Directional alignment of different cell types organizes planar cell polarity

Masaki Arata[1,*], Hiroshi Koyama[1,2], and Toshihiko Fujimori[1,2,*].

[1] Division of Embryology, National Institute for Basic Biology, Okazaki, Aichi, Japan.

[2] SOKENDAI (The Graduate University for Advanced Studies), Okazaki, Aichi, Japan.

Corresponding author: Masaki Arata and Toshihiko Fujimori

Email: fujimori@nibb.ac.jp (T.F.), arata@nibb.ac.jp (M.A.)

This PDF file includes:

Main Text
Figures 1 to 5

Abstract

Planar cell polarity (PCP) refers to unidirectional alignment of cell polarity in the plane of epithelial tissues. Previous studies in *Drosophila* wing epithelia have shown that imbalances in the amount of core PCP proteins (core proteins) between adjacent cells disrupt PCP. In contrast, various tissues maintain PCP despite containing multiple cell types with different levels of core proteins. The mechanisms by which these tissues resist imbalances in core protein levels and preserve PCP remain unclear. Here, we demonstrated that cell-type distribution plays a crucial role in resisting such imbalances. In the mouse oviduct, Secretory cells (SCCs) have low levels of core proteins, whereas multiciliated cells (MCCs) exhibit high levels. We found that SCCs are aligned along the ovary-uterus axis that is consistent with the orientation of the oviduct PCP. Furthermore, through a combination of mathematical modeling, deep learning, and statistical modeling, we showed that orientation of cell-type alignment is a key aspect of cell-type distribution that affects PCP. As predicted by these theoretical analyses, core proteins were locally misoriented where SCCs were not aligned along the ovary-uterus axis. Finally, we found that in oviducts lacking the core protein CELSR1, the alignment of SCCs was disrupted, suggesting that positive feedback regulation between core proteins and cell alignment contributes to PCP maintenance. Our findings highlight the overlooked role of spatial arrangement of cell-types in PCP maintenance that is essential for the proper function of epithelial tissues.

Significance Statement

Oocyte transport in the oviduct, driven by ovary-uterus oriented cilia beating, exemplifies how epithelial tissue functions rely on planar cell polarity (PCP)—the coordinated alignment of cell polarity within the plane of an epithelial sheet. PCP establishment requires intercellular interactions between core proteins, and mutant cells lacking these proteins can disrupt the polarity of neighboring cells. However, despite containing multiple cell types with different amounts of core proteins, various tissues establish and maintain PCP. In this study, we demonstrated that the spatial arrangement of these cells types plays a crucial role in the maintenance of PCP that is essential for functions of various tissues.

Main Text

Introduction

The proper function of various organs depends on planar cell polarity (PCP), the tissue-level alignment of cell polarity within the plane of epithelial tissue (1–8). For example, the transport of oocytes is driven by the ovary-to-uterus-directed beating of cilia in the oviduct epithelium (9–12), and misoriented hair cells in the cochlea are associated with hearing deficit (13).

Previous studies have identified an evolutionally conserved group of proteins, known as core PCP proteins (hereafter referred to as core proteins), that play key roles in PCP establishment (1–7). Core proteins form two types of complexes at cell edges (Fig. 1A and A'): one comprised of CELSR and FZD (FRIZZLED), and another comprised of CELSR and VANGL (Van Gogh-like; hereafter referred to as the FZD complex and the VANGL complex, respectively). In various developing organs, the FZD and VANGL complexes show asymmetric distributions at cell edges along body axes (1, 4, 8). For example, in the mouse oviduct, the FZD complex is localized to the uterus side (Fig. 1A and A'), while the VANGL complex is localized to the ovary side of the cell (10). This asymmetry is established before ciliogenesis, and the loss of CELSR1 results in misoriented ciliary beating (9), suggesting that the asymmetry is a prerequisite for PCP establishment. Although details remain unclear, experimental and theoretical studies suggest that the establishment of the asymmetry requires intracellular and intercellular interactions between the FZD and VANGL complexes (14–17).

Epithelial tissues are typically composed of multiple cell types. In such cases, how is PCP formed or maintained? The mouse oviduct epithelium consists of two cell types: multiciliated cells (MCCs) and secretory cells (SCCs, Fig. 1A). Notably, these two cell types differ in their levels of core proteins; CELSR1 is enriched at the cell edges of MCCs but is present at lower levels in SCCs (Fig 1B-B", (9)). This imbalance in core protein levels between adjacent cells (hereafter referred to as core protein imbalance) has the potential to disrupt PCP. Indeed, when cells lacking Frizzled (Fz, a fly homolog of FZD), Van Gogh (Vang, a fly homolog of VANGLs), or Flamingo (Fmi, also known as Starry night, a fly homolog of CELSR) are experimentally produced in the *Drosophila* wing, polarities in surrounding wild-type cells are reorganized (6, 7, 18). Nevertheless, in the oviduct, both the ciliary beating orientation and core protein distribution remain biased along the ovary-uterus axis. Such core protein imbalances have also been observed in various organs, including the mouse inner ear (19), airway (20), and *Drosophila* eye disc (21). However, the mechanisms by which tissues overcome core protein imbalance and maintain PCP remain unclear.

In the mouse oviduct epithelium, SCCs appeared to be aligned along the ovary-uterus axis (Fig. 1). This observation led us to hypothesize that the arrangement of SCCs and MCCs attenuates the disruptive effects of core protein imbalance on PCP. The impact of two-dimensional configurations of mutant cells lacking core proteins on PCP has not been thoroughly investigated *in vivo*. Moreover, previous theoretical studies have examined effects of such mutant cells on surrounding wild-type cells either in a one-dimensional cell row (22) or in a 2D cell sheet containing a mutant cell cluster with a simple shape such as a rectangle (14, 15). Therefore, the influence of cell-type distributions on PCP remains poorly understood. Furthermore, these studies primarily focused on elucidating

the molecular functions of core proteins during PCP establishment, rather than exploring the role of cell type distributions in PCP establishment and maintenance.

In this study, we analyzed the effects of various distributions of these cell types on PCP using a mathematical model to test our hypothesis. We demonstrated that the directional positioning of these cell-types plays a crucial role in maintaining PCP in the presence of core protein imbalance. Consistent with this finding, core protein distribution was misoriented in regions where SCCs were not aligned along the ovary-uterus axis. Finally, we found that the alignment of SCCs along the ovary-uterus axis was lost in *Celsr1* mutant oviducts, suggesting that positive feedback regulation between core protein and the alignment of SCCs contributes to PCP maintenance.

