fields from each replicate (mWT: 53/108, 42/170, 35/89; m*Mci* KO: 1/97, 4/121, 1/142; $^{**}p = 0.0068$; hWT: 32/145, 30/98, 22/79; h*MCI* KO: 2/103, 1/97, 1/141; $^{***}p = 0.0006$). Unpaired, two-tailed Student's *t* test was used to compare quantitative analyses. $^{**}p < 0.01$ and $^{***}p < 0.001$.

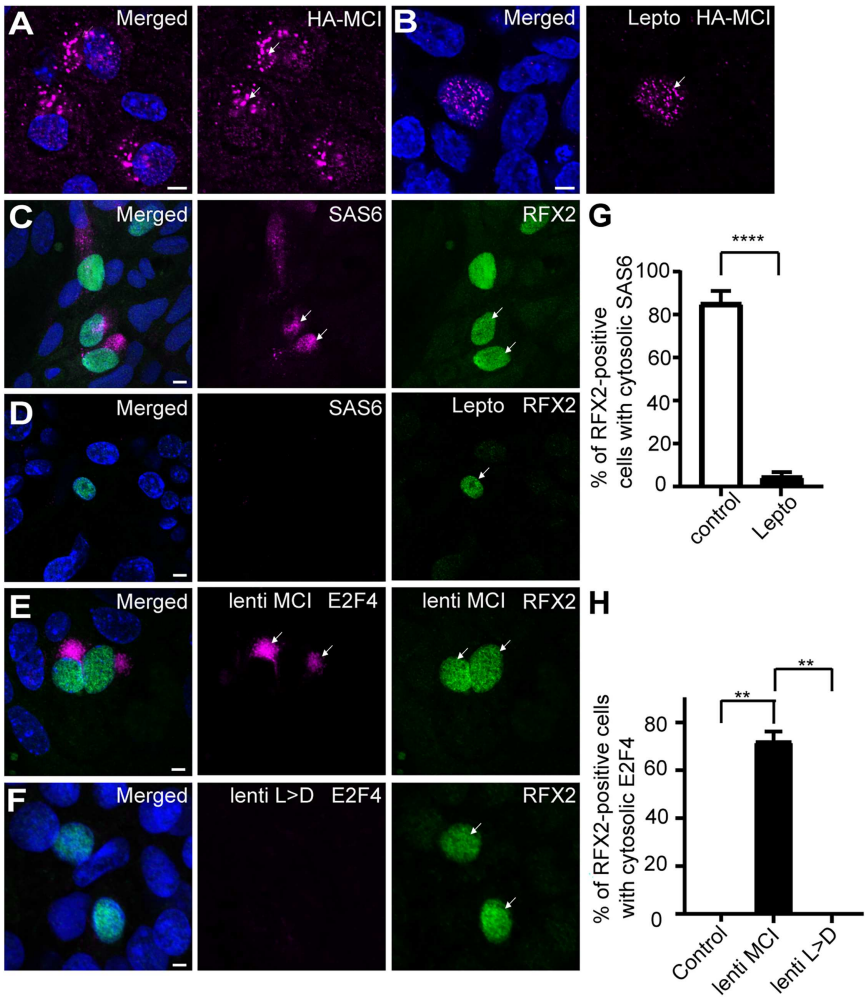


Figure 6. CRM1-mediated cytoplasmic accumulation of MCIDAS is required for centriole amplification in MCCs

(A) Cytoplasmic accumulation of HA-tagged endogenous MCIDAS (arrows) in differentiating hESC-derived MCCs. Scale bar: 5 μm.

(B) Leptomycin B caused nuclear retention of HA-tagged endogenous MCIDAS (arrow) in hESC-derived MCCs. Scale bar: 5 μm.

(C) Cytoplasmic accumulation of SAS6 (arrows) in wild-type mTECs. MCCs were identified using RFX2 staining (arrows). Scale bar: 5 μm.

(D) Cytoplasmic SAS6 was absent from MCCs on leptomycin B treatment. MCCs were identified using RFX2 staining (arrow). Scale bar: 5 μm.

(E) Lentivirus-mediated wild-type MCIDAS overexpression rescued cytoplasmic E2F4 accumulation (arrows) in *Mcidas* mutant mTECs. MCCs were identified using RFX2 staining (arrows). Scale bar: 5 μm.

(F) Lentivirus-mediated L > D mutant MCIDAS overexpression failed to rescue cytoplasmic E2F4 accumulation in *Mcidas* mutant mTECs. MCCs were identified using RFX2 staining (arrows). Scale bar: 5 μm.

DAPI was used to highlight nuclei (blue) in (A)–(F).

(G) Percentage of RFX2-positive mTECs showing cytoplasmic SAS6 with and without leptomycin B treatment. 3 biological replicates; 10 random microscope fields from each replicate (control: 89/99, 66/84, 40/46; leptomycin B treatment: 3/41, 1/36, 2/52; ****$p < 0.0001$).

(H) Percentage of RFX2-positive *Mcidas* mutant mTECs showing cytoplasmic E2F4 after lentivirus-mediated overexpression of wild type or L > D mutant MCIDAS. 2 biological replicates; 10 random microscope fields from each replicate (control: 0/46, 1/53; wild-type MCIDAS lentivirus-infected cells: 36/64, 61/78; L > D mutant MCIDAS lentivirus-infected cells: 0/46, 1/56. Control vs. wild-type MCIDAS lentivirus-infected cell: *p = 0.0268; Wild-type MCIDAS lentivirus infected cell vs. L > D mutant MCIDAS lentivirus infected cell: *p = 0.0269). Unpaired, two-tailed Student's t-test was used to compare quantitative analyses. *$p < 0.05$.

complex as an important mediator in organizing the *de novo* pathway.

Whole transcriptomics sequencing of MCIDAS mutant mTECs identifies its target genes

Delineating the spectrum of MCIDAS target genes will provide further information regarding the mechanism of centriole amplification and multiciliation in MCCs. Toward this end, gene expression profiles of *Mcidas* mutant mTECs at ALI days 0, 3, and 5 were compared against the wild type to identify significantly dysregulated genes (absolute log2-fold change > 1 and adjusted false discovery rate [FDR] p value < 0.1) (Table S2). The majority of dysregulated genes represent those that were down-regulated in the mutant relative to the wild type, indicating that MCIDAS functions mainly in transcriptional activation (Table S2). 301 genes were down-regulated in both days 3 and 5 of differentiation. Among these, 276 genes showed sequential down-regulation, which includes *Mcidas* as noted before,[15] vindicating successful inactivation of the gene. 1,097 genes (827 protein coding) showed significant down-regulation at day 5, expression of which remained unchanged until day 3. On the other hand, 18 genes were found to be significantly up-regulated in both day 3 and day 5 (Table S2), among which 12 genes, including *Gmnc*, showed sequential up-regulation, again following a trend observed previously.[15] Notably, as mentioned before, *E2f4* and *E2f5* did not exhibit significant changes in expression on MCIDAS loss (Table S1).

Overexpression studies have been used in previous attempts to determine the transcriptional landscape regulated by MCIDAS.[12,14,36,37] We compared our findings with two earlier studies to gauge the extent of congruence: (1) Stubbs et al.,[36] where Mcidas was overexpressed in *Xenopus* embryos and gene expression changes assessed with microarrays, and (2) Ma et al.,[12] where Mcidas was again overexpressed in *Xenopus* embryos, either alone or together with a dominant-negative version of E2f4 (to mimic Mcidas loss), and RNA-seq was used to examine gene expression changes. Among 95 genes up-regulated after Mcidas overexpression (having mouse orthologs)

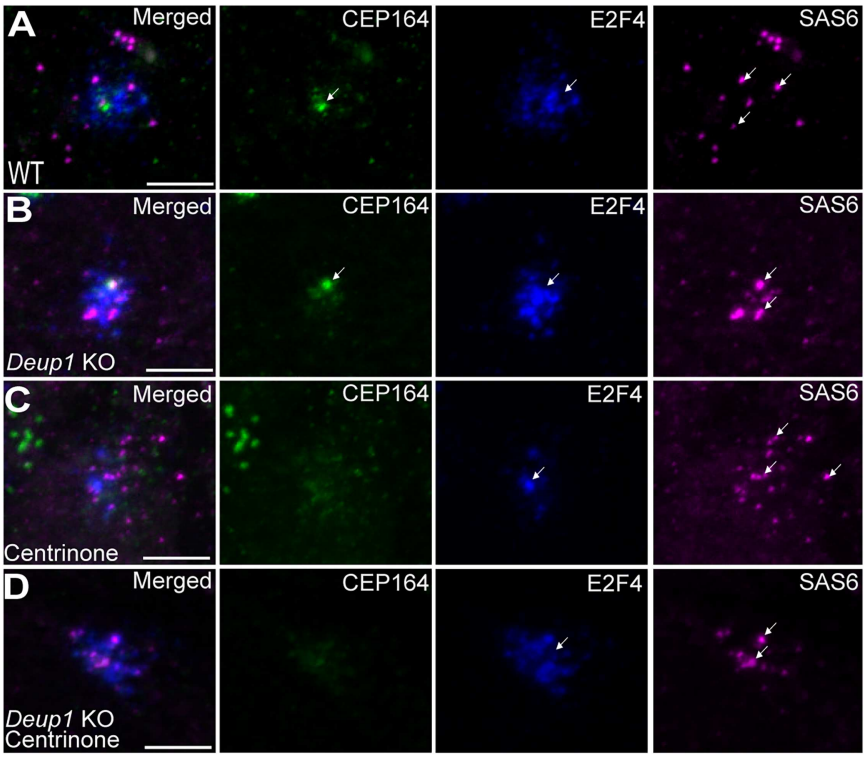


Figure 7. E2F4 localization to the PCM is unaffected in the absence of parental centrioles and deuterosomes

(A) E2F4 (arrow) accumulated around parental centrioles (CEP164 staining, arrow) and SAS6-positive procentrioles (arrows) in differentiating wild-type mouse ependymal MCCs. Scale bar, 5 μm.

(B) E2F4 cytoplasmic localization (arrow) and SAS6-positive procentrioles (arrows) were unaffected in MCCs derived from ependymal cells of *Deup1* mutant (*Deup1* KO) mice. Scale bar, 5 μm.

(C) E2F4 cytoplasmic localization (arrow) and SAS6-positive procentrioles (arrows) were unaffected in centrinone-treated wild-type mouse ependymal MCCs. Scale bar, 5 μm.

(D) E2F4 cytoplasmic localization (arrow) and SAS6-positive procentrioles (arrows) were unaffected in centrinone-treated MCCs derived *Deup1* mutant (*Deup1* KO) mouse ependymal cells. Scale bar, 5 μm.

reported by Stubbs et al., 62 genes (65%) were found to be down-regulated in our dataset (Table S2). On the other hand, among 378 genes reported to be up-regulated on overexpression of Mcidas relative to Mcidas with dominant-negative E2f4 (having mouse orthologs) reported by Ma et al., 92 (24%) were found to be down-regulated in our analysis (Table S2). Thus, we have identified a large collection of genes that have previously not been linked to MCIDAS function.

MCIDAS targets are replete with MCC-specific proteins

The down-regulated protein coding and noncoding genes (except pseudogenes and rRNAs) were mapped to the Gene Ontology (GO) database. We could identify 107 GO-biological processes (BP) to be significantly ($q < 0.05$) enriched, which were further clustered into 21 modules based on gene set similarities. 496 genes were enriched in these modules. Multiple GO-BP terms relating to microtubule, centriole, and cilia biogenesis, assembly, and functions were markedly enriched, consistent with earlier overexpression studies[12,14,36,37] and the derailment of centriole and cilia biogenesis on MCIDAS loss[15,21,36] (Figures S9 and S10; Table S3). *Rp1* showed the highest down-regulation and was present in 6 modules related to ciliary function, which we could confirm using RT-qPCR analysis (Figure S11). *Rp1* encodes a ciliary protein studied extensively in the context of the immotile photoreceptor connecting cilium and its disorder, retinitis pigmentosa.[38,39] How RP1 functions in MCC motile cilia is not known, although the protein has been detected on MCC ciliary axonemes.[40] *Dnai3*, *Dnah10*, *Dnai2*, *Dnai1*, *Ncan*, *Ccdc170*, *Cibar*, *Ttc29*, *Foxj1*, *Foxn4*, etc., related to cilium organization, microtubule-based movement, and transport and axoneme assembly, were also among the top down-regulated genes. Likewise, *Ccdc78*, *Mdm1*, and *Deup1*—genes associated with centriole formation—were also strongly down-regulated. Furthermore, genes encoding the atypical cyclin, CCNO, as well as the transcription factor E2F7, which are thought to orchestrate a pseudo cell cycle that uncouples centriole amplification from DNA synthesis and mitosis in differentiating MCCs, also figures among MCIDAS targets.[41–43]

859 down-regulated genes did not map to any of the significant GO-BP modules. We further mapped these to CellMarker (version 2024), and 414 genes could be mapped to 469 human/mouse cell types (Table S4). Of the detected cell types, five—ciliated cell trachea (mouse), ciliated epithelial cell esophagus (human), ciliated cell lung (human), ependymal cell brain (mouse), and ciliated cell lung (mouse)—were significantly enriched ($q < 0.000068$, odds ratio > 5.4) and include 201 genes with *Nhlrc4*, *Rsph14*, *Ccdc60*, *Ccdc170*, *Sntn*, *Cdhr3*, *Ect2l*, *Lrrc71*, *Ccdc81* among those that were highly down-regulated. The human orthologs were also mapped to Human Protein Atlas (HPA) (525 could be mapped; see Table S5), and we found that 98 genes are expressed in ciliated cells, while 96 are expressed in spermatocytes, early spermatids, and late spermatids, signifying links to cilia.

Finally, we examined the representation of human disease genes among MCIDAS targets. Not surprisingly, the vast majority of PCD-causing genes figures in the list (Table S2), underscoring that this collection is a valuable resource for the discovery of additional loci mutated in motile ciliopathies. The list also contains genes associated with other diseases where the contribution of MCC dysfunction to the pathological state remains to be determined (Table S2).

DISCUSSION

MCCs are an evolutionarily ancient cell type—many aquatic ciliated protozoans like *Paramecium* and *Tetrahymena* are in

essence MCCs, and in metazoans, MCCs also function in driving organismal locomotion, such as swimming in many invertebrate larval forms and the gliding locomotion of flatworms.[44,45] Although cilia and flagella are rather restricted in distribution among plants and fungi, multiciliated spermatozoa have been reported from ferns and gymnosperms (sperm of cycad *Zamia* have about 20,000 motile cilia).[46] In addition, MCCs within various tissues and organs of animals function in generating vectorial fluid flow over epithelial surfaces, and this seems to be the sole function of these cells in the vertebrates.[1,45] Impairment in MCC-driven fluid flow can cause severe disease in humans as exemplified by the pathological consequences in patients with PCD and RGMC.[2]

We have now discovered that the MCC "master" transcriptional regulator, MCIDAS, not only activates centriole and ciliary genes for MCC differentiation but also generates a cytoplasmic organizing center for *de novo* assembly of the massive numbers of centrioles required for multiciliation. By visualizing endogenous MCIDAS in differentiating human MCCs, we discovered that besides localizing in the nucleus, the protein underwent cytoplasmic translocation, clustering with forming procentrioles. In addition, cytoplasmic MCIDAS was found in the vicinity of and colocalized with E2F4/5, and its loss abolished the cytoplasmic accumulation of these proteins. Yet, in MCC precursors devoid of parental centrioles and deuterosomes, where centriole amplification can proceed unabated, cytoplasmic accumulation of E2F4 remained unaffected. Furthermore, MCIDAS loss inhibited centriole assembly at the inception stage. Thus, our findings resolve a long-standing enigma regarding the peculiar ability of post-mitotic MCC precursors to amplify hundreds of centrioles and establish that the process is directed by a linked transcriptional and organelle assembly program, orchestrated by MCIDAS.

Because MCIDAS complexes with E2F proteins for transcription,[12,15] their cytoplasmic accumulation could be mediated by the cytoplasmic translocation of MCIDAS. However, in MCIDAS mutant MCCs, where cytoplasmic clustering of the E2F proteins is not observed, we did not notice their perdurance in the nucleus (Figure 2). This suggests that E2F4/5 nuclear export is not mediated by MCIDAS, but cytoplasmic MCIDAS is critical for their cytoplasmic clustering. In fact, CRM1-mediated nuclear export of the E2F proteins has been well documented.[47,48] Likewise, based on the presence of NESs, we have proposed a CRM1-mediated nuclear export pathway also for MCIDAS: the nuclear export of MCIDAS is leptomycin B sensitive and disruption of the central NES motif can block nuclear export. Consistent with this, leptomycin B inhibited MCC formation in cultured mTECs, and the L > D NES mutant MCIDAS protein failed to rescue centriole amplification and MCC formation in mTECs derived from *Mcidas* mutant mice. Together, these data are consistent with the model that cytoplasmic MCIDAS directs cytoplasmic accumulation of the E2F proteins, and together they organize a *de novo* centriole assembly center in differentiating MCCs. But once in the cytoplasm, how does the MCIDAS-E2F4/5 complex initiate centriole biogenesis? As the E2F proteins can associate with DEUP1 and SAS6,[18] we argue that centriole biogenesis is initiated via interactions with these, and possibly other, centriole biogenesis components. Furthermore, there is emerging evidence that MCIDAS can interact with PLK4 when the proteins are overexpressed in cycling cells.[23] Using *in silico* modeling, we could confirm this interaction, and we also found that it is the second NES motif of MCIDAS that mediates interaction with PKL4 (Figures S7E and S7F). Since PLK4 phosphorylates and recruits core centriolar components like STIL and SAS6 to parental centrioles for procentriole formation during mitosis,[49] MCIDAS-PLK4 association could also operate during MCC differentiation and represent the mechanistic basis by which centriole assembly is triggered by cytoplasmic MCIDAS.