Results

Distribution of SCCs in the mouse oviduct

To investigate how epithelial tissues tolerate core protein imbalance, we focused on the mouse oviduct epithelium, which consists of only two cell types: secretory cells (SCCs) and multiciliated cells (MCCs). Previous studies reported that SCCs comprise 24% of the epithelial cells in the infundibulum of the adult mouse oviduct (11 weeks old; 1, 3), and that CELSR1 levels are lower in SCCs than in MCCs (Fig. 1B-B", 1). First, we quantitatively evaluated the imbalance in CELSR1 levels between SCCs and MCCs. To this end, we measured the intensities of CELSR1 immunofluorescent signals at cell edges. We found that CELSR1 signals at edges between MCCs were approximately twice as high as those at edges between SCCs (Fig. S1A-C'). To further validate whether SCCs have lower levels of core proteins, we examined the oviducts of transgenic mice ubiquitously expressing EGFP-fused VANGL2 (VANGL2:EGFP; *R26-VANGL2-EGFP* mice, 2). If other core proteins were expressed sufficiently in SCCs and distributed asymmetrically at cell edges, VANGL2:EGFP signals would be expected to exhibit asymmetric localization at cell edges. Consistent with the CELSR1 levels, the asymmetric distribution of VANGL2:EGFP at cell edges was not observed in SCCs (Fig. S1D-E'). These observations suggest that SCCs have fewer levels of core proteins than MCCs and do not establish the polarized distribution of core proteins at cell edges.

Next, we analyzed the distribution of SCCs in the oviduct. For this purpose, we used previously-published images in which cell edges and MCCs were labeled by ZO-1 and γ-TUBULIN immunofluorescent signals, respectively (Fig. 1C and C'; images were adopted from (11)). In the oviduct epithelium, SCCs form small clusters and these clusters scatter around the epithelium (Fig. 1C-C'). We noticed that these clusters elongate along the ovary-uterus axis. We measured the length of each SCC cluster by placing a bounding rectangle whose horizontal edges were aligned parallel to the ovary-uterus axis (Fig. 1C'). Hereafter, we denote cluster length horizontal and

vertical to the direction of PCP as l_{h} and l_{v}, respectively. In the oviduct, l_{h} and l_{v} correspond to ovary-uterus and circumferential lengths of an SCC cluster, respectively. The mean l_{h} was 2.45 cells (Standard deviation, SD: 1.16), which was significantly larger than the mean l_{v}, 1.85 cells (SD: 0.862; p<0.01, Wilcoxon rank sum test, Fig. 1D and D'). This result indicates that SCCs preferentially align along the direction of PCP in the oviduct.

Modeling PCP in a 2D cell sheet

The observation in the oviduct epithelium led us to hypothesize that directional positioning of SCCs mitigates the disruptive effect of core protein imbalance on PCP. To address this hypothesis, we developed a mathematical model (see also Supporting Information). In our model, the epithelial sheet is represented as an array of regular hexagonal cells (Fig. 2A-B'), where the FZD complex and the VANGL complex shuttle between cell edges and cytoplasmic regions (Fig. 2A, blue arrows).

$$F_{\mathrm{C}_k} \underset{R_{\mathrm{F}_{ki}}}{\overset{K_{\mathrm{F}_{ki}}}{\rightleftarrows}} F_{\mathrm{M}_{ki}} \tag{1}$$

$$V_{\mathrm{C}_k} \underset{R_{\mathrm{V}_{ki}}}{\overset{K_{\mathrm{V}_{ki}}}{\rightleftarrows}} V_{\mathrm{M}_{ki}} \tag{2}$$

Here, F_{C_k} and V_{C_k} denote the concentration of the FZD and VANGL complexes in the cytoplasmic region of *k*th cell, respectively. The concentration of the FZD and VANGL complexes at *i* th cell edge in the *k*th cell are represented as $F_{\mathrm{M}_{ki}}$ and $V_{\mathrm{M}_{ki}}$, respectively; $i = 1{\sim}6$. In addition, concentrations of the FZD and VANGL complexes were assumed to be uniform within each cell edge and cytoplasmic region. The FZD complex shuttles between cell edge i and a cytoplasmic region at rates $K_{\mathrm{F}_{ki}}$ and $R_{\mathrm{F}_{ki}}$, while the Vangl complex undergoes similar transport at rates $K_{\mathrm{V}_{ki}}$ and $R_{V_{ki}}$. Therefore, the dynamics of $F_{\mathrm{M}_{ki}}$ is describes as $dF_{\mathrm{M}_{ki}}/dt = -R_{\mathrm{F}_{ki}}F_{\mathrm{M}_{ki}} + K_{\mathrm{F}_{ki}}F_{\mathrm{C}_{ki}}$; the differential equations for other components are shown in Supplementary information.

According to previous studies, to establish the planar polarized distributions of the two complexes on cell edges, the formation of intercellular complexes composed of the two complexes was required (Fig. S2A), where the intercellular complexes have activities of both mutual inhibition and autoactivation (14–17). We assumed that these critical processes affected the above shuttling rates as follows:

$$K_{\mathrm{F}_{ki}} = \rho_{\mathrm{F}} + \omega_{\mathrm{F}}\left(F_{\mathrm{M}_{ki}}V'_{\mathrm{M}_{ki}} + d_{\mathrm{A}_{\mathrm{F}}}\sum_{j=i-1,i+1} F_{\mathrm{M}_{kj}}V'_{\mathrm{M}_{kj}}\right) \quad (3)$$

$$R_{\mathrm{F}_{ki}} = \mu_{\mathrm{F}} + \varphi_{\mathrm{F}}\left(V_{\mathrm{M}_{ki}}F'_{\mathrm{M}_{ki}} + d_{\mathrm{I}_{\mathrm{V}}}\sum_{j=i-1,i+1} V_{\mathrm{M}_{kj}}F'_{\mathrm{M}_{kj}}\right) \quad (4)$$

$$K_{\mathrm{V}_{ki}} = \rho_{\mathrm{V}} + \omega_{\mathrm{V}}\left(V_{\mathrm{M}_{ki}}F'_{\mathrm{M}_{ki}} + d_{\mathrm{A}_{\mathrm{V}}}\sum_{j=i-1,i+1} V_{\mathrm{M}_{kj}}F'_{\mathrm{M}_{kj}}\right) \quad (5)$$

$$R_{V_{ki}} = \mu_{\mathrm{V}} + \varphi_{\mathrm{V}}\left(F_{\mathrm{M}_{ki}}V'_{\mathrm{M}_{ki}} + d_{\mathrm{I}_{\mathrm{F}}}\sum_{j=i-1,i+1} F_{\mathrm{M}_{kj}}V'_{\mathrm{M}_{kj}}\right) \quad (6)$$

where the concentrations of FZD and VANGL complexes on a neighboring cell's edge juxtaposing edge i are denoted as $F'_{\mathrm{M}_{ki}}$ and $V'_{\mathrm{M}_{ki}}$,respectively (Fig. 2A). The intercellular complexes are formed between the FZD and VANGL complexes: $F_{\mathrm{M}_{ki}} - V'_{\mathrm{M}_{ki}}$ and $V_{\mathrm{M}_{ki}} - F'_{\mathrm{M}_{ki}}$. For simplicity, we assumed that the concentrations of the intercellular complexes were proportional to the terms of $F_{\mathrm{M}_{ki}}V'_{\mathrm{M}_{ki}}$ and $F_{\mathrm{M}_{ki}}V'_{\mathrm{M}_{ki}}$ in the above equations. The first term in the parenthesis of eq.3 and 5 correspond to the autoactivation activity: $K_{\mathrm{F}_{ki}}$and $K_{\mathrm{V}_{ki}}$ increase when $F_{\mathrm{M}_{ki}}V'_{\mathrm{M}_{ki}}$ and $V_{\mathrm{M}_{ki}}F'_{\mathrm{M}_{ki}}$ are elevated, respectively. The first term in the parenthesis of eq.4 and 6 corresponds to the mutual inhibition activity: $R_{\mathrm{F}_{ki}}$and $R_{\mathrm{V}_{ki}}$ increase when $V_{\mathrm{M}_{ki}}F'_{\mathrm{M}_{ki}}$ and $F_{\mathrm{M}_{ki}}V'_{\mathrm{M}_{ki}}$ are elevated, respectively. In addition, this model assumed a non-local effect, where the shuttling rates on cell edge i was affected by core proteins on neighboring edges $i-1$ and $i+1$ (Fig. S2A' and A"; the second term in the parenthesis of eq.3-6; see details in Supporting Information). A constant d determines the strength of the non-local effects (Fig. S2A' and A"). When $d_{\mathrm{A}} > 0$, the complexes on edge $i-1$ and $i+1$ elevate $K_{\mathrm{F}_{ki}}$and $K_{\mathrm{V}_{ki}}$. Similarly, when $d_{\mathrm{I}} > 0$, the complexes on edge $i-1$ and $i+1$ elevate $R_{\mathrm{F}_{ki}}$and $R_{\mathrm{V}_{ki}}$. Although non-local effects have not been experimentally validated, theoretical studies suggest they are necessary for robust PCP establishment (15, 23).

The parameter ρ represents the basal internalization rate of the FZD and the VANGL complexes from the cell edge to the cytoplasm, while μ denotes the basal trafficking rate of the FZD and VANGL complexes from the cytoplasm to the cell edge; both parameters were set to be constant. The values of parameters used in our model are listed in Table S1. We selected these values to reproduce the non-autonomous effects of mutant cells lacking core proteins, as observed *in vivo* (Fig. S3; see details in the Supporting Information).

Using the above mathematical model, the effects of core protein imbalances on PCP were assessed as follows. First, we prepared a cell sheet in which PCP was already established (Fig. 2B, left panel). Second, core protein levels were reduced in selected cells (low cells; gray cells in Fig. 2B; remaining cells are referred to as high cells in the following). The magnitude of the core protein imbalance was defined as a constant $balance_{\mathrm{LH}}$, which ranges from 0 to 1. When $balance_{\mathrm{LH}} = 1$, the amount of core proteins in low cells is equal to that in high cells, whereas, when $balance_{\mathrm{LH}} = 0$, low cells have no core proteins. Then, we performed simulations to analyze how the distribution of core proteins was reorganized. Fig. 2B is an example in the case that the low cells were randomly distributed, by which we will explain the setting of our simulations. At the initial state, the distributions of core proteins at cell edges were unidirectionally biased, with the FZD and VANGL complexes enriched at the right and left ends of each cell, respectively (Fig. 2B and S2B). In our data, a polarity vector, representing an orientation of FZD complex-enriched cell edges in each cell, is shown by an orange-headed bar (Fig. 2B", see details in Supporting Information). Length and angle of the bar represents the magnitude and the angle of the polarity, respectively. $\overline{\Delta angle_{\mathrm{high}}}$ is the mean absolute change in polarity vector angles of high cells in a cell sheet. Since the value of $\overline{\Delta angle_{\mathrm{high}}}$ is 0 at the initial state of the simulation, the value of $\overline{\Delta angle_{\mathrm{high}}}$ at the final state represents the extent of polarity reorganization. In Figure 2B, $\overline{\Delta angle_{\mathrm{high}}}$ is 36.0 degrees at the final state, indicating that the core protein imbalance disrupted PCP.

The direction of low-cell alignment is essential for mitigating the effect of low cells on PCP

To address whether the direction of low-cell alignment affects PCP, we placed rectangular low-cell clusters within a cell sheet and analyzed the degree of PCP reorganization after simulations. First, low-cell clusters were aligned in a lattice at intervals of 3 cells and $balance_{\mathrm{LH}}$ was set to 0 (i.e., no core proteins in the low cells). We then examined the effects of low-cell clusters with varying values of l_{h} and l_{v} (Fig. 2C-C'). Here, l_{h} and l_{v} represent the lengths of the cluster in the horizontal and vertical directions, respectively, relative to the X-axis, which corresponds to the initial direction of PCP. Our analysis revealed that when low-cell clusters were elongated perpendicular to the X-axis, the value of $\overline{\Delta angle_{\mathrm{high}}}$ increased (Fig. 2C, compare $l_{\mathrm{h}} = 3$, $l_{\mathrm{v}} = 3$ with $l_{\mathrm{h}} = 3$, $l_{\mathrm{v}} = 5$). In contrast, when low-cell clusters were elongated parallel to the X-axis, the values of $\overline{\Delta angle_{\mathrm{high}}}$ did not increase (Fig. 2C, compare $l_{\mathrm{h}} = 3$, $l_{\mathrm{v}} = 3$ with $l_{\mathrm{h}} = 5$, $l_{\mathrm{v}} = 3$).

Since low-cell clusters were not regularly positioned in the oviduct, we introduced stochastic shifts from the lattice points (Fig. 2D-D'). The parameter $\sigma_{\max}$ determines the magnitude of this shift. When $\sigma_{\max} = 1$, low cell clusters could be displaced by up to one cell along both the x- and y-axis.

Fig. 2D shows that elongation of low-cell clusters along the Y-axis increased the mean $\overline{\Delta angle_{\mathrm{high}}}$, while elongation of low-cell clusters along the X-axis did not ($l_{\mathrm{h}} = 3$, $l_{\mathrm{v}} = 5$ vs. $l_{\mathrm{h}} = 5$, $l_{\mathrm{v}} = 3$). We obtained similar results under the conditions of various parameter values, $\sigma_{\max}$ (1~3), l_{h}, l_{v}, and $balance_{\mathrm{LH}}$ (Fig. S4A and B). When low cell clusters were regularly aligned, misorientation of polarity vectors was observed only in the immediate vicinity of low cell clusters (Fig. 2C). This is consistent with a previous theoretical study, which report that that core proteins were eliminated from cell edges between Fmi (Flamingo, a fly homolog of CELSRs)-deficient cells and wild type cells (22). In contrast, when clusters were shifted, misorientation of polarity vectors was not limited to the vicinity of low cell clusters (Fig. 2D, middle panel), suggesting that shifting positions of clusters disrupts PCP. We further noted that the influence of low cells on $\overline{\Delta angle_{\mathrm{high}}}$ was particularly evident when $balance_{\mathrm{LH}}$ was below 0.6 (Fig. S4A). This indicates that PCP is affected by low cells even when core proteins are only partially lost in these cells.