Using transcriptomics of *Mcidas* mutant mTEC-derived MCCs, we have also generated a comprehensive list of MCIDAS target genes. This list represents genes dysregulated in a stable genetic loss-of-function condition, as opposed to earlier MCIDAS overexpression-based assays.[12,14,36,37] Nevertheless, consistent with earlier studies, the most prominent differentially expressed gene modules contain GO-BP terms pertaining to cilia and centriole biogenesis and function. Although many of these genes have already been investigated in centrioles and cilia in MCCs, many others have been shown to function in these organelles elsewhere, but their activities in MCCs have not been investigated, such as RP1. Also, the cilia and centriole GO-BP-enriched genes include the majority of those linked with PCD, reiterating the importance of MCC abnormalities in the pathology of this disorder. In addition, we have been able to identify many genes that have previously not been identified as MCIDAS targets, and interrogation of their transcript/protein distribution in existing databases revealed that a significant proportion are indeed MCC specific, possibly encoding many "novel" centriole and cilia components. Rather interestingly, on parsing through the topmost down-regulated genes, we have noted that there could be several MCIDAS targets with MCC-specific functions that may not involve cilia or centrioles. For instance, *Cdhr3*, encoding a Cadherin-related protein, figures among the most down-regulated genes (Table S2; Figure S11). CDHR3 localizes to the apical cytoplasm of developing MCCs and has been implicated in asthma, chronic obstructive pulmonary disease, as well as rhinovirus infections.[50,51] CDHR3 loss does not affect multiciliation but impairs rhinovirus tropism to MCCs,[50] highlighting that MCIDAS targets may have roles beyond cilia and centrioles, and their dysfunction can lead also to common MCC-based diseases like asthma and viral infections, besides the rare disorders like PCD and RGMC.

Limitations of the study

Although our findings greatly favor *de novo* centriole biogenesis in MCCs, with cytoplasmic MCIDAS being an important player in organizing this pathway, the parental centrioles and deuterosomes, while dispensable, could make auxiliary contributions. When available, they could serve as alternative procentriole nucleation centers, ensuring spatial and temporal coordination as well as fidelity in a process that involves the generation of a large number of highly structured organelles, the centrioles, within a short time frame. Thus, our study prompts a more careful evaluation of centriole amplification in MCCs that are devoid of the deuterosomes and parental centrioles, for example, using live imaging with fluorophore-tagged proteins, to uncover if

certain elements in the dynamics of centriole amplification can be perturbed in their absence.

RESOURCE AVAILABILITY

Lead contact

Further information and requests for resources and reagents should be directed to and will be fulfilled by the lead contact, Dr. Sudipto Roy (sudipto_roy@a-star.edu.sg).

Materials availability

All CRISPR/Cas9 edited hESC lines generated in this study will be made available upon request.

Data and code availability

- This study did not generate any original code.
- Raw RNA-seq data of *Mcidas* mutant and wild-type mTECs have been deposited at the European Nucleotide Archive (ENA) with accession number ENA: PRJEB77353.

ACKNOWLEDGMENTS

We thank the IMCB central imaging and the A*STAR and Lee Kong Chian School of Medicine fluorescence-activated cell sorting (FACS) facilities for microscopy and cell sorting, respectively; M. Vignuzzi from the IDLabs (A*STAR) for support to computational analyses; and T. Stracker, D. Rayamajhi, S. Heng, and M.G. Jones for discussion and insightful comments on the manuscript. This work was funded by the Agency for Science, Technology and Research (A*STAR) of Singapore-University of Sheffield Research Attachment Program (ARAP) graduate studentship to C.T.J. and funds from A*STAR and a National Medical Research Council (NMRC) of Singapore grant (OFIRG19nov-0037) to S.R. S.R. dedicates this paper in memory of his mother, Sudha Roy, who passed away during the revision process, for her support and encouragement to pursue a career in science.

AUTHOR CONTRIBUTIONS

Conceptualization, H.L., K.J.G., and S.R.; investigation, H.L., K.J.G., E.K.T, C. T.J., and A.-R.B.; analysis and visualization, H.L., K.J.G., E.K.T., C.T.J., A.G., A.-R.B., P.A.L., S.M.-S., N.K.B., C.B.D., N.R.D., A.M., and S.R.; writing, K.J.G., A.G., N.K.B., P.A.L., A.-R.B., C.D.B., A.M., and S.R.; supervision, N.K.B., C.D.B., A.M., N.R.D., and S.R.

DECLARATION OF INTERESTS

The authors declare no competing interests.

STAR★METHODS

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

SUPPLEMENTAL INFORMATION

Supplemental information can be found online at https://doi.org/10.1016/j.celrep.2025.116321.

Received: August 22, 2024
Revised: May 29, 2025
Accepted: August 29, 2025

REFERENCES

1. Lyu, Q., Li, Q., Zhou, J., and Zhao, H. (2024). Formation and function of multiciliated cells. J. Cell Biol. *223*, e202307150. https://doi.org/10.1083/jcb.202307150.
2. Lucas, J.S., Davis, S.D., Omran, H., and Shoemark, A. (2020). Primary ciliary dyskinesia in the genomics age. Lancet Respir. Med. *8*, 202–216. https://doi.org/10.1016/S2213-2600(19)30374-1.
3. Anderson, R.G., and Brenner, R.M. (1971). The formation of basal bodies (centrioles) in the Rhesus monkey oviduct. J. Cell Biol. *50*, 10–34. https://doi.org/10.1083/jcb.50.1.10.
4. Sorokin, S.P. (1968). Reconstructions of centriole formation and ciliogenesis in mammalian lungs. J. Cell Sci. *3*, 207–230. https://doi.org/10.1242/jcs.3.2.207.
5. Zhao, H., Zhu, L., Zhu, Y., Cao, J., Li, S., Huang, Q., Xu, T., Huang, X., Yan, X., and Zhu, X. (2013). The Cep63 paralogue Deup1 enables massive de novo centriole biogenesis for vertebrate multiciliogenesis. Nat. Cell Biol. *15*, 1434–1444. https://doi.org/10.1038/ncb2880.
6. Al Jord, A., Lemaître, A.I., Delgehyr, N., Faucourt, M., Spassky, N., and Meunier, A. (2014). Centriole amplification by mother and daughter centrioles differs in multiciliated cells. Nature *516*, 104–107. https://doi.org/10.1038/nature13770.
7. Mercey, O., Levine, M.S., LoMastro, G.M., Rostaing, P., Brotslaw, E., Gomez, V., Kumar, A., Spassky, N., Mitchell, B.J., Meunier, A., and Holland, A.J. (2019). Massive centriole production can occur in the absence of deuterosomes in multiciliated cells. Nat. Cell Biol. *21*, 1544–1552. https://doi.org/10.1038/s41556-019-0427-x.
8. Rayamajhi, D., and Roy, S. (2020). Multiciliated Cells: Rise and Fall of the Deuterosomes. Trends Cell Biol. *30*, 259–262. https://doi.org/10.1016/j.tcb.2020.02.003.
9. Zhao, H., Chen, Q., Fang, C., Huang, Q., Zhou, J., Yan, X., and Zhu, X. (2019). Parental centrioles are dispensable for deuterosome formation and function during basal body amplification. EMBO Rep. *20*, e46735. https://doi.org/10.15252/embr.201846735.

10. Mercey, O., Al Jord, A., Rostaing, P., Mahuzier, A., Fortoul, A., Boudjema, A.R., Faucourt, M., Spassky, N., and Meunier, A. (2019). Dynamics of centriole amplification in centrosome-depleted brain multiciliated progenitors. Sci. Rep. *9*, 13060. https://doi.org/10.1038/s41598-019-49416-2.

11. Lewis, M., and Stracker, T.H. (2021). Transcriptional regulation of multiciliated cell differentiation. Semin. Cell Dev. Biol. *110*, 51–60. https://doi.org/10.1016/j.semcdb.2020.04.007.

12. Ma, L., Quigley, I., Omran, H., and Kintner, C. (2014). Multicilin drives centriole biogenesis via E2f proteins. Genes Dev. *28*, 1461–1471. https://doi.org/10.1101/gad.243832.114.

13. Terre, B., Piergiovanni, G., Segura-Bayona, S., Gil-Gomez, G., Youssef, S. A., Attolini, C.S., Wilsch-Brauninger, M., Jung, C., Rojas, A.M., Marjanovic, M., et al. (2016). GEMC1 is a critical regulator of multiciliated cell differentiation. EMBO J. *35*, 942–960. https://doi.org/10.15252/embj.201592821.

14. Lewis, M., Terré, B., Knobel, P.A., Cheng, T., Lu, H., Attolini, C.S.O., Smak, J., Coyaud, E., Garcia-Cao, I., Sharma, S., et al. (2023). GEMC1 and MCIDAS interactions with SWI/SNF complexes regulate the multiciliated cell-specific transcriptional program. Cell Death Dis. *14*, 201. https://doi.org/10.1038/s41419-023-05720-4.

15. Lu, H., Anujan, P., Zhou, F., Zhang, Y., Chong, Y.L., Bingle, C.D., and Roy, S. (2019). Mcidas mutant mice reveal a two-step process for the specification and differentiation of multiciliated cells in mammals. Development *146*, dev172643. https://doi.org/10.1242/dev.172643.

16. Mori, M., Hazan, R., Danielian, P.S., Mahoney, J.E., Li, H., Lu, J., Miller, E. S., Zhu, X., Lees, J.A., and Cardoso, W.V. (2017). Cytoplasmic E2f4 forms organizing centres for initiation of centriole amplification during multiciliogenesis. Nat. Commun. *8*, 15857. https://doi.org/10.1038/ncomms15857.

17. Nakazawa, Y., Hiraki, M., Kamiya, R., and Hirono, M. (2007). SAS-6 is a cartwheel protein that establishes the 9-fold symmetry of the centriole. Curr. Biol. *17*, 2169–2174. https://doi.org/10.1016/j.cub.2007.11.046.

18. Hazan, R., Mori, M., Danielian, P.S., Guen, V.J., Rubin, S.M., Cardoso, W. V., and Lees, J.A. (2021). E2F4's cytoplasmic role in multiciliogenesis is mediated via an N-terminal domain that binds two components of the centriole replication machinery, Deup1 and SAS6. Mol. Biol. Cell *32*, ar1. https://doi.org/10.1091/mbc.E21-01-0039.

19. Hawkins, F.J., Suzuki, S., Beermann, M.L., Barillà, C., Wang, R., Villacorta-Martin, C., Berical, A., Jean, J.C., Le Suer, J., Matte, T., et al. (2021). Derivation of Airway Basal Stem Cells from Human Pluripotent Stem Cells. Cell Stem Cell *28*, 79–95.e8. https://doi.org/10.1016/j.stem.2020.09.017.

20. Soh, B.S., Zheng, D., Li Yeo, J.S., Yang, H.H., Ng, S.Y., Wong, L.H., Zhang, W., Li, P., Nichane, M., Asmat, A., et al. (2012). CD166(pos) subpopulation from differentiated human ES and iPS cells support repair of acute lung injury. Mol. Ther. *20*, 2335–2346. https://doi.org/10.1038/mt.2012.182.

21. Boon, M., Wallmeier, J., Ma, L., Loges, N.T., Jaspers, M., Olbrich, H., Dougherty, G.W., Raidt, J., Werner, C., Amirav, I., et al. (2014). MCIDAS mutations result in a mucociliary clearance disorder with reduced generation of multiple motile cilia. Nat. Commun. *5*, 4418. https://doi.org/10.1038/ncomms5418.

22. Pefani, D.E., Dimaki, M., Spella, M., Karantzelis, N., Mitsiki, E., Kyrousi, C., Symeonidou, I.E., Perrakis, A., Taraviras, S., and Lygerou, Z. (2011). Idas, a novel phylogenetically conserved geminin-related protein, binds to geminin and is required for cell cycle progression. J. Biol. Chem. *286*, 23234–23246. https://doi.org/10.1074/jbc.M110.207688.

23. Arbi, M., Skamnelou, M., Bournaka, S., Zitouni, S., Tsaridou, S., Karayel, O., Vasilopoulou, C.G., Tsika, A.C., Giakoumakis, N.N., Spyroulias, G.A., et al. (2022). McIdas localizes at centrioles and controls centriole numbers through PLK4-dependent phosphorylation. Preprint at bioRxiv. https://doi.org/10.1101/2022.09.30.510086.

24. Arbi, M., Pefani, D.E., Kyrousi, C., Lalioti, M.E., Kalogeropoulou, A., Papanastasiou, A.D., Taraviras, S., and Lygerou, Z. (2016). GemC1 controls multiciliogenesis in the airway epithelium. EMBO Rep. *17*, 400–413. https://doi.org/10.15252/embr.201540882.

25. Zhou, F., Narasimhan, V., Shboul, M., Chong, Y.L., Reversade, B., and Roy, S. (2015). Gmnc Is a Master Regulator of the Multiciliated Cell Differentiation Program. Curr. Biol. *25*, 3267–3273. https://doi.org/10.1016/j.cub.2015.10.062.

26. Li, M., Wang, S., Cai, M., and Zheng, C. (2011). Identification of nuclear and nucleolar localization signals of pseudorabies virus (PRV) early protein UL54 reveals that its nuclear targeting is required for efficient production of PRV. J. Virol. *85*, 10239–10251. https://doi.org/10.1128/JVI.05223-11.

27. Kosugi, S., Yanagawa, H., Terauchi, R., and Tabata, S. (2014). NESmapper: accurate prediction of leucine-rich nuclear export signals using activity-based profiles. PLoS Comput. Biol. *10*, e1003841. https://doi.org/10.1371/journal.pcbi.1003841.

28. Vladar, E.K., and Stearns, T. (2007). Molecular characterization of centriole assembly in ciliated epithelial cells. J. Cell Biol. *178*, 31–42. https://doi.org/10.1083/jcb.200703064.

29. Zhao, H., Chen, Q., Li, F., Cui, L., Xie, L., Huang, Q., Liang, X., Zhou, J., Yan, X., and Zhu, X. (2021). Fibrogranular materials function as organizers to ensure the fidelity of multiciliary assembly. Nat. Commun. *12*, 1273. https://doi.org/10.1038/s41467-021-21506-8.

30. Hall, E.A., Kumar, D., Prosser, S.L., Yeyati, P.L., Herranz-Pérez, V., García-Verdugo, J.M., Rose, L., McKie, L., Dodd, D.O., Tennant, P.A., et al. (2023). Centriolar satellites expedite mother centriole remodeling to promote ciliogenesis. eLife *12*, e79299. https://doi.org/10.7554/eLife.79299.

31. Gopalakrishnan, J., Guichard, P., Smith, A.H., Schwarz, H., Agard, D.A., Marco, S., and Avidor-Reiss, T. (2010). Self-assembling SAS-6 multimer is a core centriole building block. J. Biol. Chem. *285*, 8759–8770. https://doi.org/10.1074/jbc.M109.092627.

32. Ossareh-Nazari, B., Bachelerie, F., and Dargemont, C. (1997). Evidence for a role of CRM1 in signal-mediated nuclear protein export. Science *278*, 141–144. https://doi.org/10.1126/science.278.5335.141.

33. LoMastro, G.M., Drown, C.G., Maryniak, A.L., Jewett, C.E., Strong, M.A., and Holland, A.J. (2022). PLK4 drives centriole amplification and apical surface area expansion in multiciliated cells. eLife *11*, e80643. https://doi.org/10.7554/eLife.80643.

34. Habedanck, R., Stierhof, Y.D., Wilkinson, C.J., and Nigg, E.A. (2005). The Polo kinase Plk4 functions in centriole duplication. Nat. Cell Biol. *7*, 1140–1146. https://doi.org/10.1038/ncb1320.

35. Graser, S., Stierhof, Y.D., Lavoie, S.B., Gassner, O.S., Lamla, S., Le Clech, M., and Nigg, E.A. (2007). Cep164, a novel centriole appendage protein required for primary cilium formation. J. Cell Biol. *179*, 321–330. https://doi.org/10.1083/jcb.200707181.

36. Stubbs, J.L., Vladar, E.K., Axelrod, J.D., and Kintner, C. (2012). Multicilin promotes centriole assembly and ciliogenesis during multiciliate cell differentiation. Nat. Cell Biol. *14*, 140–147. https://doi.org/10.1038/ncb2406.

37. Kim, S., Ma, L., Shokhirev, M.N., Quigley, I., and Kintner, C. (2018). Multicilin and activated E2f4 induce multiciliated cell differentiation in primary fibroblasts. Sci. Rep. *8*, 12369. https://doi.org/10.1038/s41598-018-30791-1.

38. Liu, Q., Zhou, J., Daiger, S.P., Farber, D.B., Heckenlively, J.R., Smith, J.E., Sullivan, L.S., Zuo, J., Milam, A.H., and Pierce, E.A. (2002). Identification and subcellular localization of the RP1 protein in human and mouse photoreceptors. Investig. Ophthalmol. Vis. Sci. *43*, 22–32.

39. Sullivan, L.S., Heckenlively, J.R., Bowne, S.J., Zuo, J., Hide, W.A., Gal, A., Denton, M., Inglehearn, C.F., Blanton, S.H., and Daiger, S.P. (1999). Mutations in a novel retina-specific gene cause autosomal dominant retinitis pigmentosa. Nat. Genet. *22*, 255–259. https://doi.org/10.1038/10314.

40. Ostrowski, L.E., Blackburn, K., Radde, K.M., Moyer, M.B., Schlatzer, D. M., Moseley, A., and Boucher, R.C. (2002). A proteomic analysis of human cilia: identification of novel components. Mol. Cell. Proteomics *1*, 451–465. https://doi.org/10.1074/mcp.m200037-mcp200.

41. Damaa, M.K., Serizay, J., Balagué, R., Boudjema, A.-R., Faucourt, M., Delgehyr, N., Goh, K.J., Lu, H., Tan, E.K., James, C.T., et al. (2024). Cyclin O controls entry into the cell-cycle variant required for multiciliated cell

differentiation. Preprint at bioRxiv. https://doi.org/10.1101/2024.05.22.595363.

42. Serizay, J., Damaa, M.K., Boudjema, A.-R., Balagué, R., Faucourt, M., Delgehyr, N., Noûs, C., Zaragosi, L.-E., Barbry, P., Spassky, N., et al. (2024). Identification of a new cell cycle variant during multiciliated cell differentiation. Preprint at bioRxiv. https://doi.org/10.1101/2024.05.22.595357.
43. Choksi, S.P., Byrnes, L.E., Konjikusic, M.J., Tsai, B.W.H., Deleon, R., Lu, Q., Westlake, C.J., and Reiter, J.F. (2024). An alternative cell cycle coordinates multiciliated cell differentiation. Nature *630*, 214–221. https://doi.org/10.1038/s41586-024-07476-z.
44. Zhou, F., and Roy, S. (2015). SnapShot: Motile Cilia. Cell *162*, 224–224.e1. https://doi.org/10.1016/j.cell.2015.06.048.
45. Brooks, E.R., and Wallingford, J.B. (2014). Multiciliated cells. Curr. Biol. *24*, R973–R982. https://doi.org/10.1016/j.cub.2014.08.047.
46. Mizukami, I., and Gall, J. (1966). Centriole replication. II. Sperm formation in the fern, Marsilea, and the cycad, Zamia. J. Cell Biol. *29*, 97–111. https://doi.org/10.1083/jcb.29.1.97.
47. Gaubatz, S., Lees, J.A., Lindeman, G.J., and Livingston, D.M. (2001). E2F4 is exported from the nucleus in a CRM1-dependent manner. Mol. Cell Biol. *21*, 1384–1392. https://doi.org/10.1128/MCB.21.4.1384-1392.2001.
48. Apostolova, M.D., Ivanova, I.A., Dagnino, C., D'Souza, S.J.A., and Dagnino, L. (2002). Active nuclear import and export pathways regulate E2F-5 subcellular localization. J. Biol. Chem. *277*, 34471–34479. https://doi.org/10.1074/jbc.M205827200.
49. Arquint, C., and Nigg, E.A. (2016). The PLK4-STIL-SAS-6 module at the core of centriole duplication. Biochem. Soc. Trans. *44*, 1253–1263. https://doi.org/10.1042/BST20160116.
50. Everman, J.L., Sajuthi, S., Saef, B., Rios, C., Stoner, A.M., Numata, M., Hu, D., Eng, C., Oh, S., Rodriguez-Santana, J., et al. (2019). Functional genomics of CDHR3 confirms its role in HRV-C infection and childhood asthma exacerbations. J. Allergy Clin. Immunol. *144*, 962–971. https://doi.org/10.1016/j.jaci.2019.01.052.
51. Wolters, A.A.B., Kersten, E.T.G., and Koppelman, G.H. (2024). Genetics of preschool wheeze and its progression to childhood asthma. Pediatr. Allergy Immunol. *35*, e14067. https://doi.org/10.1111/pai.14067.
52. Dobin, A., Davis, C.A., Schlesinger, F., Drenkow, J., Zaleski, C., Jha, S., Batut, P., Chaisson, M., and Gingeras, T.R. (2013). STAR: ultrafast universal RNA-seq aligner. Bioinformatics *29*, 15–21. https://doi.org/10.1093/bioinformatics/bts635.
53. Anders, S., Pyl, P.T., and Huber, W. (2015). HTSeq–a Python framework to work with high-throughput sequencing data. Bioinformatics *31*, 166–169. https://doi.org/10.1093/bioinformatics/btu638.
54. Love, M.I., Huber, W., and Anders, S. (2014). Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biol. *15*, 550. https://doi.org/10.1186/s13059-014-0550-8.
55. Shannon, P., Markiel, A., Ozier, O., Baliga, N.S., Wang, J.T., Ramage, D., Amin, N., Schwikowski, B., and Ideker, T. (2003). Cytoscape: a software environment for integrated models of biomolecular interaction networks. Genome Res. *13*, 2498–2504. https://doi.org/10.1101/gr.1239303.
56. Bindea, G., Mlecnik, B., Hackl, H., Charoentong, P., Tosolini, M., Kirilovsky, A., Fridman, W.H., Pagès, F., Trajanoski, Z., and Galon, J. (2009). ClueGO: a Cytoscape plug-in to decipher functionally grouped gene ontology and pathway annotation networks. Bioinformatics *25*, 1091–1093. https://doi.org/10.1093/bioinformatics/btp101.
57. Bindea, G., Galon, J., and Mlecnik, B. (2013). CluePedia Cytoscape plugin: pathway insights using integrated experimental and in silico data. Bioinformatics *29*, 661–663. https://doi.org/10.1093/bioinformatics/btt019.
58. Hu, C., Li, T., Xu, Y., Zhang, X., Li, F., Bai, J., Chen, J., Jiang, W., Yang, K., Ou, Q., et al. (2023). CellMarker 2.0: an updated database of manually curated cell markers in human/mouse and web tools based on scRNA-seq data. Nucleic Acids Res. *51*, D870–D876. https://doi.org/10.1093/nar/gkac947.
59. Thul, P.J., and Lindskog, C. (2018). The human protein atlas: A spatial map of the human proteome. Protein Sci. *27*, 233–244. https://doi.org/10.1002/pro.3307.
60. Hamosh, A., Scott, A.F., Amberger, J.S., Bocchini, C.A., and McKusick, V.A. (2005). Online Mendelian Inheritance in Man (OMIM), a knowledgebase of human genes and genetic disorders. Nucleic Acids Res. *33*, D514–D517. https://doi.org/10.1093/nar/gki033.
61. Thumuluri, V., Almagro Armenteros, J.J., Johansen, A.R., Nielsen, H., and Winther, O. (2022). DeepLoc 2.0: multi-label subcellular localization prediction using protein language models. Nucleic Acids Res. *50*, W228–W234. https://doi.org/10.1093/nar/gkac278.
62. Nguyen Ba, A.N., Pogoutse, A., Provart, N., and Moses, A.M. (2009). NLStradamus: a simple Hidden Markov Model for nuclear localization signal prediction. BMC Bioinf. *10*, 202. https://doi.org/10.1186/1471-2105-10-202.
63. Scott, M.S., Troshin, P.V., and Barton, G.J. (2011). NoD: a Nucleolar localization sequence detector for eukaryotic and viral proteins. BMC Bioinf. *12*, 317. https://doi.org/10.1186/1471-2105-12-317.
64. Du, Z., Su, H., Wang, W., Ye, L., Wei, H., Peng, Z., Anishchenko, I., Baker, D., and Yang, J. (2021). The trRosetta server for fast and accurate protein structure prediction. Nat. Protoc. *16*, 5634–5651. https://doi.org/10.1038/s41596-021-00628-9.
65. Crooks, G.E., Hon, G., Chandonia, J.M., and Brenner, S.E. (2004). WebLogo: a sequence logo generator. Genome Res. *14*, 1188–1190. https://doi.org/10.1101/gr.849004.
66. Christoffer, C., Chen, S., Bharadwaj, V., Aderinwale, T., Kumar, V., Hormati, M., and Kihara, D. (2021). LZerD webserver for pairwise and multiple protein-protein docking. Nucleic Acids Res. *49*, W359–W365. https://doi.org/10.1093/nar/gkab336.
67. Weng, G., Wang, E., Wang, Z., Liu, H., Zhu, F., Li, D., and Hou, T. (2019). HawkDock: a web server to predict and analyze the protein-protein complex based on computational docking and MM/GBSA. Nucleic Acids Res. *47*, W322–W330. https://doi.org/10.1093/nar/gkz397.
68. Berman, H.M., Westbrook, J., Feng, Z., Gilliland, G., Bhat, T.N., Weissig, H., Shindyalov, I.N., and Bourne, P.E. (2000). The Protein Data Bank. Nucleic Acids Res. *28*, 235–242. https://doi.org/10.1093/nar/28.1.235.
69. UniProt Consortium, T. (2018). UniProt: the universal protein knowledgebase. Nucleic Acids Res. *46*, 2699. https://doi.org/10.1093/nar/gky092.
70. Montague, T.G., Cruz, J.M., Gagnon, J.A., Church, G.M., and Valen, E. (2014). CHOPCHOP: a CRISPR/Cas9 and TALEN web tool for genome editing. Nucleic Acids Res. *42*, W401–W407. https://doi.org/10.1093/nar/gku410.
71. Untergasser, A., Cutcutache, I., Koressaar, T., Ye, J., Faircloth, B.C., Remm, M., and Rozen, S.G. (2012). Primer3–new capabilities and interfaces. Nucleic Acids Res. *40*, e115. https://doi.org/10.1093/nar/gks596.
72. Vallier, L., Mendjan, S., Brown, S., Chng, Z., Teo, A., Smithers, L.E., Trotter, M.W.B., Cho, C.H.H., Martinez, A., Rugg-Gunn, P., et al. (2009). Activin/Nodal signalling maintains pluripotency by controlling Nanog expression. Development *136*, 1339–1349. https://doi.org/10.1242/dev.033951.
73. McCauley, K.B., Hawkins, F., Serra, M., Thomas, D.C., Jacob, A., and Kotton, D.N. (2017). Efficient Derivation of Functional Human Airway Epithelium from Pluripotent Stem Cells via Temporal Regulation of Wnt Signaling. Cell Stem Cell *20*, 844–857.e6. https://doi.org/10.1016/j.stem.2017.03.001.
74. UniProt Consortium (2023). UniProt: the Universal Protein Knowledgebase in 2023. Nucleic Acids Res. *51*, D523–D531. https://doi.org/10.1093/nar/gkac1052.
75. Yamagishi, R., and Kaneko, H. (2016). Data from comprehensive analysis of nuclear localization signals. Data Brief *6*, 200–203. https://doi.org/10.1016/j.dib.2015.11.064.

STAR★METHODS

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Antibodies		
ALCAM Phycoerythrin monoclonal antibody (Clone 105902)	R&D Systems	Cat# FAB6561P; RRID: AB_2223887
APC anti-human CD271 (NGFR) antibody	Biolegend	Cat# 345108; RRID: AB_10645515
PE anti-human CD326 (EpCAM) antibody	Biolegend	Cat# 324206; RRID: AB_756080
Anti-NANOG antibody	R&D systems	Cat# AF1997; RRID: AB_355097
Anti-HA antibody	Cell Signaling	Cat# CST 3724; RRID: AB_1549585
Anti-α-tubulin antibody	Sigma-Aldrich	Cat# T9026; RRID: AB_477593
Anti-E2F5 antibody	Santa Cruz	Cat# sc-374268; RRID: AB_10988935
Anti-E2F5 antibody	Atlas	Cat# HPA065441
Anti-E2F4 antibody	EMD millipore	Cat# MABE160; RRID: AB_10845939
Anti-SAS6 antibody	Santa Cruz	Cat# SC-81431; RRID: AB_1128357
Anti-CENTRIN antibody	EMD millipore	Cat# 04-1624; RRID: AB_11211820
Anti-PCM1 antibody	Cell Signaling	Cat# 5213; RRID: AB_10069824
Anti-PCNT antibody	Ab-Cam	Cat# Ab4448; RRID: AB_304461
Anti-FOXJ1 antibody	eBiosecience	Cat# 14-9965-80; RRID: AB_1548835
Anti-RFX2 antibody	Sigma-Aldrich	Cat# HPA048969; RRID: AB_2756226
Anti-RFX3 antibody	Sigma-Aldrich	Cat# HPA035689; RRID: AB_10671224
Anti-Acetylated tubulin antibody	Sigma-Aldrich	Cat# T6793; RRID: AB_10061791
Anti-Acetylated tubulin antibody	Cell signaling	Cat# 5335; RRID: AB_
Anti-γ-tubulin antibody	Sigma-Aldrich	Cat# T6557; RRID: AB_477584
Anti-γ-tubulin antibody	Sigma-Aldrich	Cat# T5192; RRID: AB_261690
Anti-CEP164 antibody	Sigma-Aldrich	Cat# SAB3500022; RRID: AB_10603614
Anti-Rabbit IgG (H + L), Alexa Fluor 546	Invitrogen	Cat# A10040; RRID: AB_11181145
Anti-Rabbit IgG (H + L), Alexa Fluor 555	Invitrogen	Cat# A-21428; RRID: AB_2535849
Anti-mouse IgG (H + L), Alexa Fluor 555	Invitrogen	Cat# A-28180; RRID: AB_2536164
Anti-Rabbit IgG (H + L), Alexa Fluor 488	Invitrogen	Cat# A-11034; RRID: AB_2576217
Anti-Mouse IgG (H + L), Alexa Fluor Plus 488	Invitrogen	Cat# A32766; RRID: AB_2762823
Anti-Goat IgG (H + L), Alexa Fluor 488	Invitrogen	Cat# A11001; RRID: AB_2534069
Anti-Mouse HRP	Dako	Cat# P044701; RRID: AB_2617137
Anti-Rabbit HRP	Dako	Cat# P044801; RRID: AB_2617138
Hoechst 33342	Sigma-Aldrich	Cat# B2261; CAS: 875756-97-1
Bacterial and virus strains		
XL-10 gold ultracompetent *E. coli* cells	Agilent Technologies	Cat# 200315
Chemicals, peptides, and recombinant proteins		
Penicillin/Streptomycyin	Life Technologies	Cat# 15140122
Leptomycin B	Sigma-Aldrich	Cat# L2913
Lipofectamine 2000	Invitrogen	Cat# 11668019
EGTA	Sigma Aldrich	Cat# E3889
HEPES Buffer	Sigma Aldrich	Cat# H3375-25G
Polybrene	Sigma-Aldrich	Cat# H9268
Hanks Balanced Salt Solution	HyClone	Cat# SH30588.01
Activin A	Peprotech	Cat# 120-14P-50
FGF2	Peprotech	Cat# 100-18B-100
LY294002	Promega	Cat# V1201

(*Continued on next page*)

Continued

REAGENT or RESOURCE	SOURCE	IDENTIFIER
CHIR99021	Tocris	Cat# 4423/10
BMP4	Peprotech	Cat# 120-05-50
SB431542	Tocris	Cat# 1614
Dorsomorphin dihydrochloride	Tocris	Cat# 3093
GlutaMAX	Life Technologies	Cat# 35050-061
N-2 Supplement	Life Technologies	Cat# 17502-048
B27 Supplement	Life Technologies	Cat# 17504-044
Non essential amino acids	Gibco	Cat# 11140-035
FGF7	Peprotech	Cat# 100-19-100
FGF10	Peprotech	Cat# 100-26-100
Retinoic Acid	Sigma	Cat# R2625-50MG
Dexamethasone	Sigma	Cat# D4902-100MG
8-bromoadenosine 30,50-cyclic monophosphate sodium salt	Sigma	Cat# B5386-100MG
3-isobutyl-1-methylxanthine	Sigma	Cat# 17018-100MG
Y-27632	Selleck Chem	Cat# S1049-SEL-10mg
hESC-qualified growth factor reduced Matrigel	Corning	Cat# 354277
Biocoat Growth Factor Reduced Basement Membrane Matrigel (Corning)	Corning	Cat# 354230
Vitronectin XF	Stem Cell Technologies	Cat# 07180
Leibovitz L-15 Medium	Thermo Fisher	Cat# 11415064
Ovomucoid	Sigma-Aldrich	Cat# O2769
Poly-L-Lysine	Sigma-Aldrich	Cat# P8290
Centrinone	Tocris	Cat# 5065
Fluoromount	Southern Biotech	Cat# 0100-01
PET Transwell membrane	Life Science	Cat# 353095
6.5 mm Transwell, 0.4 μm Pore Polyester Inserts	Stem Cell Technologies	Cat# 38024
mTEC plus Medium	Merck Millipore	Cat# MTEC/Plus
Fetal Bovine Serum	Hyclone	Cat# SH30071.03HI
Opti-Mem	Life Technologies	Cat# 31985070
Advanced DMEM F12 Medium	Life Technologies	Cat# 12634028
mTeSR1 Medium	Stem Cell Technologies	Cat# 85850
Essential 8 Medium	Life Technologies	Cat# A1517001
Gentle Cell Dissociation Buffer	Stem Cell Technologies	Cat# 07174
Blasticidin	Life Technologies	Cat# R210-01
IGEPAL CA630	Sigma Aldrich	Cat# I8896
Proteinase K	Qiagen	Cat# 19133
1 x TE Buffer	Thermo Fisher	Cat# 12090015
Cryostor CS10	Stem Cell Technologies	Cat# 07930
Gelatin	Sigma	Cat# G1890-100G
Accutase	Stem Cell Technologies	Cat# 07920
F-12 Medium	Life Technologies	Cat# 31765068
IMDM	Life Technologies	Cat# 21980065
Polyvinyl Alcohol	Sigma	Cat# P8136-250G
Concentrated Lipids	Life Technologies	Cat# 11905031
Monothioglycerol	Sigma	Cat# M6145-100ML
Transferrin	Sigma	Cat# T1147
Insulin	Roche	Cat# 1376497