These findings support our hypothesis that the direction of low-cell alignment plays a key role in mitigating the disruptive effects of low cells on PCP. However, in our analysis, modifying the lengths of low-cell clusters inevitably altered other aspects of cell-type distributions. For instance, changes in l_{h} and l_{v} also affected the frequency of low cells. To dissect effects of cluster length and other aspects of cell-type distributions on PCP, we constructed a generalized linear model (GLM), a framework for modeling the relationship between an object variable and explanatory variables. Here, the object variable was $\overline{\Delta angle_{\mathrm{high}}}$ and candidates for explanatory variables included $balance_{\mathrm{LH}}$, the frequency of low cells ($freq_{\mathrm{low}}$), $\sigma_{\max}$, l_{h}, and l_{v}. In short, by using GLM, we can expect to evaluate to what extent each explanatory variable contributes to the $\overline{\Delta angle_{\mathrm{high}}}$. Based on the value of AIC (Akaike Information Criterion), we selected the model that best predicted $\overline{\Delta angle_{\mathrm{high}}}$ in our simulation data (Fig. S4C) as follows.

$$\log \overline{\Delta angle_{\mathrm{high}}} =$$

$$-0.951 + 0.529 \cdot freq'_{\mathrm{low}} - 0.237 \cdot l'_{\mathrm{h}} + 0.289 \cdot l'_{\mathrm{v}} - 0.118 \cdot balance'_{\mathrm{LH}} + 0.0418 \cdot \sigma'_{\max}$$

where values of explanatory variables were normalized so that their mean and SD were 0 and 1, respectively. Normalized variables are marked with an apostrophe (e.g., $freq'_{\mathrm{low}}$). The Pearson correlation coefficient between the actual and predicted values of $\overline{\Delta angle_{\mathrm{high}}}$ was 0.822 (Fig. S4C), indicating that the model accurately described the relationship between the object variable and explanatory variables. Here, the absolute value of each coefficients indicates the strength of the influence of explanatory variables on $\overline{\Delta angle_{\mathrm{high}}}$. Among all coefficients, the absolute value of $freq'_{\mathrm{low}}$ is the largest, indicating that frequency of low cells has the most significant effect on $\overline{\Delta angle_{\mathrm{high}}}$. The model shows that l_{v} and l_{h} are also essential parameters that affect PCP, as their

coefficients have the second and third highest absolute values, respectively. As expected, the model shows that an increase in vertical length (l_{v}) elevates $\overline{\Delta angle_{\mathrm{high}}}$, an increase in horizontal length (l_{h}) reduces $\overline{\Delta angle_{\mathrm{high}}}$. In comparison, the absolute value of the coefficient for σ'_{max} is much smaller than those of l'_{h} and l'_{v}. Therefore, the effect of cluster distribution (σ_{max}) on $\overline{\Delta angle_{\mathrm{high}}}$ is less significant than that of cluster lengths (l_{h} and l_{v}). These results were conserved in a different type of cell array (Fig. S2B' and S5). Taken together, we concluded that the directional low-cell alignment is a key aspect of cell-type distributions that mitigates the effects of low cells on PCP.

Exploring features of cell-type distributions that affect PCP

In the previous section, we focused on the effects of direction of low-cell alignment on PCP. However, cell-type distributions can vary widely across different organs and disease states. To gain further insights into which aspects of cell-type distributions influence PCP, we generated cell-type distributions by randomly placing low cells, expecting these distributions to encompass diverse features of cell-type distributions. Through our simulations, we evaluated their effects on PCP and searched for features of cell-type distributions that significantly impact PCP.

We began by generating random distributions with low-cell frequencies ($freq_{\mathrm{low}}$) ranging from 0.05 to 0.95 in increments of 0.05. At $freq_{\mathrm{low}} = 0$, all cells were high cells, whereas at $freq_{\mathrm{low}} = 1$, all cells were low cells. For each value of $freq_{\mathrm{low}}$, we generated 10 random distributions and performed simulations (Fig. 3A-B). Fig. 3A-A" shows examples of our simulations at $balance_{\mathrm{LH}} = 0.5$. $freq_{\mathrm{low}}$ significantly influenced the values of $\overline{\Delta angle_{\mathrm{high}}}$ (Fig. 3B; see the boxed column for these values at $balance_{\mathrm{LH}} = 0.5$), indicating that $freq_{\mathrm{low}}$ is an essential parameter influencing PCP. Additionally, changes in $balance_{\mathrm{LH}}$ significantly altered the value of $\overline{\Delta angle_{\mathrm{high}}}$ (Fig. 3B and S6A-B'; e.g., compare Fig S6A with S6A').

Although $freq_{\mathrm{low}}$ and $balance_{\mathrm{LH}}$ appeared to be essential parameters influencing $\overline{\Delta angle_{\mathrm{high}}}$, they might not be the only determinants of $\overline{\Delta angle_{\mathrm{high}}}$. For example, at $freq_{\mathrm{low}} = 0.4$ and $balance_{\mathrm{LH}}$ = 0.5, the maximum and the minimum values of $\overline{\Delta angle_{\mathrm{high}}}$ were 36.46 and 16.34 degrees, respectively (variance: 40.54 degrees; Fig. S6B). Variances of $\overline{\Delta angle_{\mathrm{high}}}$ at different values of $freq_{\mathrm{low}}$ and $balance_{\mathrm{LH}}$ are summarized in Fig. S6C. We expected that such a variance can be explained by differences in features of cell-type distribution between cell sheets. One candidate is the orientation of the alignment of low and high cells. As highlighted by boxes in Fig. 3A-A', polarity vectors of high cells were strongly reoriented where these cells were aligned along

the Y-axis of the cell sheets. In addition, we noticed that when high cells were dominant in a cell sheet, polarity vectors of high cells were reoriented only in the vicinity of low cells (see a boxed region in Fig. 3A"), suggesting that differences in the frequency of contacts between high and low cells underlie the variations in $\overline{\Delta angle_{\mathrm{high}}}$ between cell sheets. We analyzed effects of these parameters on the values of $\overline{\Delta angle_{\mathrm{high}}}$ later in this section.

To validate whether features of cell-type distribution can explain differences in values of $\overline{\Delta angle_{\mathrm{high}}}$ between cell sheets, we trained the neural network to predict the extent of PCP disorganization in a cell sheet using only information on cell-type distributions (Fig. 3C). If the network was successfully trained, it indicates that the cell-type distribution is a crucial factor in explaining the value of $\overline{\Delta angle_{\mathrm{high}}}$ in a cell sheet. Furthermore, by using an explainable artificial intelligence technique, Grad-CAM (24), we can visualize regions of a cell sheet where the network focused its attention to make predictions (Fig. 3C). We expected that this technique would allow us to identify key features of cell-type distributions that influence $\overline{\Delta angle_{\mathrm{high}}}$.

To this end, we first generated a large number of simulation data with $freq_{\mathrm{low}} = 0.4$ and $balance_{\mathrm{LH}} = 0.5$, as we anticipated that this parameter set would produce a wide range of $\overline{\Delta angle_{\mathrm{high}}}$ values. A total of 6,000 random cell-type distributions were generated, and those in the top 12.5% and bottom 12.5% of $\overline{\Delta angle_{\mathrm{high}}}$ were labeled as "top" and "bottom" distributions, respectively (Fig. 3C). In addition to $\overline{\Delta angle_{\mathrm{high}}}$, we calculated the averaged angle of polarity vectors over low cells ($\overline{\Delta angle_{\mathrm{low}}}$) and all cells ($\overline{\Delta angle_{\mathrm{all}}}$). However, using $\overline{\Delta angle_{\mathrm{low}}}$ or $\overline{\Delta angle_{\mathrm{all}}}$ for labeling did not significantly alter the classification of each distribution as "top" or "bottom", since these values were closely correlated with $\overline{\Delta angle_{\mathrm{high}}}$ (Fig. S6D and E).

Then, a neural network based on the ResNet18 architecture (25) was trained to categorize distributions as either "top" or "bottom" (Fig. 3D). Among the 750 "top" distributions and 750 "bottom" distributions, 100 from each group were used as validation data, while the remaining distributions were used for training. After training, the model achieved a classification accuracy of 0.925 (Fig. 3E), indicating that $\overline{\Delta angle_{\mathrm{high}}}$ is determined by cell-type distributions.

We then used Grad-CAM (23) to analyzed how the network categorized the distributions. Notably, in distributions predicted as "top," Grad-CAM highlighted the edges between high and low cells that were aligned along the Y-axis of the cell sheet (Fig. 3F). In contrast, in distributions predicted as "bottom," Grad-CAM highlighted the edges between high and low cells that were aligned along the X-axis of the cell sheet (Fig. 3F'). To quantify these patterns, we measured the total lengths of horizontal and vertical alignment of low cells that are adjacent to high cells, denoted as L_{h} and L_{v},

respectively (Fig. 3G). The values of L_{V} of "top" distributions were significantly higher than those of "bottom" distributions (p<0.01; Student's *t*-test; Fig. 3H). In contrast, the values of L_{H} of "top" distributions were significantly lower than those of "bottom" distributions (p<0.01; Student's *t*-test; Fig. 3H). As discussed earlier, we expected that the frequency of contacts between high cells and low cells ($freq_{\mathrm{edge}}$) would influence $\overline{\Delta angle_{\mathrm{high}}}$. Although the difference was subtle, we found that the value of $freq_{\mathrm{edge}}$ was higher in "top" distributions than in "bottom" distributions (p<0.05; Student's *t*-test; Fig. 3G'). As discussed later, the contribution of this parameter to $\overline{\Delta angle_{\mathrm{high}}}$ is not negligible.

Finally, to assess the generalizability of the relationships between these parameters and values of $\overline{\Delta angle_{\mathrm{high}}}$, we constructed a GLM using the dataset shown in Fig. 3B. Data with $balance_{\mathrm{LH}}$ >0.6 were excluded from the GLM analysis, as $\overline{\Delta angle_{\mathrm{high}}}$ was weakly affected by cell distributions when the $balance_{\mathrm{LH}}$ was high. The candidate explanatory variables were L_{h}, L_{v}, $balance_{\mathrm{LH}}$, $freq_{\mathrm{low}}$, and $freq_{\mathrm{edge}}$. However, since L_{h} and L_{v} were strongly correlated (i.e., positive correlation) in the dataset, their direct inclusion in the model was problematic. To address this, we converted them into a single variable, $L_{\mathrm{h/v}} = L_{\mathrm{h}}/L_{\mathrm{v}}$. The optimal model was selected based on the value of AIC:

$$\log \overline{\Delta angle_{\mathrm{high}}} = -0.834 + 0.615 \cdot freq'_{\mathrm{low}} - 0.0390 \cdot L'_{\mathrm{h/v}} + 0.327 \cdot freq'_{\mathrm{edge}} - 0.355 \cdot balance'_{\mathrm{LH}}$$

where all explanatory variables are normalized. The values of $\overline{\Delta angle_{\mathrm{high}}}$ predicted by this model were closely correlated with actual values of $\overline{\Delta angle_{\mathrm{high}}}$ (Fig. 3I; Pearson correlation coefficient: 0.859). However, the model's accuracy decreased when the value of $freq_{\mathrm{low}}$ was high, particularly at $freq_{\mathrm{low}}$= 0.95 (yellow dots in Fig. 3I'). The coefficients indicate that $freq_{\mathrm{low}}$ and $balance_{\mathrm{LH}}$ have a significant impact on $\overline{\Delta angle_{\mathrm{high}}}$, as expected. In addition, the model suggests that $freq_{\mathrm{edge}}$ is a key parameter. Although the absolute value of the coefficient for $L'_{\mathrm{h/v}}$ was smaller than those of the other explanatory variables, a model excluding $L_{\mathrm{h/v}}$ was less optimal according to AIC. Since the coefficient of $L_{\mathrm{h/v}}$ is negative, the model suggests that an increase in the ratio of total lengths of horizontal to vertical alignment of low cells decreases $\overline{\Delta angle_{\mathrm{high}}}$. These results indicate that the frequency of low cells ($freq_{\mathrm{low}}$), the frequency of contacts between high and low cells ($freq_{edge}$), and the directional arrangement of cell types ($L_{\mathrm{h/v}}$) are key aspects of cell-type distributions that influence PCP.

Roles of low cells in the maintenance of PCP

Given that low cells potentially disrupt PCP, what is the biological significance of the presence of low cells in relation to PCP? One intriguing hypothesis is that a cluster of low cells provide a directional cue that contribute to PCP maintenance. To test this idea, we disrupted the distribution of core proteins in cells located in a right half of the cell sheet, and examined whether low-cell clusters facilitate PCP reconstruction. In certain conditions, low-cell clusters enhanced reconstruction of PCP. For example, when polarity vectors were oriented 60 degrees relative to the X-axis, polarities were not strongly reorganized when all cells were high cells (Fig. S7A and A'). This may be due to the periodic edge conditions of the given cell sheet, which caused the reorganization of polarity vectors to become trapped in a local minimum. In contrast, when low-cell clusters were elongated along the X-axis of the cell sheet, polarity vectors were reoriented along the direction of cluster elongation (Fig. S7B and B'). Notably, in regions where high cells were flanked between the long sides of two elongated clusters, polarity vectors aligned with the cluster orientation, suggesting that low-cell alignment provides a local directional cue (Fig. S7B', a boxed region). However, when the initial cell-edge distributions of the FZD and VANGL complexes were randomized in the right half of the cell sheet, the presence of low-cell clusters did not improve PCP reconstruction (Fig. S7A, A", B and B"). In the absence of low-cell clusters, polarity vectors became unidirectionally aligned by the end of the simulation. Conversely, when low-cell clusters were present, misoriented polarity vectors persisted (Fig. S7B", a boxed region). These results suggest that low-cell clusters provide a local directional cue, but it does not improve PCP maintenance robustly.

Core protein mislocalization in regions lacking directional alignment of secretory cells

Theoretical analyses predicted that the orientation of polarity vectors in high cells is correlated with the direction of the alignment of adjacent low cells. To test whether this relationship exists *in vivo*, we examined a region of the oviduct, the ampulla, where MCCs are sparsely distributed (Fig. 4A-A"). We specifically searched for ampulla regions where SCCs were circumferentially aligned and analyzed the distribution of VANGL2:EGFP in MCCs adjacent to these SCCs (three examples are shown in Fig. 4B-D"). As expected, VANGL2:EGFP localization at cell edges was biased along the direction of SCC alignment in neighboring MCCs (arrowheads in Fig. 4B-D"). In regions where SCCs were aligned along the ovary-uterus axis, VANGL2:EGFP was enriched at cell edges oriented circumferentially in adjacent MCCs (yellow arrowheads in Fig. 4B-B" and 4-D-D"). In contrast, in regions where SCCs aligned circumferentially, VANGL2:EGFP was enriched at cell edges running along the ovary-uterus axis in adjacent MCCs (white arrowheads in Fig. 4B-D).