(*Continued on next page*)

Continued

REAGENT or RESOURCE	SOURCE	IDENTIFIER
RPMI Medium	Life Technologies	Cat# 61870-010
PneumaCULT Ex Plus Basal Medium	Stem Cell Technologies	Cat# 05040
A 83-01	Stem Cell Technologies	Cat# 100-0245
PneumaCULT-ALI Basal Medium	Stem Cell Technologies	Cat# 05022
Tryple Express	Thermo Fisher	Cat# 12604021
S.O.C Medium	Thermo Fisher	Cat# 15544034
Ampicillin Sodium Salt	Sigma-Aldrich	Cat# A19518
Kannamycin Sulfate	Millipore	Cat# 420311
CloneR	Stem Cell Technologies	Cat# 05888
L-Glutamine	Life Technologies	Cat# 25030024
Sodium Bicarbonate	Sigma-Aldrich	Cat# S6014
DNase I	Invitrogen	Cat# AM2222
D-Glucose	Sigma-Aldrich	Cat# G7258
cOmplete protease inhibitor cocktail	Calbiochem	Cat# 539131
Pierce RIPA BUffer	Thermo Fisher	Cat# 89900
Tween 20	Sigma-Aldrich	Cat# P9416
Critical commercial assays		
SuperScript III First-Strand Synthesis System	Invitrogen	Cat# 18080051
EXPRESS SYBR GreenER Super Mix	Invitrogen	Cat# A10315
NEBNext Ultra II directional library preparation kit	New England Biolabs	Cat# E7760S
NEBNext Poly(A) mRNA isolation kit	New England Biolabs	Cat# E7490S
SuperSignal West Femto Maximum Sensitivity Substrate	Thermo Fisher	Cat# 34095
QuickChange II XL Site Directed Mutagenesis Kit	Agilent Technologies	Cat# 200521
Nucleofection Amaxa P3 primary Cell 4D Nucleofector Kit	Lonza	Cat# V4XP-3024
PrimeStar Max Mastermix	Takara Bio Inc	Cat# R045A,
ViraPower Lentiviral Expression Systems Version C	Invitrogen	Cat# 25-0501
Lentiviral Packaging Mix	Invitrogen	Cat# K4975-00
QIAprep spin miniprep	Qiagen	Cat# 27106
Zymopure II plasmid maxiprep	Zymopure	Cat# D4203-B
Deposited data		
Raw RNA-seq data of *Mcidas* mutant and wild-type mTECs have been deposited at the European Nucleotide Archive (ENA) with	European Nucleotide Archive (ENA)	https://www.ebi.ac.uk/ena/browser/home (accession number PRJEB77353)
Experimental models: Cell lines		
H9 human embryonic stem cells	WiCell	RRID:CVCL_9773, Cat# WA09
HEK293T cells	ATCC	Cat# CRL-3216
Experimental models: Organisms/strains		
C57BL/6 mice	The Jackson Laboratory	Cat# 000664
Oligonucleotides		
hE2F5 sgRNA: GCCTGCTGGCCC GAGCTCGC	This Paper	N/A
hMCI sgRNA: GTCGGTGTACGAG GATCCCC	This Paper	N/A
HA-GMNC sgRNA: GCAGGTATTC CACTACAACGTCAG	This Paper	N/A

(*Continued on next page*)

Continued

REAGENT or RESOURCE	SOURCE	IDENTIFIER
HA-MCIDAS agRNA: GCAGTCGG GCCTTCGACAGCATCT	This Paper	N/A
hMCI F: CCATCTCTCAGCACCTCCTC	This Paper	N/A
hMCI R: CAGGTCTATGGTGGTGAGGG	This Paper	N/A
hE2F5 F: AGCCCCTTCTCCTCGTTTC	This Paper	N/A
hE2F5 R: CAGCGACACGAACTTGGTAG	This Paper	N/A
MCI-HA-G-F1: CCTTGGAGAGGTGTTGGCTT	This Paper	N/A
MCI-HA-G-R1: CCAGATGAGCCCGACTGTATC	This Paper	N/A
GMNC-HA-G-F1: CCAAAGCTGGGCTCGCTAA	This Paper	N/A
GMNC-HA-G-R1: TTTGGGGTTCCTTTGTTTGTGTCT	This Paper	N/A
MCI-HA-XhoI-pLvx: GATCGATCC TCGAGGCCACCATGT ACCCATA CGACGTGCCAGACTACG CAATG CAGGCGTGCGGGGG	This Paper	N/A
MCI-C-XbaI-pLvx: GATCGATCTC TAGATCAACTGGG GACCCAGCG	This Paper	N/A
MCI-NES-N: ACCCGGCACGACG CCTCGGTGGACGATAAGGACAT GATCAC	This Paper	N/A
MCI-NES-C: GTGATCATGTCCTT ATCGTCCACCGAGGCGTCGTGCCGGGT	This Paper	N/A
mRP1-F: TCAACCAAGTAGTAAGAG	This Paper	N/A
mRP1-R: ATCTTCTTGTTCTCTACTGCT	This Paper	N/A
mCDHR3-F: GAGGCCATGGTGGAATC	This Paper	N/A
mCDHR3-R: TTATCTCCCTGTGTGTTTAG	This Paper	N/A
HA-MCI-5′-HomoArm-F1: AGTGGAGGAGGGTTCGTAGTGT	This Paper	N/A
HA-MCI-5′-HomoArm-R1: GCGCC GCCCCCGCACGCCTGAGCGTAA TCTGGAACATCGTATGGGTAAGC GTAATCTGGAACATCGTATGGGT AAGCGTAATCTGGAACATCGTAT GGGTACATTGTGCCTCCTGCCTC	This Paper	N/A
HA-MCI-3′HomoArm-F1: CAGGCGTGCGGGGGCGGCGC	This Paper	N/A
HA-MCI-3′-HomoArm-R1: TTGGTCAAATCAGCTCCAGTGGA	This Paper	N/A
HA-GMNC-5′-HomoArm-F1: GGGCACTGAGCCCTAGTGAT	This Paper	N/A
HA-GMNC-5′-HomoArm-R1: TCTTG GCAAGGCAGAATAGCGTAATCTGGA ACATCGTATGGGTAAGCGTAATCTGG AACATCGTATGGGTAAGCGTAATCTG GAACATCGTATGGGTAGGTGTTCTG TGAAATTCA	This Paper	N/A
HA-GMNC-3′HomoArm-F1: ATTCTGCCTTGCCAAGACCAG	This Paper	N/A
HA-GMNC-3′HomoArm-R1: ATACTGCTGGTATGCCACAGATC	This Paper	N/A

(*Continued on next page*)

Continued

REAGENT or RESOURCE	SOURCE	IDENTIFIER
HA-E2F5-5′-HomoArm-F1: TGGCTCCACATTCAAAATAGGC	This Paper	N/A
HA-E2F5-5′-HomoArm-R1: CCTC ACAGCCAAAGTATCAGCAGCGT AATCTGGAACATCGTATGGGTA AGCGTAATCTGGAACATCGTAT GGGTAAGCGTAATCTGGAACAT CGTATGGGTAAGCCTGTTCAGG AAAGACAAC	This Paper	N/A
HA-E2F5-3′HomoArm-F1: GCTGATACTTTGGCTGTGAGG	This Paper	N/A
HA-E2F5-3′-HomoArm-R1: CATCTGGAAGGCCAGACTGG	This Paper	N/A
pCAG-HF1-Seq-F1: CCTTACATGTTTTACTAGCCAGA	This Paper	N/A
Recombinant DNA		
PLVX	Addgene	Cat# 632159
pMD2.G	Addgene	Cat# 12259
pRSV-REV	Addgene	Cat# 12253
pMDLg	Addgene	Cat# 12251
pRRE	Addgene	Cat# 12251
pCAG-Csy4-T2A-Cas9-HF1-IRES-BSD-U6	Addgene	Cat# 79144
Zero Blunt Vector	Life Technologies	Cat# K2700-20
Plasmid: pCAG-HF1-Cas9-MCI-sgRNA	This Paper	N/A
Plasmid: pCAG-HF1-Cas9-GMNC-sgRNA	This Paper	N/A
Plasmid: pCAG-HF1-Cas9-E2F5-sgRNA	This Paper	N/A
Plasmid: pCR-Blunt-HA-MCI	This Paper	N/A
Plasmid: pCR-Blunt-HA-GMNC	This Paper	N/A
Plasmid: pCR-Blunt-HA-E2F5	This Paper	N/A
Software and algorithms		
Applied BioSystems 7900HT Fast Real-Time PCR SDS2.4 software	Thermo Fisher Scientific	Cat #7900 SDS v2.4.1
ImageJ	NIH/ImageJ Community	Version 1.53
FastQC (version 0.11.8)	Babraham Bioinformatics	https://www.bioinformatics.babraham.ac.uk/projects/fastqc/
Reference genome (*Mus musculus*) GENCODE Release M36 (GRCm39)	GENCODE	https://www.gencodegenes.org/mouse
STAR aligner (version 2.6.0c)	Dobin et al.[52]	https://github.com/alexdobin/STAR
HTSeq (version 0.11.0)	Anders et al.[53]	https://htseq.readthedocs.io/en/latest/index.html
DESeq2 package (version 1.22.2) With R version 3.5.1	Love et al.[54]	https://bioconductor.org/packages/devel/bioc/html/DESeq2.html
Cytoscape package (version 3.10.1)	Shannon et al.[55]	https://cytoscape.org
ClueGO (version 2.5.10)	Bindea et al.[56]	https://apps.cytoscape.org/apps/cluego
CluePedia (version 1.5.10)	Bindea et al.[57]	https://apps.cytoscape.org/apps/cluepedia
CellMarker 2.0 database	Hu et al.[58]	http://117.50.127.228/CellMarker/index.html
Human Protein Atlas	Thul et al.[59]	https://www.proteinatlas.org
OMIM database	Hamosh et al.[60]	https://www.omim.org/
DeepLoc 2.0	Thumuluri et al.[61]	https://services.healthtech.dtu.dk/services/DeepLoc-2.0
NLSeer	NLSeer Developers	https://github.com/auerbachsd/NLSeer

(*Continued on next page*)

Continued

REAGENT or RESOURCE	SOURCE	IDENTIFIER
NLStradamus	Nguyen Ba et al.[62]	http://www.moseslab.csb.utoronto.ca/NLStradamus/
NESmapper	Kosugi et al.[27]	https://mybiosoftware.com/nesmapper-prediction-of-crm1-dependent-nuclear-export-signals.html
NoD detector	Scott et al.[63]	https://www.compbio.dundee.ac.uk/www-nod/
trROSETTA	Du et al.[64]	https://yanglab.qd.sdu.edu.cn/trRosetta/
WebLogo3	Crooks et al.[65]	https://weblogo.threeplusone.com/https://github.com/gecrooks/weblogo
LZerD Protein Docking Web Server	Christoffer et al.[66]	https://lzerd.kiharalab.org/about/
HawkDock Server	Weng et al.[67]	https://cadd.zju.edu.cn/hawkdock/
RCSB database	Berman et al.[68]	https://www.rcsb.org/
UniProt Database	The UniProt Consortium[69]	https://www.uniprot.org/
CHOPCHOP	Montague et al.[70]	https://chopchop.cbu.uib.no/
FV31S-AW FLUOVIEW software	Olympus Global	Cat #FV31S-DT
GraphPad Prism software	GraphPad Software	Version 9.5.1
BD FACSDiva software	BD Biosciences	Version 8.0
Primer3	Untergasser et al.[71]	http://primer3.ut.ee
SnapGene	Domatics	Version 7.1.0
Other		
Scepter Cell Counter	Merck Millipore	Cat# PHCC20060
FACSAria Fusion Cell Sorter	BD Biosciences	Special Order Research Product (SORP)
ChemiDoc MP Imaging System	Bio-Rad	Cat# 12003154
Olympus FV3000 upright confocal microscope	Olympus	Cat# SKU: fv3000

EXPERIMENTAL MODELS AND STUDY PARTICIPANT DETAILS

Animals and ethics

All animal studies were performed in accordance with the guidelines of the Singapore National Advisory Committee on Laboratory Animal Research (protocol number: 211617) and the European Community and French Ministry of Agriculture and approved by the Direction départementale de la protection des populations de Paris (approval number Ce5/2012/107; APAFiS #9343). C57BL/6 mice were used as wild-type controls. *Mcidas* and *Deup1* mutant mouse strains, generated in the C57BL/6 background, have been described previously.[7,15]

In vitro culture of human embryonic stem cells

The parent cell line used for all experiments was the female human embryonic stem cell line (hESC) H9 (WiCell, Madison, WI, https://www.wicell.org; WA09, RRID: CVCL_9773). H9 is a well-characterized female hESC line originally derived from a blastocyst-stage embryo and all experiments were approved by the institutional biosafety committee (approval number: HSE-IBC +GMAC-SR-10). Cell lines were routinely tested and confirmed to be mycoplasma-free. Cells used in experiments were between passages 30–50. For experimental studies, H9 hESCs were differentiated into multiciliated epithelial cells using a directed differentiation protocol (described in method details). For specific experiments, gene knockouts (KOs) and knock-ins (KIs) were introduced using CRISPR-Cas9 genome editing, and edited clones were validated by Sanger sequencing and/or qPCR as described. Modified lines were expanded from single-cell clones and used between passages 30–50 for downstream analyses. Detailed genotyping, editing strategy, and sequence information are available in method details and key resources table.

METHOD DETAILS

Mouse primary ependymal cell cultures and centrinone treatment

Cell culture and centrinone treatment, described previously,[7] was performed as follows: Newborn mice (P0–P2) were sacrificed by decapitation. The brains were dissected in Hank's solution (10% HBSS (#SH30588.01, HyClone), 5% HEPES (H3375-25G,

Sigma-Aldrich), 5% sodium bicarbonate (S6014, Sigma-Aldrich), 1% penicillin/streptomycin (P/S; #1514012, Life Technologies) in pure water) and the extracted ventricular walls were cut into pieces manually, followed by enzymatic digestion (DMEM glutamax, 33% papain (Worthington 3126), 17% DNase at 10 mg/mL, 42% cysteine at 12 mg/mL) for 45 min at 37°C in a humidified 5% CO_2 incubator. Digestion was stopped by addition of a solution of trypsin inhibitors (Leibovitz Medium L15 (#11415064, Thermo Fisher), 10% ovomucoid at 1 mg/mL (#O2769, Sigma-Aldrich), 2% DNase at 10 mg/mL (#AM2222, Invitrogen). The cells were then washed in L15 and resuspended in DMEM glutamax supplemented with 10% fetal bovine serum (FBS; #SH30071.03HI, HyClone) and 1% P/S in a 1X Poly-L-lysine (PLL; #P8290, Sigma-Aldrich)-coated flask. Ependymal progenitors were left to proliferate for 5 days until confluence. Centrinone (#5065, Tocris) was added on day 3 of the proliferation phase, at a final concentration of 0.6 μM. On day 5 of proliferation, flasks were shaken (250 rpm) overnight. Then, centrinone was washed out 3 times with 1X PBS just before trypsinization and replating at high confluency ($1.8x10^5$ cells/20 μL) for MCC differentiation corresponding to days *in vitro* (DIV) in DMEM glutamax, 10% FBS, 1% P/S on 1X PLL-coated coverslips for immunocytochemistry experiments. The medium was replaced the following day with serum-free DMEM glutamax 1% P/S, to trigger ependymal differentiation *in vitro* (DIV0).

Immunostaining of cultured mouse ependymal MCCs

Cells were fixed at DIV4 in methanol at −20°C for 10 min. Coverslips were pre-blocked in 1X PBS with 0.2% Triton X-100 and 10% FBS before incubation with primary and secondary antibodies. Cells were counterstained with 10 μg/mL Hoechst (#B2261, Sigma) and mounted in Fluoromount (#0100-01, Southern Biotech). All antibodies used are listed in key resources table.

ALI culture of mTECs into multiciliated epithelia

The mTEC culture followed a standardized protocol as described before.[15] Briefly, mTECs were cultivated on transparent PET membranes of transwells (#353095, Life Science) in mTEC plus+RA medium (#MTEC/Plus, Merck Millipore). Once the cells on the apical side of the transwell chambers achieved 100% confluence, ALI was established by removing the culture medium from the transwell chambers. Subsequently, differentiation medium (mTEC Plus medium without fetal bovine serum and Y-27632 ROCK inhibitor (#S1049-SEL-10mg, Selleck Chem)) was added to the basal chamber of 24-well plates. The mTEC cells were maintained on transwells by replenishing the differentiation medium in the basal chamber every two days.