These observations support our hypothesis that the direction of low-cell alignment influences the PCP of the tissue.

CELSR1 is required for aligning low cells in the direction of PCP

Our analyses revealed that the orientation of low-cell alignment is a key feature of cell-type distributions that mitigate the effect of low cells on PCP. In the oviduct, orientation of low-cell alignment is consistent with the orientation of PCP. What regulates the orientation? We hypothesized that core proteins direct the low-cell alignment.

To test this hypothesis, we analyzed lengths of SCC clusters in oviducts lacking the core protein, CELSR1 (Fig. 5A-B'). For this analysis, images were adopted from (9). In *Celsr1* mutants, l_{h} and l_{v} values were comparable (Fig. 5C-D, gray plots in 5D); the mean l_{h} and l_{v} values of SCC clusters were 2.49 cells (SD: 1.15) and 2.55 cells (SD: 0.97), respectively. While the mean l_{v} value increased in *Celsr1* mutant oviducts, the mean l_{h} value remained unchanged (Fig. 5D, compare white and gray plots). In contrast, neither the frequency of MCCs (Fig. 5E) nor the frequency of edges between SCCs and MCCs (Fig. 5F) was affected in *Celsr1* mutants. These results indicate that CELSR1 is required to align low cells in the direction of PCP.

Since CELSR1 levels were low in SCCs, defects of SCC alignment in *Celsr1* mutant oviduct may result from an absence of CELSR1 in surrounding MCCs. This suggests that low cells were non-autonomously aligned by surrounding high cells. Consistent with this idea, clusters of *Celsr1* mutant cells were elongated along the ovary-uterus axis in wild-type oviduct epithelia (Fig. S8).

The theoretical analyses in previous sections demonstrated that low-cell alignment reorganizes the distribution of core proteins in adjacent high cells. Taken together, our results suggest that a positive feedback mechanism between core proteins and cell-type distributions contributes to PCP maintenance.

Discussion

The establishment of PCP requires interplays of core proteins between adjacent cells. Previous studies showed that experimentally induced low cells disrupt polarity of surrounding wild-type cells (18). Nevertheless, several tissues naturally exhibit imbalances in core protein levels between cells, while still maintaining PCP. In this study, we demonstrated that cell-type distribution determines the extent to which low cells affect PCP. Specifically, the disruptive effects of low cells on PCP

were mitigated when low and high cells were aligned in the direction of PCP. We believe that our study provides important perspectives on the biological significance of cell-type distributions that contribute to normal development and homeostasis by stabilizing PCP.

The oviduct epithelium consists of two cell types: SCCs and MCCs, corresponding to low and high cells, respectively. Small clusters of SCCs are scattered throughout the epithelium, and these clusters are elongated along the direction of PCP. Our theoretical analyses predicted that this directional alignment of low cells helps counteract the effects of core protein imbalance on PCP. In contrast, forming small clusters of low cells increases the number of edges between low and high cells, potentially compromising PCP maintenance. In other words, the distribution of SCCs in the oviduct is not optimized for maintaining PCP. The function of SCCs extends beyond providing essential components for oocyte survival and growth. They have also been proposed to act as stem cells in the oviduct epithelium (26). Thus, their dispersed distribution may be necessary to ensure uniform secretion activity and tissue renewal. The cell-type distribution of the oviduct epithelium may not be ideal for PCP but it appears to be well-suited for supporting multiple functions of the oviduct epithelium.

An open question remains: what is the biological significance of the absence of core proteins in specific cell types? We tested the hypothesis that low-cell clusters contribute to PCP maintenance by providing a local directional cue. Near the boundaries of low-cell clusters, polarity vectors of high cells aligned parallel to the cluster boundaries, suggesting that low-cell clusters may serve as a local directional guide. In certain conditions, the presence of low-cell clusters rescued PCP defects; however, the efficiency of this rescue depended on the initial conditions. Therefore, we could not conclusively determine whether the presence of low cells improves PCP maintenance. It is also possible that the absence of core proteins is essential for the specific functions of low cells, but this remains to be investigated.

We found that CELSR1 is required for aligning low cells in the direction of PCP. Our results suggest that low cells are aligned non-autonomously by CELSR1 in surrounding high cells, although the underlying mechanisms remain unclear. The well-established view of roles of core proteins in PCP establishment is to coord the distribution of core proteins between adjacent cells. In addition to this function, our study suggests that core proteins may also facilitate the coordination of cell polarity by regulating cell-type distribution, particularly in tissues where core protein imbalance is present. This implies that a positive feedback mechanism acts in the maintenance of PCP: core proteins align low cells, and the alignment of low cells refines the distribution of core PCP proteins in adjacent high cells. Unlike the oviduct, cilia beating orientation exhibits modular organization in the ventricle, where beating orientations are coordinated within each subdivision of the ventricle, but

not between subdivisions (27). It would be interesting to investigate whether the positive feedback regulation between core proteins and cell-type distributions contribute to the generation of divergent cilia beating directions in the ventricle.

In this study, we focused on a specific type of low cells in which the levels of the FZD complex and the VANGL complex are equal. However, different values of $balance_{\mathrm{LH}}$ can be assigned to each complex. A comprehensive understanding of the effects of cell-type distribution on PCP would require considering all possible combinations of FZD and VANGL complex levels in low cells, as well as the coexistence of different low-cell types within a tissue. However, such an approach would necessitate an extensive number of simulations and analyses. Therefore, we chose to focus on a scenario that closely aligns with the oviduct, though this may limit the generalizability of our conclusions. If the low cells in a given tissue differ from those analyzed in our study, our findings may not fully apply. Nevertheless, our approach—including mathematical modeling and GLM analysis—can be readily adapted to examine the effects of other types of low cells on PCP.

Although our study focused on cell-type distribution, we do not exclude the possibility that other mechanisms also mitigate the effects of core protein imbalance on PCP. For example, in the mouse airway, the frequency of MCCs (0.37 on average; (28)) is much lower than in the oviduct (0.76 on average; (9)). If non-ciliated cells in the airway are low cells, our model suggests that PCP maintenance would not be feasible in the airway epithelium. However, several studies have reported that fluid flow can align cilia beating orientation in the ventricle (29), the airway (30), and the *Xenopus* skin (31). These mechanisms might also play a role in compensating for core protein imbalance.

Materials and Methods

Simulation

We used Octave (32) for numerical simulations and visualized the data by using Fiji (33). The epithelial tissue was modeled as an array of regular hexagonal cells, with a width and height of 24 cells each. The boundary conditions were periodic in all directions. See Supporting Information for details on our mathematical modeling and simulation.

Imaging fluorescent signals

Oviducts were dissected and opened longitudinally along the ovary uterus axis. For imaging VANGL2:EGFP signals, oviducts were fixed in 4% PFA in PBS (-) at 4℃ overnight. After washing, samples were incubated with Blocking One (Nacalai Tesque, 03953-95) at room temperature for 30 min. The samples were then incubated with anti-acetylated-TUBULIN antibody (1/500, Sigma-Aldrich, T7451) in Blocking One (Nacalai Tesque, 03953-95) at 4℃ overnight. After washing, samples were incubated with phalloidin (1/500, Thermo Fisher Scientific, A12381) and anti-Mouse IgG antibody (1/500, Thermo Fisher Scientific, A21235) in Blocking One (Nacalai Tesque, 03953-95) at 4℃ overnight. Finally, after washing, samples were mounted in Fluoromount G (SouthernBiotech, 0100-01). All washing steps were performed as follows: 3 x rinse and 3 x 10 min washes with 0.1% Triton X-100 in PBS (-) at room temperature. Images were acquired using Nikon ECLIPSE Ti2 microscope equipped with CSUW1 spinning disk confocal unit (Yokogawa).

Quantification of low-cell distributions *in vivo*

In Fig. 1C-D, data adopted from (11) were reanalyzed. In Fig. 5, data adopted from (9) were reanalyzed. Cluster lengths were measured as follows: First, cells were segmented using Cellpose2 (34), and the segmented data were saved as ImageJ ROIs. ROIs marking SCCs were selected based on signals of MCC markers (γ-TUBULIN or acetylated-TUBULIN). These ROIs were then dilated and filled with white, resulting in fusions of areas of adjacent SCCs. The images were then rotated so that the ovary-uterus axis aligned with the X-axis of the image. In the rotated images, white areas were converted to ROIs using the analyze particle function in Fiji. Using these ROIs, the O-U and circumferential lengths of each cluster were measured by the measurement function in Fiji. Clusters that crossed the image edges or epithelial edges were excluded from our analysis. Since oviduct epithelial cells elongate along the ovary-uterus axis (9), we normalized the O-U and circumferential lengths of each SCC cluster. To achieve this, we divided cluster lengths by the average length of SCCs that comprise each cluster. These values were plotted in Fig. 1D-D' and Fig. 5C-D. In Fig. S8, lengths of clusters composed of wild-type or *Celsr1* mutant cells were quantified using the same method, except that nuclear H2B:EGFP signals were used as a marker for these cells.

In Fig. 5, frequencies of MCCs were measured by calculating the ratio of MCCs to all cells in each image. To measure the frequency of edges between SCCs and MCCs, we counted the number of these edges and divided the value by the number of all cell edges in each image.

Mouse

Female mice (*R26-VANGL2-EGFP* transgenic mice, (10)) were used to study VANGL2:EGFP distributions in wild-type oviducts. Details of *R26-VANGL2-EGFP* transgenic mice (10) were described in these papers. Animal care and experiments were conducted following the Guidelines of Animal Experimentation of National Institutes for Natural Sciences. All animal experiments were approved by either the Animal Research Committee of National Institutes for Natural Sciences. Animals were maintained in a light- and temperature-controlled room using a 12 h:12 h light:dark cycle at 23±2°C.

Generalized linear model

All following procedures were performed in R. Generalized linear models (GLM) were fit to data by using *glm* function of *stats* package (version 4.4.1). Explanatory variables were standardized before fitting GLM by using *scale* function of *base* package (version 4.4.1), which enabled comparisons of coefficients of explanatory variables. Gamma distribution was selected as a probability distribution, because $\overline{\Delta angle_{\mathrm{high}}}$ takes positive continuous values. The object variable was related to explanatory variables via a log link function. To avoid the multicollinearity, we calculated variance inflation factors (VIF) by using *vif* function of *car* package (version 3.1-2). We selected explanatory variables so that the value of VIF does not to exceed 10. Then, models were evaluated based on values of AIC (Akaike information criterion) by using *dredge* function of *MuMIN* package (version 1.48.4). The model that took the lowest value of AIC was selected and used in the manuscript. We used *r2* function of package *performance* version 0.12.4 and *cor* function of package *stats* version 4.4.1 for calculating these values for calculating Pearson correlation coefficient and Nagelkerke's R2, respectively.

Exploring features of cell-type distribution that affect PCP by deep learning

1. Data generation and labeling

For generating random distribution of low cells in **Fig. 3**, random numbers that range from 0 to 1 were assigned to each cell, and cells with top 100*$freq_{\mathrm{low}}$ % values were defined as low cells. Threshold values of $\overline{\Delta angle_{\mathrm{high}}}$ for labeling “top” or “bottom” were determined by the accuracy of prediction made by the trained network. More specifically, at first, 3,000 distributions were generated at $freq_{\mathrm{low}} = 0.4$ and $balance_{\mathrm{LH}} = 0.5$ and data that took top 25% and bottom 25% values of $\overline{\Delta angle_{\mathrm{high}}}$ were labeled as “top” and “bottom”, respectively. We used these data for training the network, but the accuracy of the prediction made by the network did not exceed 0.85.

We additionally generated 3,000 distributions and selected distributions that took top 12.5% and bottom 12.5% values of $\overline{\Delta angle_{\mathrm{high}}}$ for deep learning.

Data augmentation is a technique to increase the number and diversity of training data. Since our simulation adopted the periodic edge condition in all directions, shifting a cell-type distribution along the X and Y axis did not affect the value of $\overline{\Delta angle_{\mathrm{high}}}$. Taking advantage of this, we generated three cell-type distributions from one distribution: one shifted 12 cells along the X-axis, one shifted 12 cells along the Y-axis, and one shifted 12 cells along both axes.

2. Training a neural network

We used the ResNet18 architecture (25) that was pre-trained with the ImageNet-1K dataset for training. We re-trained the network (i.e., transfer learning) by using Google colaboratory. Hyperparameters used for training were as follows: the optimizer was SGD with a learning rate of 0.01 and a momentum of 0.9. batch size, 50; epoch number, 10. Since ResNet18 accepts images as input, each cell-type distribution was converted to an RGB image with 224 pixels in height and 224 pixels in width each. Since each cell was drawn as a hexagon, the converted image had margins at edges. The red channel was assigned to high cells, the green to low cells, and the blue to the margin. Then intensities of images were normalized so that the mean and SDs of intensities of each channel matched with these of the ImageNet-1K dataset. The accuracy of the trained network was evaluated by counting the number of validation data whose labels were correctly predicted by the network.

3. Grad-CAM visualization of the neural network attention

Grad-CAM is an explainable AI technique that can visualize which regions of images the network focuses on when making a classification of an input image (24). We used pytorch-gradcam package in Python to visualize the network attention. Layer2[-1] of PyTorch ResNet18 were selected as a target layer to compute the attention map.

Acknowledgments

We thank Dr. Goshi Ogita for discussion; members of the T.F laboratory for discussion and technical assistance. This work was supported by JSPS KAKENHI Grant Numbers, 21H02494, 24K02039, 22H05168 (to T.F.), 22K15130 and the Takeda Science Foundation (to M.A.).

Declaration of interests

The authors declare no competing interests.

Author Contributions

M.A., H.K. and T. F. designed research; M.A. performed research; M.A. analyzed data; and M.A., H.K. and T. F. wrote the paper.