Generation of wild-type and L > D mutant MCIDAS

Coding sequence of human MCIDAS was cloned into pLVX vector (#632159, Addgene) with *Xho*I and *Xba*I restriction enzymes sites and a single HA tag sequence fused to the N terminus of the protein. L > D mutation of MCIDAS was generated using site-directed mutagenesis.

HEK293T cell culture and transfection

HEK293T cells were cultured in DMEM with 4500 mg/L glucose (Sigma-Aldrich, #G7258) and 10% FBS. When cells reached approximately 80% confluence, 3 μg of plasmid DNA was diluted into 300 μL opti-MEM (Life Technologies, #31985070) together with 8 μL of lipofectamine 2000 (Invitrogen, #11668019) and incubated for 10 min at room temperature. Subsequently, the mixed solution was applied to the cells. After overnight transfection, the transfection medium was replaced with normal DMEM medium.

Leptomycin B treatment of HEK293T cells, mTECs and hESCs

Cells were treated with leptomycin B (Sigma-Aldrich, #L2913) at a final concentration of 20 ng/mL. For HEK293T cells, after transfection overnight, leptomycin B was applied to cells for 24 h. Subsequently, the cells were fixed for immunostaining or harvested for RNA extraction. For mTEC and hESCs, leptomycin B was administered at the end of ALI day 1 and maintained till the end of ALI day 2.5, when the cells were fixed for immunostaining.

RTqPCR analysis of MCIDAS targets genes

SuperScript III First-Strand Synthesis System (Invitrogen, #18080051) was used for the generation of cDNA. Gene-specific primers for qPCR were designed using Primer3 software (Primer3 v.0.4.0) and are listed in key resources table. qPCR was conducted using EXPRESS SYBR GreenER Super Mix (Invitrogen, #A10315) on an Applied BioSystems 7900HT Fast Real-Time PCR System using SDS2.4 software. Each sample was analyzed in technical triplicate. Gene expression fold differences were calculated from the Ct values using Microsoft Excel, with normalization against the internal control, *Gapdh*/*GAPDH*.

Lentivirus generation and infection

Gene expression lentiviruses were produced with 3rd generation lentiviral system. Briefly, PLVX vector containing the coding sequences of specific genes was transfected together with pMD2.G (Addgene, #12259), pRSV-REV (Addgene, #12253) and pMDLg/pRRE (Addgene, #12251) vectors into HEK 293T cells. 3 days after transfection, the viruses were harvested. Viral titration was performed by infecting HEK293T cells with a GFP expression lentivirus which was generated together with the other gene expression lentiviruses. Viral titer was determined according to the percentage of GFP positive HEK293T cells three-days post-infection. For infection of mTECs, the cells were treated with 12mM EGTA (Sigma-Aldrich, #E3889) in 10 mM HEPES buffer (Sigma-Aldrich, #H3375-25G), pH 7.4, for 25 min at 37°C. Following PBS washes, cells were exposed to a mixture of lentivirus and Polybrene

(Sigma-Aldrich, #H9268; final concentration 5 μg/mL) in the culture medium. The cells and virus mixture were then centrifuged at 1500Xg for 80 min at 32°C, and subsequently incubated at 37°C in a cell culture incubator.

Whole transcriptome sequencing to identify MCIDAS regulated genes

Whole mRNA was extracted from 18 samples (3 replicates for each of 3 time points for wild-type and *Mcidas* mutant mTECs) and sequencing library was prepared with NEBNext Ultra II directional library preparation kit (NEB, #E7760S), combined with the NEBNext Poly(A) mRNA isolation kit (NEB, #E7490S). mRNA sequence data was generated through Illumina NextSeq and ~40 million, 76bp single-end reads were obtained.

hESC culture

H9 hESCs, obtained from WiCell Research Institute, Inc., were maintained in feeder-free conditions on either Biocoat Growth Factor Reduced Basement Membrane Matrigel (Corning, #354230) coated plates in mTeSR1 basal medium (Stem Cell Technologies, #85850) or vitronectin XF (Stem Cell Technologies, #07180) coated plates in Essential 8 (E8) basal medium (Life Technologies, #A1517001). Cells were passaged with Gentle Cell Dissociation Buffer (Stem Cell Technologies, #07174) every 4–5 days and kept in an incubator at 37°C in 5% CO_2.

Construction of CRISPR single guide RNA (sgRNA) vectors and donor templates

CRISPR single guide (sgRNAs) targeting human *E2F5*, *MCIDAS* and *GMNC* genes were designed using CHOPCHOP (https://chopchop.cbu.uib.no/) and cloned into a pCAG-Csy4-T2A-Cas9-HF1-IRES-BSD-U6 (Addgene, #79144) plasmid backbone. All gRNA sequences can be found in the key resources table. Donor templates containing a 3 x HA tag and 800 bp flanking homology arms were generated by individual PCR. The PCR amplicons for both 5′ and 3′ homology arms were PCR purified and used as a DNA template for a fusion PCR. The resulting amplicon was ligated into a zero blunt vector (Life Technologies, #K2700-20) and sequenced to confirm correct DNA insertion. Site directed mutagenesis using QuickChange II XL (Agilent Technologies, #200521) was used to alter the PAM (NGG) sequence on the corresponding donor vector, without altering the resulting codon. Homology arm and SDM primers can be found in key resources table.

Nucleofection of hESCs

Editing and donor template vectors were delivered to H9 hESCs using a Nucleofection Amaxa P3 primary Cell 4D Nucleofector Kit (Lonza, #V4XP-3024). Briefly, $2x10^6$ cells were disassociated in Accutase (Stem Cell Technologies, #07920) and resuspended at the appropriate cell density in P3 solution in a nucleofection column. 8 μg of plasmid DNA was added to the cells, the column was placed in a the Lonza Amaxa 4D Nucleofector (Lonza) and run under the program CA137. 24 h post-nucleofection, edited cell were selected by supplementing mTeSR1 medium with 5.25 μg/mL blasticidin (Life Technologies, #R210-01) for 48 h. Fresh mTeSR1 media was then fed to the cells every day until colonies got large enough for picking and genotyping.

Genotyping of CRISPR edited hESCs

Colonies were picked and incubated at 55°C for 1 h followed by a 5 min incubation at 95°C in a 20μL reaction containing 1X detergent (0.05% IGEPAL CA630 (Sigma-Aldrich, #I8896), 0.05% Tween 20), proteinase K (Qiagen, #19133) and 1X TE buffer (Thermo Fisher, #12090015) to extract genomic DNA. PCR reactions containing 1X Primestar Max Mastermix (Takara Bio, #R045A) and genotyping primers were set up with 1μL of genomic DNA. PCR products were subject to Sanger sequencing and analyzed using Inference of CRISPR Edits (ICE) software for mutations and their frequency. For the *E2F5* gene, a PCR amplicon containing the edited region of the compound heterozygous mutant was further cloned into pCR-Blunt vector and analyzed with Sanger sequencing to confirm each mutation. Genotyping and sequencing primers are listed in Table S6.

Clonal expansion of CRISPR edited hESCs

H9 clones that were found to contain frame-shifting mutations were subcloned by dissociating them into single cells using Accutase and seeded at very low cell densities onto matrigel-coated plates in mTeSR1 medium supplemented with CloneR (Stem Cell Technologies, #05888) for 48-72h. Cells were subsequently fed with fresh mTeSR1 medium daily until colonies grew large enough for picking and genotyping. Once genotypes were confirmed, clones were expanded in culture and then frozen stocks were made using Cryostor (Stem Cell Technologies, #07930).

Karyotype analysis

All H9 hESC lines (*E2F5* and *MCIDAS* mutant, *HA-MCIDAS* and *HA-GMNC* knock-in and their wild-type counterparts) were cultured to 70–80% confluency and sent to the KK Women's and Children's Hospital (Singapore) cytogenetics department for karyotype analyses.

Differentiation of hESCs into multiciliated airway epithelium

H9 hESCs were dissociated into single cells with Accutase for 5 min at 37°C after which $4x10^5$ cells were seeded as monolayer into each well of a 0.1% gelatin (Sigma Aldrich) and MEF coated 12-well plate (Falcon). The next day (Day −1, (D-1)) cells were fed fresh E8

media. On D0 the induction of the definitive endoderm (DE) lineage was intiated[72] by culture for 48h in a base medium (hereafter called CDM-PVA) consisting of F-12 (Life Technologies, #31765068), IMDM (Life Technologies, #21980065), polyvinyl alcohol (Sigma Aldrich, P8136-250G), concentrated lipids (Life Technologies, #11905031), monothioglycerol (Sigma Aldrich, #M61545-100ML), Transferrin (Sigma Aldrich, #T1147) and insulin (Roche, #1376497); which was supplemented with 100 ng/mL Activin A (Peprotech, #120-14P-50), 80 ng/mL FGF2 (Peprotech, #100-18B-100), 10 ng/mL BMP4 (Peprotech, #120-05-50), 3 μM CHIR99021 (Tocris, #4423/10) and LY294002 (Promega, #V1201). For the following 24h, cells were cultured in RPMI medium (Life Technologies, #61870-010) supplemented with B27 (Life Technologies, #17504-044), non-essential amino acids (Gibco, #11140-035), 100 ng/mL Activin A and 80 ng/mL FGF2.

The DE cells were then differentiated into Anterior Foregut Endoderm (AFE) and subsequently into lung progenitors.[73] To drive DE cells toward the AFE, cells were dissociated in gentle cell dissociation buffer, passaged into MEF-coated tissue culture plates, cultured for 72 h in a base medium (hereafter called BFG) consisting of Advanced DMEM F12 (Life Technologies, #12634028), B27, retinoic acid (RA; Sigma Aldrich, #R2625-50MG), *N*-2 (Life Technologies, #17502-048), 10 mM HEPES, GlutaMAX (Life Technologies, #35050-061) and supplemented with P/S, 2 μM Dorsomorphin dihydrochloride (DSM; Tocris, #3093) and 10 μM SB431542 (Tocris, #1614). To differentiate AFE cells into lung progenitors, cells were cultured in the BFG medium supplemented with 3 μM CHIR, 10 ng/mL BMP4, 10 ng/mL FGF7 (Peprotech, #100—19-100), 10 ng/mL FGF10 (Peprotech, #100-26-100) and 50 nM RA for 9 days.

Lung progenitors were differentiated into proximal lung organoids according to published protocol.[19] The D15-sorted NKX2-1+ve lung progenitors were resuspended in BFG medium containing 250 ng/mL FGF2, 100 ng/mL FGF10, 50 nM dexamethasone (Sigma-Aldrich, #D4902-100MG), 0.1 mM 8-bromoadenosine 30,50-cyclic monophosphate sodium salt (cAMP; Sigma-Aldrich, #B5386-100MG) and 0.1 mM 3-isobutyl-1-methylxanthine (IBMX; (Sigma-Aldrich, #17018-100MG), hereafter called proximal airway medium. The cell suspension was mixed with undiluted hESC-qualified growth factor-reduced matrigel (Corning, #354277) at a 1:1 ratio, resulting in a density of 1000 cells per μL and 40 μL of this mixture was pipetted in droplets onto 16 mm glass coverslips (Marienfeld Superior) in the base 24-well plate to create a matrigel domes. Matrigel domes were then incubated at at 37°C for 30 min to solidify, before 500 μL proximal airway medium was overlaid over each dome. Medium was supplemented with 10 μM Y-27632 for the first 24 h to improve cell survival. Proximal lung organoids were fed fresh media every other day for 9 days. Once confluent the Matrigel domes were passaged at a ratio of 1:6 into PneumaCULT Ex Plus Basal Medium (PEX; Stem Cell Technologies, #05040) supplemented with 1 μM A 83-01 (Stemcell Technologies, #100–0245), 1 μM DSM and 10 μM Y-27632 for a further 7–9 days, before they were harvested and sorted for basal cells.

$1.5x10^5$ basal cells collected from sorting proximal lung organoids were seeded onto 6.5 mm matrigel-coated transwell inserts (Stem Cell Technologies, #38024) with PEX medium in the apical and basal chambers. When the transwells reached 100% confluence, media was removed from both chambers and PneumaCULT ALI medium (Stem Cell Technologies, #05022) was added only to the basal chamber. Medium was refreshed every other day for 2 weeks or till cells were harvested for downstream processing/analyses at specific ALI time points.

Fluorescence-activated cell sorting (FACS)

Cells were harvested by incubation with Tryple Express (Thermofisher, #12604021) for 10–15 min at 37°C. Airway basal cells embedded in Matrigel-based 3D culture were additionally disassociated into single cells by resuspension using a p1000 pipette. The Tryple-cell mixture was diluted in Advanced DMEM F12 medium containing 10% FBS at a ratio of 1:2. Harvested cells were centrifuged at 1200 rpm for 5 min and resuspended in a FACS buffer consisting of Hank's Balanced Salt Solution (HBSS), 2% FBS, 25mM HEPES 2mM EDTA, P/S and supplemented with 10μM Y-27632. Cell clumps were removed by passing the cell suspension through a 40 μm cell strainer (Fisher Brand, #22363547) before cells were counted using the Scepter Cell Counter (Merck Millipore, #PHCC20060) and concentrations were adjusted to $1x10^6$/100 μL FACS buffer. For purification of lung progenitors at Day 15, 0.5 μL isotype control or 0.5 μL human ALCAM/CD166 phycoerythrin (PE)-conjugated antibody (Clone 105902, R&D Systems, #FAB6561P) was added per 100 μL cell suspension. Airway basal cells were purified on D25-30, where 0.5 μL of each isotype control or 0.5 μL human NGFR/CD271 (APC)-conjugated antibody (Biolegend, #345108) and 0.5 μL EpCAM/CD326 (PE)-conjugated antibody (Clone REA764; Miltenyibiotec, #324206) were added to 100 μL of cell suspension. Cells were stained for 30 min in the dark at 4°C, washed with FACS buffer and resuspended in 500 μL FACS buffer for sorting. Unstained and isotype-stained H9-derived lung progenitors or lung cells were used as negative controls for the gating parameters on the cytometer. Cells were analyzed for forward and side scatter and gated to exclude dead cells/debris. This population of live cells were further analyzed for FSC-H and FSC-A to isolate single cells, which were subsequently gated for CD166 or NGFR/EpCAM. For these sorting protocols a BD FACSAria Fusion Cell Sorter and BD FACSDiva software (BD Biosciences) was used and cells were collected in basal medium supplemented with 10 μM Y-27632.

Immunofluorescence analysis of mTEC and hESC-derived airway MCCs

Cells were fixed using either 4% paraformaldehyde (PFA) for 20 min at room temperature or ice-cold 100% methanol for 15 min at 4°C, depending on the specific antibodies used for staining. Following fixation, cells were treated with blocking and permeabilizing buffer (3% BSA, 0.1% Triton X-100 in DPBS) overnight at 4°C. Next, cells were incubated with primary antibodies diluted in staining buffer (3% BSA, 0.1% Triton X-100 in DPBS) for 1 h at room temperature. After 3 washes with PBS, cells were incubated with fluorophore-conjugated secondary antibodies diluted in staining buffer (3% BSA, 0.1% Triton X-100 in DPBS) for 1 h in the dark at room

temperature. Subsequently, cells were subjected to 3 additional washes with PBS. Following washing, the PET membrane with cells were mounted on glass slides and imaged using a Leica FV-3000 confocal microscope. Image processing and analyses were done using ImageJ. The E2F4 antibody has been previously validated for use on human and mouse cells (see company data sheet and ref.[16]). The E2F5 antibodies were validated for use with human cells in this study (see Figure S1). The Atlas E2F5 antibody (HPA065441) was also used to detect E2F5 expression in mTECs using immunofluorescence, and the antigen sequence is 96% identical between the mouse and human E2F5 proteins.

Western blotting

MTECs (wild-type control and *Mcidas* knock-out) and hESC-derived ALI cultures (wild-type control, *E2F5* knock-out, *HA-MCIDAS* knock-in and *HA-GMNC* knock-in) grown on transwells (9 transwells for each genotype) were lysed in 100 μL Pierce RIPA buffer (ThermoFisher Scientific, #89900) supplemented with complete protease inhibitor cocktail (Calbiochem, #539134) to prepare whole cell lysates. 10 μg protein was resolved on a 10% SDS PAGE gel and then transferred onto PVDF membranes. Membranes were blocked with 3% BSA, 0.1% Tween 20 (Sigma-Aldrich, #P9416) in PBS at 4°C overnight. The membranes were then incubated with specific primary antibodies for 1 h at room temperature, washed 3 times in PBS and then incubated with HRP-linked secondary antibodies for 1 h at room temperature. Pierce ECL Plus Substrate or SuperSignal West Femto Maximum Sensitivity Substrate (ThermoFisher Scientific, #340925) were used to visualize proteins on Chemidoc (Bio-Rad, #12003154).

Live imaging of human MCC ciliary motility

Transwell membranes with live, mature MCCs (ALI day14) derived from H9 hESCs were excised and mounted on glass slides. The mounted cells were examined using a Olympus FV3000 upright confocal microscope at 25°C. 20X objective was used to locate motile cilia, followed by recording of ciliary motility with a 100X oil emersion lens. The movies were recorded at 100 frames/s and processed with FV31S-AW FLUOVIEW software.

QUANTIFICATION AND STATISTICAL ANALYSIS

Whole transcriptome sequence analysis to identify MCIDAS regulated genes

Primary QC was performed on the raw data through FastQC (https://www.bioinformatics.babraham.ac.uk/projects/fastqc/). The QC passed data were aligned to GRCm39 primary assembly (https://ftp.ebi.ac.uk/pub/databases/gencode/Gencode_mouse/release_M26/GRCm39.primary_assembly.genome.fa.gz) with Gencode vM26 (GRCm39) transcript model (https://ftp.ebi.ac.uk/pub/databases/gencode/Gencode_mouse/release_M26/gencode.vM26.annotation.gtf.gz) and STAR aligner (version 2.6.0c).[52] An average of 88% uniquely mapped reads (range: 87.67–88.64%) were obtained. The raw count data for 53647 genes were generated through HTSeq (version 0.11.0)[53] from the aligned reads. A total of 15482 genes (13098 protein coding and 2384 non-coding) that were expressed at least one-copy in at least 2 of 3 replicates in either case or control of at least 1 out of 3 timepoints were considered for downstream analysis. Differential gene expression (DGE) analysis was performed for these genes between *Mcidas* mutant and wild-type control separately for each of the 3 time points using method encoded in DESeq2 package.[54] A gene was considered to be differentially expressed if absolute fold-change of >2 with FDR-corrected p-value <0.1 was obtained.

Gene set enrichment and cellular colocalization analysis

Gene set enrichment analysis was performed using Clugo (https://ieeexplore.ieee.org/document/1565762) and Cluepedia[57] modules available in Cytoscape (https://cytoscape.org/) package (version 3.10.1) against GO gene sets (tree level >3 and <8) for biological processes, cellular and molecular functions. Initially, a given mouse gene set was mapped to GO terms for enrichment through right-sided hypergeometric test for significant association, followed by Bonferroni step down p-value correction. GO terms with at least 4% enrichment and minimum of 3 genes enriched with $p < 0.05$ were considered for further grouping through ClueGO and Cluepedia. Enriched GO terms with 50% overlapping genes were grouped, and the name of the merged group was chosen based on the highest kappa score. A fraction of our query genes was not mapped to GO-BP. We further mapped these genes to (a) CellMarker 2.0 database[58] and (b) Human Protein Atlas (HPA: https://www.proteinatlas.org/) (using human orthologous genes converted through Ensembl BioMart). To generate further insights into the possible cellular colocalization of proteins, we have utilized a machine learning (ML)-based algorithm DeepLoc 2.0 (https://services.healthtech.dtu.dk/services/DeepLoc-2.0/) with protein sequence fasta (longest isoform of the human orthologues) downloaded from NCBI with default parameters. All deregulated genes in our *Mcidas* mutant versus wild-type analysis were queried with OMIM (https://www.omim.org/) dataset to identify association with genetic disorders.

NLS, NoLS, NES and protein-protein interaction (PPI) predictions

Protein sequences of human and mouse GMNC (accession numbers A6NCL1, Q3URY2) and MCIDAS (accession numbers D6RGH6, Q3UZ45) were downloaded from Uniprot.[74] To predict NLS, NoLS and NES motifs, we utilized several bioinformatics methods based on Hidden Markov Models (HMMs) or machine learning. We employed NLSeer, NLStradamus and DeepLoc 2.0 to detect NLS motifs.[61,62,75] For NES prediction, we utilized NESmapper and DeepLoc 2.0, whereas for NoLS, we utilized NoD detector.[27,61,63] Protein 3D structure of the NES top candidate motif was generated by trROSETTA.[64] Saliency logos were generated by WebLogo3.[65]

To build the MCIDAS-PLK4 and the CRM1-NES interaction models, we obtained the X-ray crystal structures of the human CRM1/SNURPORTIN-1 complex (PDB ID: 3GB8), human PLK4 (PDB ID: 3COK) and of the MCIDAS:GEMININ heterodimeric parallel coiled-coil (PDB ID: 4BRY) from Protein DataBank (RCSB PDB, https://www.rcsb.org/).[68] Protein-protein docking predictions of the complexes were done running the LZerD Protein Docking Web Server.[66] The binding energies of the obtained complexes by LZerD were then calculated by using the HawkDock server based on MM/GBSA free energy decomposition.[67]

Statistical analysis

The experimental data were analyzed using GraphPad Prism software (GraphPad Software Inc., La Jolla, CA, USA). Numerical values are presented as mean ± standard error. For all experiments, at least 3 biological replicates were examined, except for the following analyses: 1 biological sample with 2 technical replicate for Figure 6; 1 biological sample, 2 technical replicates for Figures S3A–S3D, 1 technical replicate for Figures S3E–S3G 1 biological sample for Figures S8G and S8H and 3 technical replicates for Figure S10. Unpaired, two-tailed Student's *t* test was utilized for quantitative analyses. A significance threshold of $p < 0.05$ was applied to determine statistical significance ($^{*}p < 0.05$, $^{**}p < 0.01$ and $^{***}p < 0.001$).

Figure assembly

All figures were assembled using Adobe Illustrator CS4.

Cell Reports, Volume *44*

Supplemental information

Nuclear-cytoplasmic translocation of MCIDAS couples transcription with massive *de novo* centriole biogenesis in multiciliated cells

Hao Lu, Kim Jee Goh, Ee Kim Tan, Cameron T. James, Arnab Ghosh, Amélie-Rose Boudjema, Paolo Alberto Lorenzini, Colin D. Bingle, Sebastian Maurer-Stroh, Nidhan K. Biswas, Alice Meunier, N. Ray Dunn, and Sudipto Roy

Supplementary figures and legends

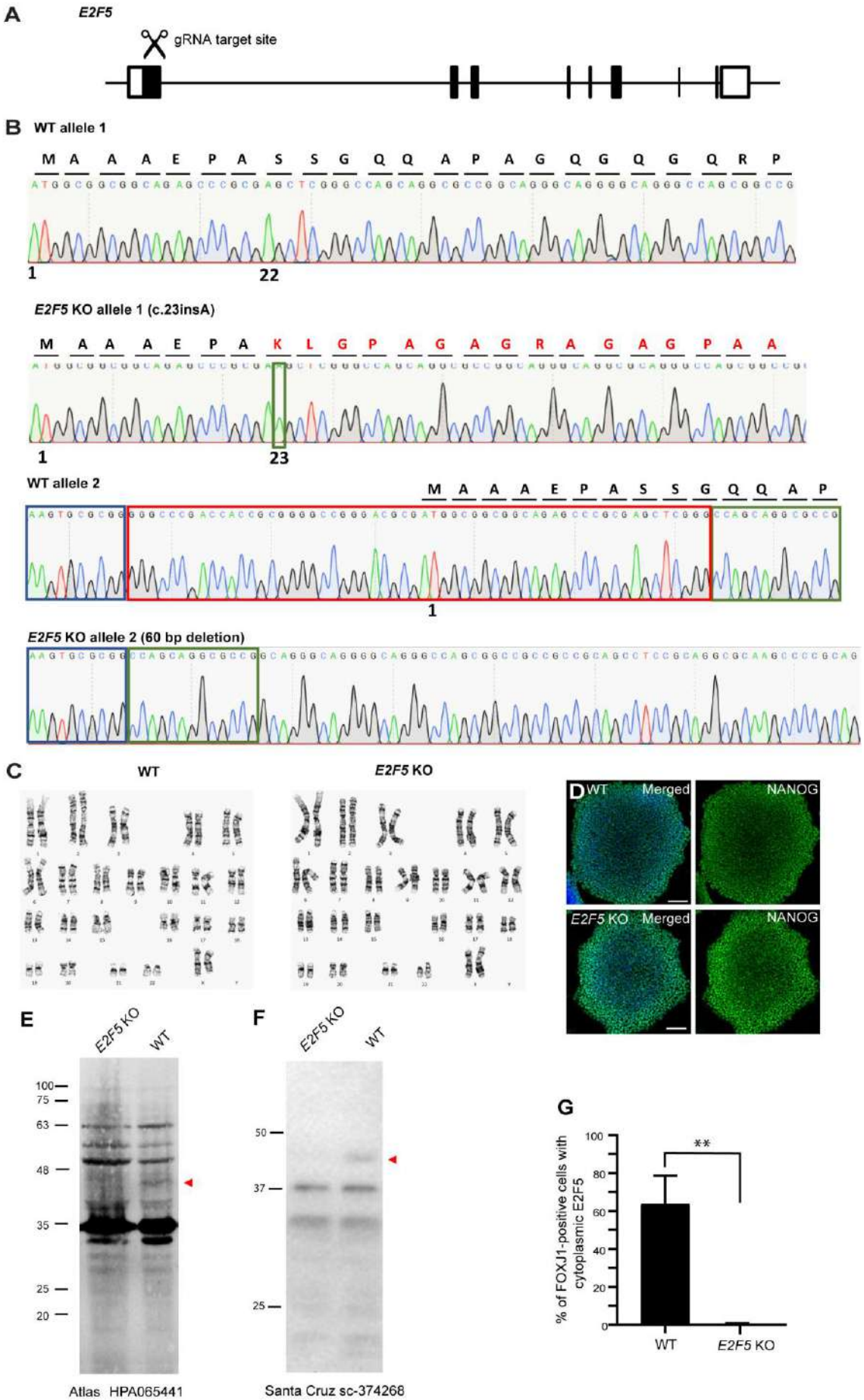

A
E2F5
gRNA target site
B
WT allele 1
M A A A E P A S S G Q Q A P A G Q G Q G Q R P
1
22
E2F5 KO allele 1 (c.23insA)
M A A A E P A K L G P A G A G R A G A G P A A
1
23
WT allele 2
M A A A E P A S S G Q Q A P
1
E2F5 KO allele 2 (60 bp deletion)
C
WT
E2F5 KO
D
WT
Merged
NANOG
E2F5 KO
Merged
NANOG
E
E2F5 KO
WT
100
75
63
48
35
25
20
Atlas HPA065441
F
E2F5 KO
WT
50
37
25
Santa Cruz sc-374268
G
% of FOXJ1-positive cells with cytoplasmic E2F5
100
80
60
40
20
0
**
WT
E2F5 KO

Figure S1. Generation of CRISPR edited *E2F5* mutant hESCs and validation with western blot. Related to Figure 1.

(A) Schematic of gene editing strategy showing sgRNA targeting exon 1 of the *E2F5* gene.

(B) Sanger sequencing chromatograms of *E2F5* sequence in wild-type (WT) and *E2F5* mutant (*E2F5* KO) H9 hESCs. Green box highlights 1 bp insertion in *E2F5* KO allele 1 at position 23. Red box highlights 60 bp deletion in *E2F5* KO allele 2. Blue and green boxes indicate sequences flanking the 60 bp deletion.

(C) Karyotype analyses of WT and *E2F5* KO hESCs.

(D) Immunostaining of pluripotency marker NANOG in WT and *E2F5* KO H9 hESCs. Scale bars = 100 μm. DAPI was used to highlight nuclei (blue).

(E, F) Western blot analysis of endogenous E2F5 protein expression in WT and *E2F5* KO hESCs with Atlas and Santa Cruz E2F5 antibodies. Antigen sequences recognized by the two antibodies are GCNTKEVIDRLRYLKAEIEDLELKERELDQQKLWLQQSIKNVMDDSINNRF and epitope mapping between amino acids 303-331 at the C-terminus, respectively, and are predicted to be lost from the mutant E2F5 proteins.

(G) Percentage of FOXJ1-positive cells with cytoplasmic E2F5. 3 biological replicates, 10 random microscope fields from each replicate (WT: 82/121, 67/143, 85/111; *E2F5* KO: 0/116, 0/78, 1/162; $**p=0.0019$). Unpaired, two-tailed Student's *t* test was used to compare quantitative analyses. $**p < 0.01$.

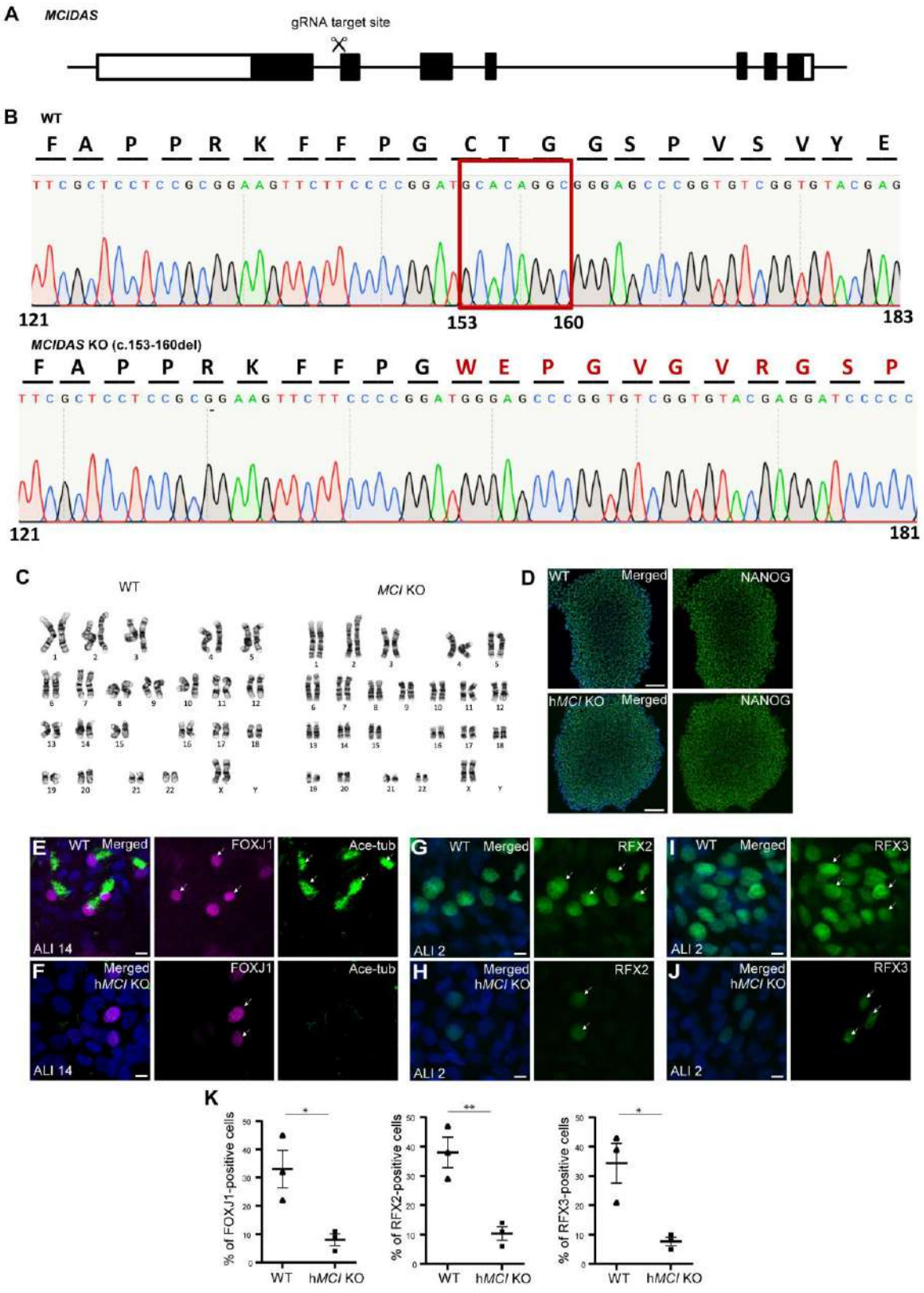


Figure S2. Generation and characterization of CRISPR edited *MCIDAS* mutant hESCs. Related to Figure 2.

(A) Schematic of gene editing strategy showing sgRNA targeting exon 2 of the *MCIDAS* gene.

(B) Sanger sequencing chromatograms of *MCIDAS* sequence in WT and CRISPR-edited *MCIDAS* mutant (*MCIDAS* KO) H9 hESCs. Red box indicates the 8 bp sequence deleted in both alleles.

(C) Karyotype analyses of WT and *MCIDAS* KO hESCs.

(D) Immunostaining of pluripotency marker NANOG in WT and *MCIDAS* mutant (*MCI* KO) ESCs. Scale bars= 100 μm.

(E) Wild-type hESCs differentiated into airway epithelia produced MCCs, labelled with FOXJ1 (arrows) and acetylated tubulin (Ace-tub) (arrows). Scale bar: 5 μm.

(F) *MCIDAS* mutant hESCs failed to differentiate into MCCs. MCC precursors expressed FOXJ1 (arrows), but no multiciliation was observed. Scale bar: 5 μm.

(G, H) RFX2 was expressed in hESC-derived *MCIDAS* mutant differentiating MCCs, but expression level appeared lower compared to wild-type (arrows). Scale bars: 5 μm.

(I, J) RFX3 was expressed in hESC-derived *MCIDAS* mutant differentiating MCCs, but expression level appeared lower compared to wild-type (arrows). Scale bars: 5 μm.

DAPI was used to highlight nuclei (blue) in panels D-J.

(K) Statistical analysis of percentage of FOXJ1-, RFX2- and RFX3-positive cells in wild-type (WT) and *MCIDAS* mutant (*MCI* KO) cells. 3 biological replicates, 8 random microscope fields from each replicate (WT FOXJ1: 82/181, 34/156, 57/179; *MCI* KO FOXJ1: 20/186, 19/211, 8/193; **p*=0.0231; WT RFX2: 74/201, 90/196, 50/178; *MCI* KO RFX2: 18/164, 19/235, 25/182; ***p*=0.0094; WT RFX3: 52/168, 92/214, 70/179; *MCI* KO RFX3: 18/182, 10/207, 15/188; **p*=0.0182). Unpaired, two-tailed Student's *t* test was used to compare the quantitative analyses. **p* < 0.05 and ***p* < 0.01.

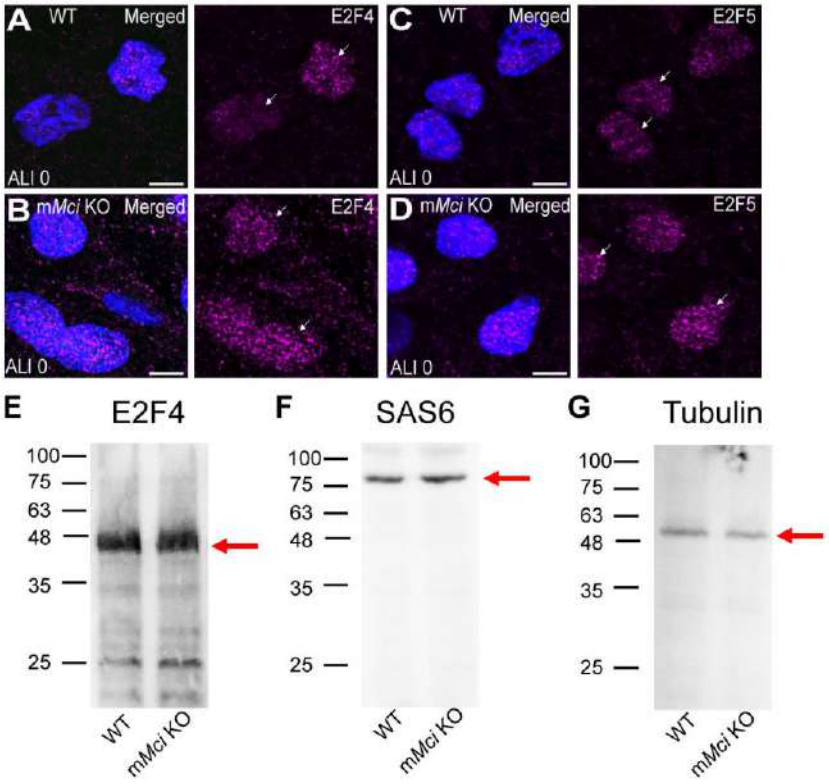


Figure S3. Expression of E2F4, E2F5 and SAS6 is not affected by loss of MCIDAS. Related to Figure 2.

(A, B) E2F4 expression (arrows) in wild-type (A) and *Mcidas* mutant (m*Mci* KO) mTECs (B) at ALI day 0. Note the nuclear localization, which is unaffected on loss of MCIDAS. Scale bars: 5 μm.

(C, D) E2F5 expression (arrows) in wild-type (C) and *Mcidas* mutant (m*Mci* KO) mTECs (D) at ALI day 0. Note the nuclear localization, which is unaffected on loss of MCIDAS. Scale bars: 5 μm.

DAPI was used to highlight nuclei (blue) in panels A-D.

(E-G) Western blot analysis of E2F4 (E) and SAS6 (F) expression level between wild-type and *Mcidas* mutant (m*Mci* KO) mTECs at ALI day 3. Tubulin was used as loading control (G).

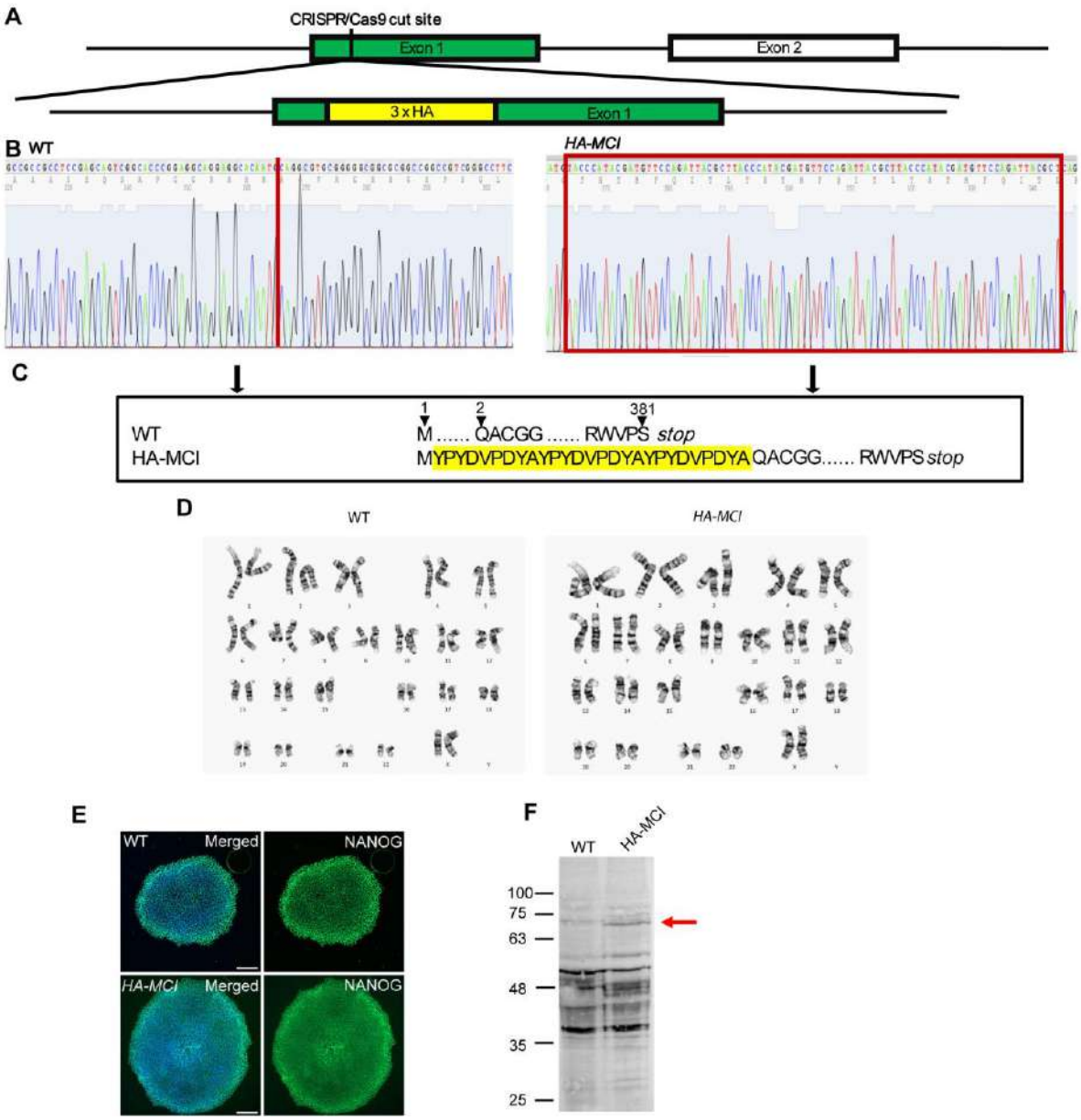


Figure S4. Generation of CRISPR edited HA-MCIDAS knock-in hESCs and validation with western blot. Related to Figure 3.

(A) Schematic of gene editing strategy, showing introduction of triple HA epitope tag in frame with the N-terminus of MCIDAS in hESCs.

(B) Sanger sequencing chromatograms of *MCIDAS* in wild-type (WT) and CRISPR-edited *HA-MCIDAS* (*HA-MCI*) hESCs.

(C) Amino acid sequence of HA-MCIDAS. The triple HA epitope sequence is highlighted in yellow.

(D) Karyotype analyses of WT and *HA-MCIDAS* hESCs.

(E) Immunostaining of pluripotency marker NANOG in WT and *HA-MCIDAS* hESCs. Scale bars= 100 μm. DAPI was used to highlight nuclei (blue).

(F) Validation of HA tagged MCIDAS protein using western blot.

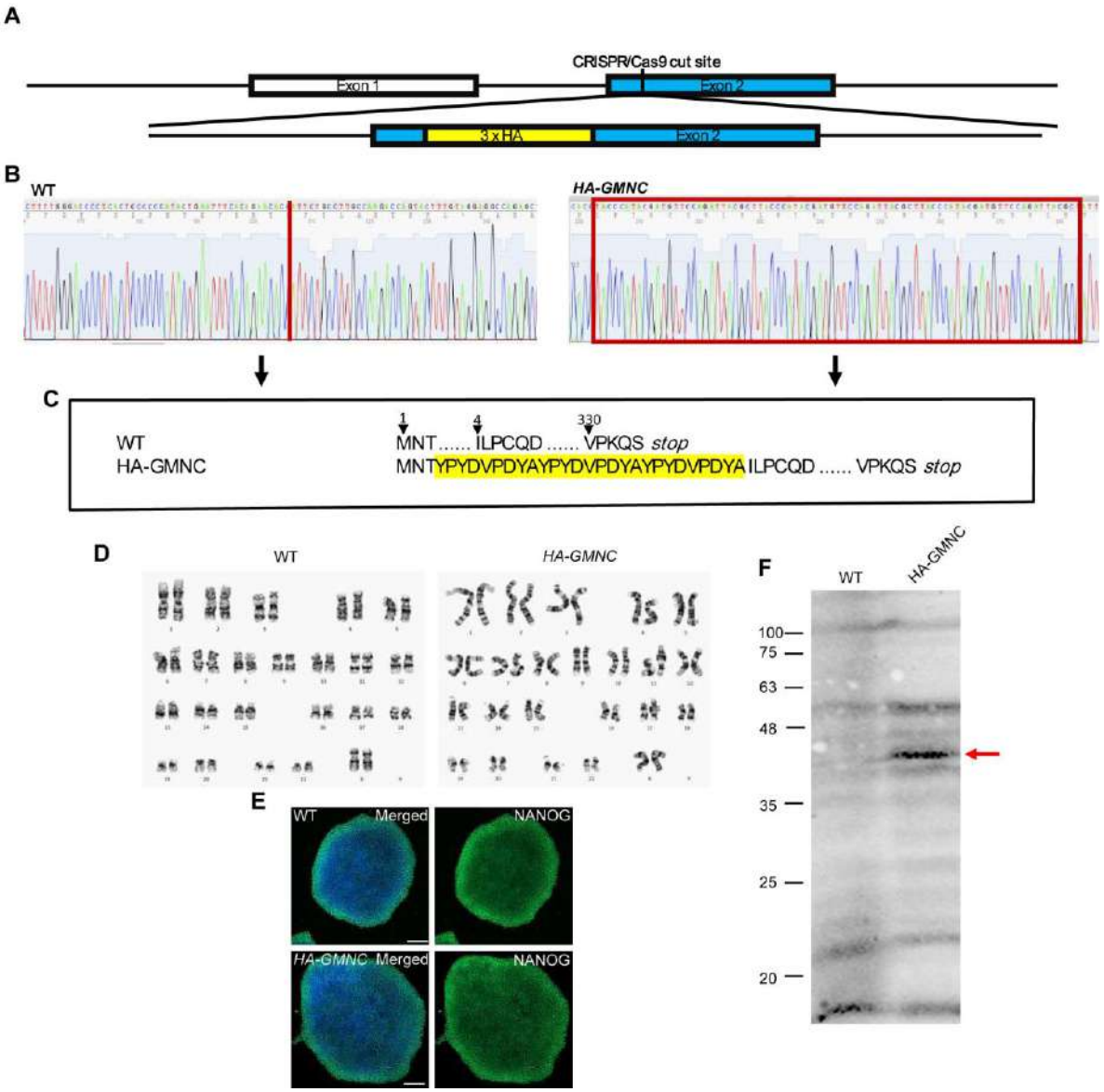


Figure S5. Generation of CRISPR edited HA-GMNC knock-in hESCs and validation with western blot. Related to Figure 3.

(A) Schematic of gene editing strategy showing introduction of triple HA epitope tag in frame with the N-terminus of GMNC in hESCs.

(B) Sanger sequencing chromatograms of *GMNC* sequence in wild-type (WT) and CRISPR-edited *HA-GMNC* hESCs.

(C) Amino acid sequence of HA-GMNC. The triple HA epitope is highlighted in yellow.

(D) Karyotype analyses of WT and HA-GMNC hESCs.

(E) Immunostaining of pluripotency marker NANOG in WT and *HA-GMNC* H9 hESCs. Scale bars= 100 μm. DAPI was used to highlight nuclei (blue).

(F) Validation of HA tagged GMNC protein using western blot.

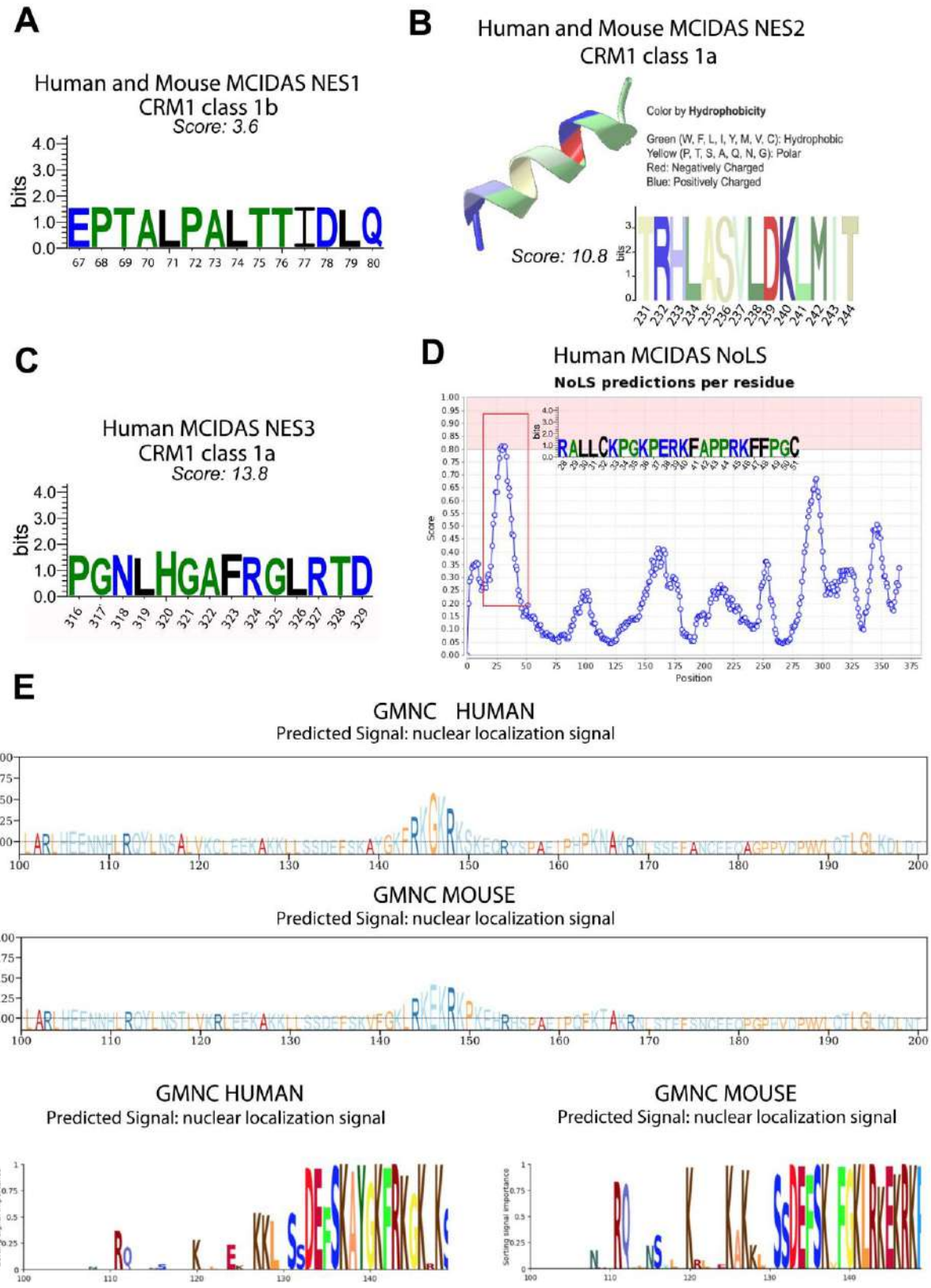


Figure S6. Motif analysis of NLS, NoLS, NES in human and mouse MCIDAS and GMNC proteins. Related to Figure 3.

(A, B) N terminal NES1 motif$^{67\text{-}80}$ (A) and NES2 top candidate motif$^{231\text{-}244}$ (B) as predicted by NESmapper in both human and mouse MCIDAS proteins. 3D protein structure of the top candidate NES2 motif was generated by trROSETTA. Amino acids are colored by type. Green indicates hydrophobic residues.
(C) C terminal NES3 motif$^{316\text{-}229}$ as predicted by NESmapper only in human MCIDAS protein.
(D) Per residue NoLS prediction by NoD detector for human MCIDAS.
(E) Saliency logos of the NLS motifs in human and mouse GMNC as predicted by DeepLoc 2.0 (top). Saliency logos of the NLS motifs in human and mouse GMNC as predicted by NLSeer (bottom).

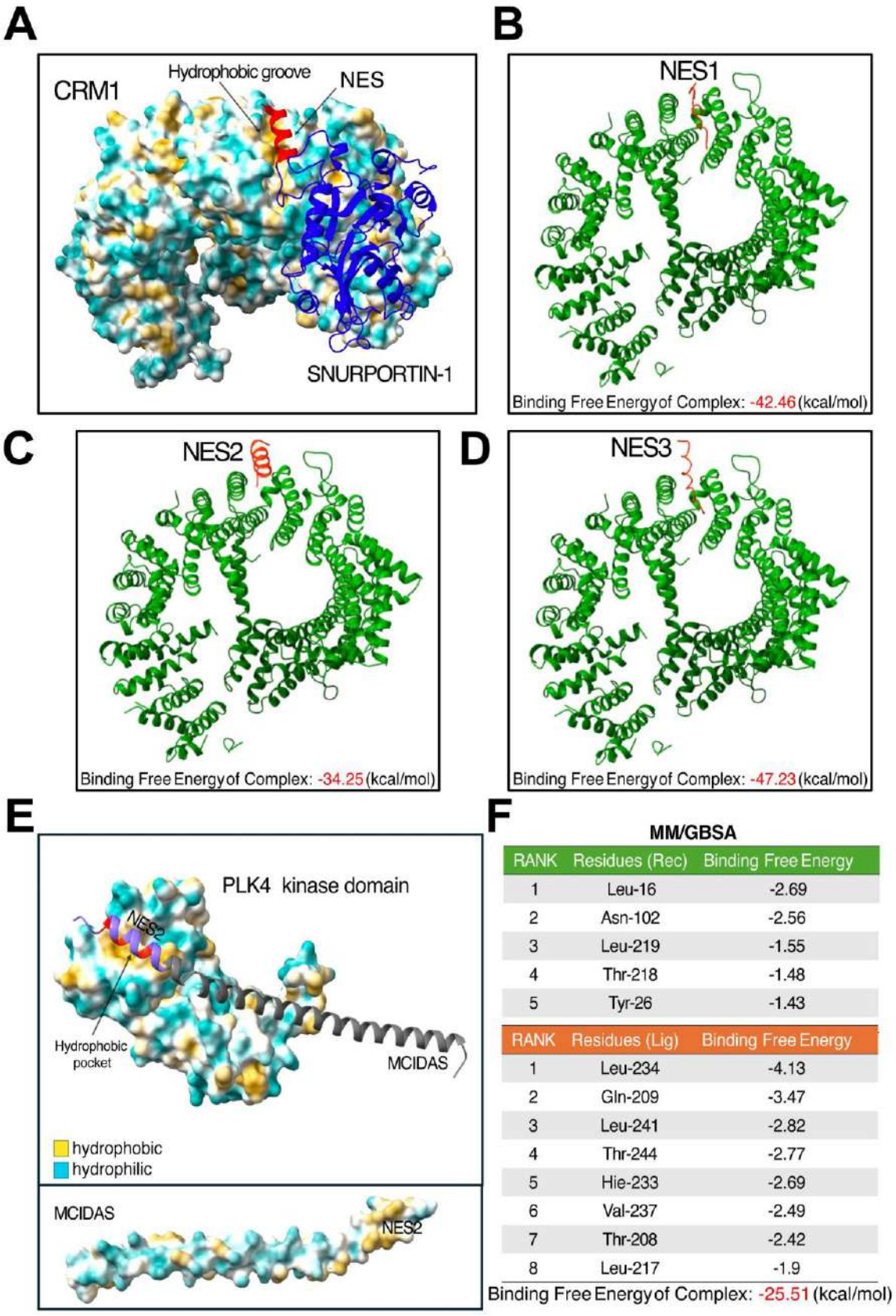


Figure S7. Protein-protein interaction models. Related to Figure 3.

(A) 3D x-ray crystal structure of human CRM1/SNURPORTIN-1 complex (3GB8 RCSB PDB). CRM1 is visualized as hydrophobicity surface (light blue and yellow), whereas SNURPORTIN-1 as ribbon style (dark blue), showing the location of its NES motif highlighted in red.
(B-D) Protein interaction models with binding free energy values in kca/mol of the three human MCIDAS NES motifs visualized as ribbon style (red), docked to human CRM1, visualized as ribbon style (green).
(E) Protein interaction model of human MCIDAS (4BRY RSCB PDB) visualized as ribbon style (gray) docked to human PLK4 kinase (3COK RSCB PDB) visualized as hydrophobicity surface (light blue and yellow). The location of the MCIDAS NES2 motif is colored in dark purple, whereas red highlights the position of the three Leucines (L) inside the NES2 motif.
(E') Side view of the hydrophobicity surface of human MCIDAS, highlighting the location of the hydrophobic NES motif.
(F) Per residue binding free energies in kcal/mol calculated by MM/GBSA of human PLK4-MCIDAS complex. On top (green), energies in kcal/mol of the receptor (PLK4), while at the bottom (orange), energies of the ligand (MCIDAS).

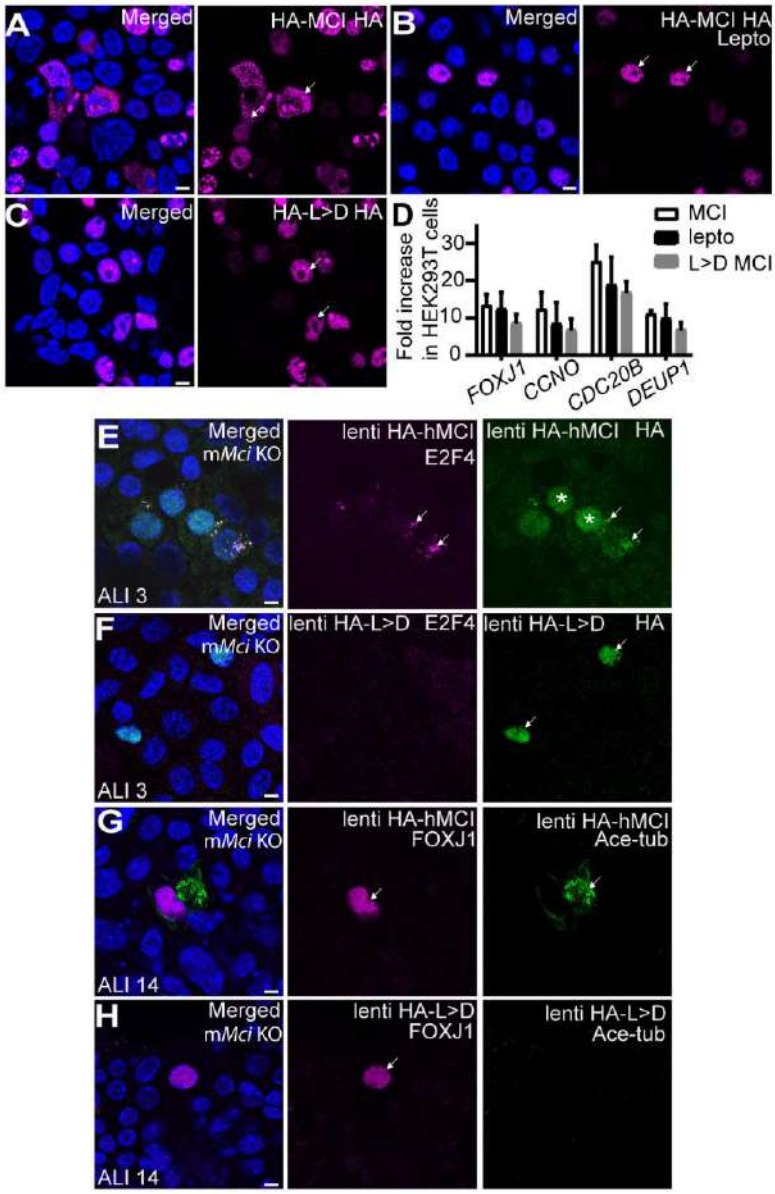


Figure S8. Centriole amplification in MCCs is activated by CRM1-dependent cytoplasmic accumulation of MCIDAS. Related to Figure 6.

(A) Nuclear and cytoplasmic localization of overexpressed wild-type MCIDAS in HEK293T cells. Cytoplasmic localization is indicated (arrows). Scale bar = 5 μm.

(B) Overexpressed MCIDAS was exclusively nuclear localized after leptomycin B treatment (arrows). Scale bar = 5 μm.

(C) L>D mutant MCIDAS was exclusively nuclear localized. Scale bar –5 μm.

(D) RTqPCR analysis of MCIDAStarget gene expression on overexpression of wild-type MCIDAS, overexpression of wild-type MCIDAS and leptomycin B treatment and overexpression of L>D mutant MCIDAS in HEK293T cells.
(E) Lentivirus-mediated overexpression of wild-type MCIDAS in *Mcidas* mutant mTECs rescued centriole assembly as revealed by cytoplasmic E2F4 accumulation (arrows). HA-MCIDAS was observed in the nucleus (asterisks) as well as the cytoplasm (arrows). Scale bar = 5 μm.
(F) Lentivirus-mediated overexpression of L>D mutant MCIDAS in *Mcidas* mutant mTECs failed to rescue centriole assembly as revealed lack of cytoplasmic E2F4 accumulation. L>D HA-MCIDAS was observed exclusively in the nucleus (arrows). Scale bar = 5 μm.
(G) Lentivirus-mediated overexpression of wild-type MCIDAS in *Mcidas* mutant mTECs rescued MCC formation as revealed by FOXJ1 expression (arrow) and multiple cilia formation (acetylated tubulin staining, arrow). Scale bar = 5 μm.
(H) Lentivirus-mediated overexpression of L>D mutant MCIDAS in *Mcidas* mutant mTECs induced high level of FOXJ1 expression (arrow) but failed to rescue multiciliation. Scale bar = 5 μm.
DAPI was used to highlight nuclei (blue) in panels A-C and E-H.

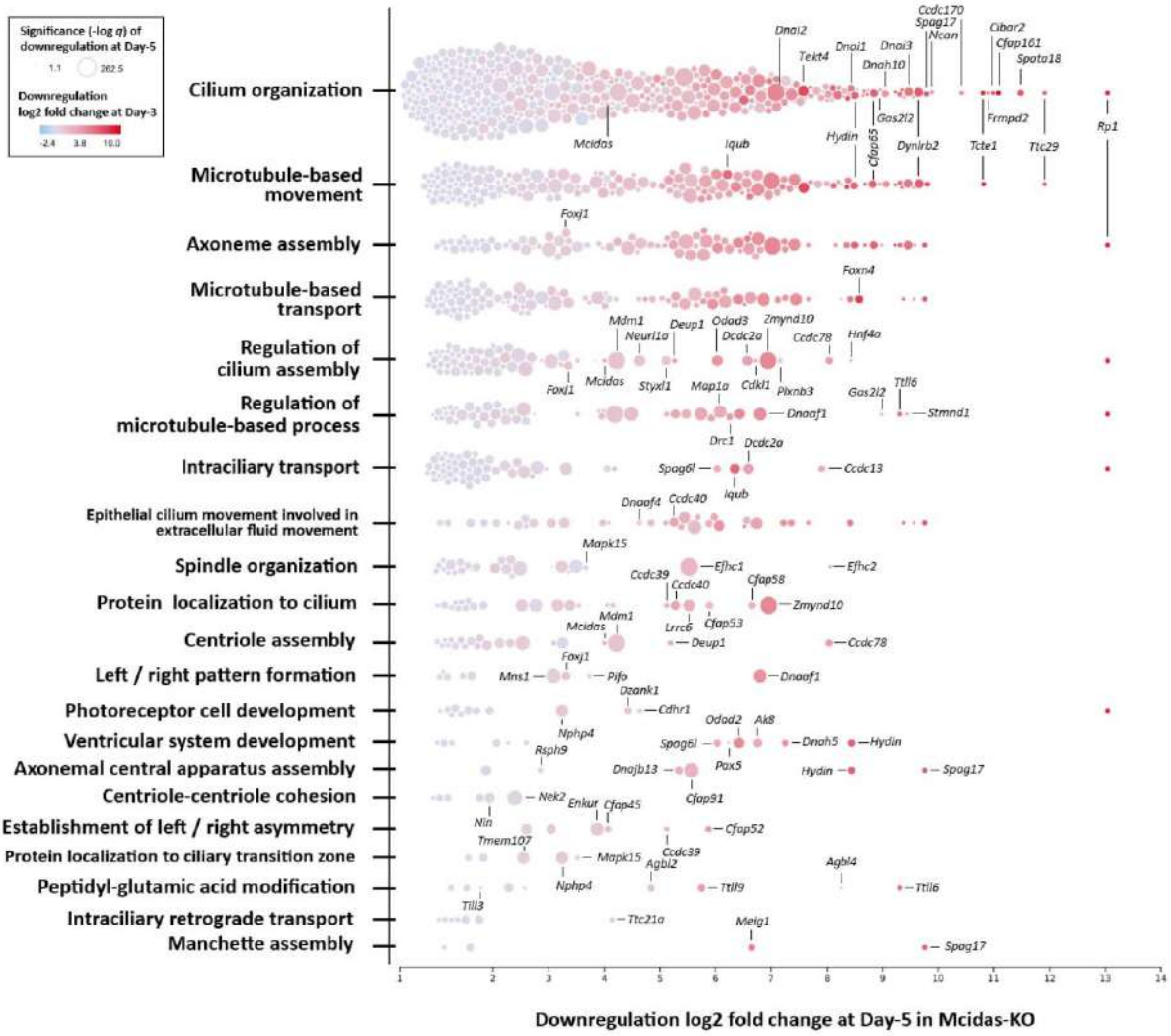


Figure S9. Time series gene expression data of significantly downregulated genes in *Mcidas* mutant mTEC culture with respect to 21 significantly enriched GO-BP modules.

The most significant GO-BP terms in each of these modules are highlighted. Each dot in the figure represents a gene in significantly enriched module, with position along X-axis quantifying increasing order of downregulation fold change (in log 2 scale) at ALI day 5 of *Mcidas* mutant (relative to wild-type) while the radius indicates significance level. The color gradient (blue to red) represents downregulation fold change status of the genes at ALI day 3. The top downregulated genes from each module are highlighted.

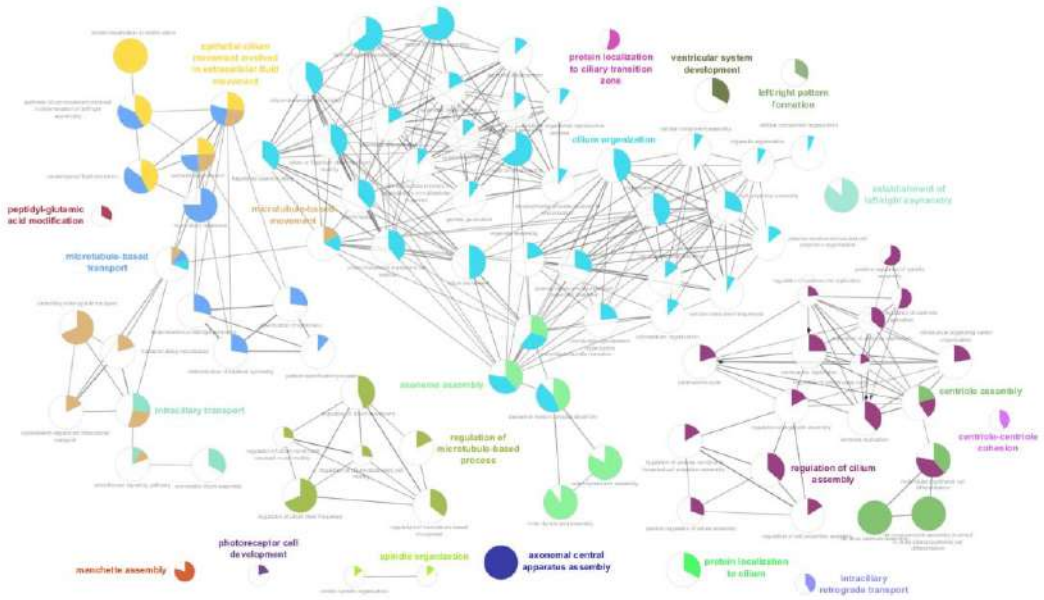


Figure S10. Enrichment of GO-BP with significantly downregulated genes (at ALI day 5) through ClueGO and CluePedia identified 21 significantly enriched modules, consisting of 107 GO-BP terms.

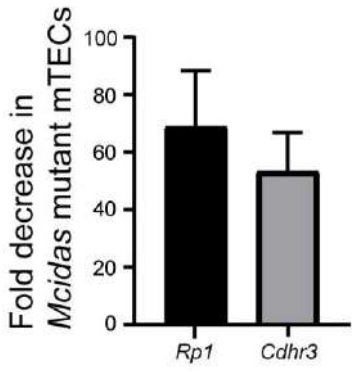


Figure S11. qRT-PCR analysis of MCIDAS target genes in mTECs.

Rp1 and *Cdhr3* expression in *Mcidas* mutant mTEC cultures was significantly lower than the wild-type.