Use of generative AI and AI-assisted technologies

Gemini (Google) was used for writing R, Fiji, and Python scripts. ChatGPT (Open AI) was used for improving language and readability of the manuscript. After using these tools, we reviewed and edited the content of the scripts and the manuscript.

References

1. M. T. Butler, J. B. Wallingford, Planar cell polarity in development and disease. *Nat Rev Mol Cell Biol* **18**, 375–388 (2017).
2. J. D. Axelrod, Planar cell polarity signaling in the development of left–right asymmetry. *Curr Opin Cell Biol* **62**, 61–69 (2020).
3. P. N. Adler, The frizzled/stan Pathway and Planar Cell Polarity in the Drosophila Wing. *Curr Top Dev Biol* **101**, 1–31 (2012).
4. Y. Koca, G. M. Collu, M. Mlodzik, Wnt-frizzled planar cell polarity signaling in the regulation of cell motility. *Curr Top Dev Biol* **150**, 255–297 (2022).
5. W. Y. Aw, D. Devenport, Planar cell polarity: global inputs establishing cellular asymmetry. *Curr Opin Cell Biol* **44**, 110–116 (2017).
6. H. Strutt, D. Strutt, How do the Fat-Dachsous and core planar polarity pathways act together and independently to coordinate polarized cell behaviours? *Open Biol* **11**, 200356 (2021).
7. P. A. Lawrence, J. Casal, Planar cell polarity: Two genetic systems use one mechanism to read gradients. *Development (Cambridge)* **145**, dev168229 (2018).
8. H. Koyama, D. Shi, T. Fujimori, Biophysics in oviduct: Planar cell polarity, cilia, epithelial fold and tube morphogenesis, egg dynamics. *Biophys Physicobiol* **16**, 89–107 (2019).
9. D. Shi, *et al.*, Celsr1 is required for the generation of polarity at multiple levels of the mouse oviduct. *Development (Cambridge)* **141**, 4558–4568 (2014).
10. D. Shi, *et al.*, Mechanisms of Development Dynamics of planar cell polarity protein Vangl2 in the mouse oviduct epithelium. *MOD* **141**, 78–89 (2016).
11. F. M. Usami, *et al.*, Intercellular and intracellular cilia orientation is coordinated by CELSR1 and CAMSAP3 in oviduct multi-ciliated cells. *J Cell Sci* **134**, jcs257006 (2021).
12. M. Arata, F. M. Usami, T. Fujimori, Coordination of Cilia Movements in Multi-Ciliated Cells. *J Dev Biol* **10**, 47 (2022).

13. C. O. Copley, J. S. Duncan, C. Liu, H. Cheng, M. R. Deans, Postnatal refinement of auditory hair cell planar polarity deficits occurs in the absence of Vangl2. *Journal of Neuroscience* **33**, 14001–14016 (2013).
14. K. Amonlirdviman, *et al.*, Mathematical modeling of planar cell polarity to understand domineering nonautonomy. *Science (1979)* **307**, 423–426 (2005).
15. K. Abley, *et al.*, An intracellular partitioning-based framework for tissue cell polarity in plants and animals. *Development (Cambridge)* **140**, 2061–2074 (2013).
16. B. Cho, G. Pierre-louis, A. Sagner, S. Eaton, Clustering and Negative Feedback by Endocytosis in Planar Cell Polarity Signaling Is Modulated by Ubiquitinylation of Prickle. *PLoS Genet*. **11**, e1005259 (2015).
17. S. J. Warrington, H. Strutt, K. H. Fisher, D. Strutt, A Dual Function for Prickle in Regulating Frizzled Stability during Feedback-Dependent Amplification of Planar Polarity. *Current Biology* **27**, 2784-2797 (2017).
18. W. Chen, *et al.*, Asymmetric Homotypic Interactions of the Atypical Cadherin Flamingo Mediate Intercellular Polarity Signaling. *Cell* **133**, 1093–1105 (2008).
19. J. S. Duncan, *et al.*, Celsr1 coordinates the planar polarity of vestibular hair cells during inner ear development. *Dev Biol* **423**, 126–137 (2017).
20. E. K. Vladar, R. D. Bayly, A. M. Sangoram, M. P. Scott, J. D. Axelrod, Microtubules enable the planar cell polarity of airway cilia. *Current Biology* **22**, 2203–2212 (2012).
21. G. Das, J. Reynolds-Kenneally, M. Mlodzik, The Atypical Cadherin Flamingo Links Frizzled and Notch Signaling in Planar Polarity Establishment in the Drosophila Eye. *Dev Cell* **2**, 655–666 (2002).
22. K. H. Fisher, D. Strutt, A. G. Fletcher, Experimental and Theoretical Evidence for Bidirectional Signaling via Core Planar Polarity Protein Complexes in Drosophila. *iScience* 17, 49–66 (2019).
23. S. Shadkhoo, M. Mani, The role of intracellular interactions in the collective polarization of tissues and its interplay with cellular geometry. *PLoS Comput Biol* **15**, e1007454 (2019).
24. R. R. Selvaraju, M. Cogswell, A. Das, R. Vedantam, D. Parikh and D. Batra, "Grad-CAM: Visual Explanations from Deep Networks via Gradient-Based Localization," *2017 IEEE International Conference on Computer Vision (ICCV)*, Venice, Italy, 2017, pp. 618-626
25. K. He, X. Zhang, S. Ren and J. Sun, "Deep Residual Learning for Image Recognition," *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Las Vegas, NV, USA, 2016, pp. 770-778
26. A. Ghosh, S. M. Syed, P. S. Tanwar, In vivo genetic cell lineage tracing reveals that oviductal secretory cells self-renew and give rise to ciliated cells. *Development (Cambridge)* **144**, 3031–3041 (2017).

27. R. Faubel, C. Westendorf, E. Bodenschatz, G. Eichele, Cilia-based flow network in the brain ventricles. *Science (1979)* **353**, 176–178 (2016).
28. G. R. Ramirez-San Juan, *et al.*, Multi-scale spatial heterogeneity enhances particle clearance in airway ciliary arrays. *Nat Phys* **16**, 958–964 (2020).
29. B. Guirao, *et al.*, Coupling between hydrodynamic forces and planar cell polarity orients mammalian motile cilia. *Nat Cell Biol* **12**, 341–350 (2010).
30. N. Sone, *et al.*, Multicellular modeling of ciliopathy by combining ips cells and microfluidic airway-on-a-chip technology. *Sci Transl Med* **13**, eabb1298 (2021).
31. B. Mitchell, R. Jacobs, J. Li, S. Chien, C. Kintner, A positive feedback mechanism governs the polarity and motion of motile cilia. *Nature* **447**, 97–101 (2007).
32. J. W. Eaton, D. Bateman, S. Hauberg, R. Wehbring, GNU Octave version 8.2.0 manual: a high-level interactive language for numerical computations. https://www.gnu.org/software/octave/doc/v8.2.0/ (2023).
33. J. Schindelin, *et al.*, Fiji: An open-source platform for biological-image analysis. *Nat Methods* **9**, 676-82 (2012).
34. M. Pachitariu, C. Stringer, Cellpose 2.0: how to train your own model. *Nat Methods* **19**, 1634–1641 (2022).

Figures

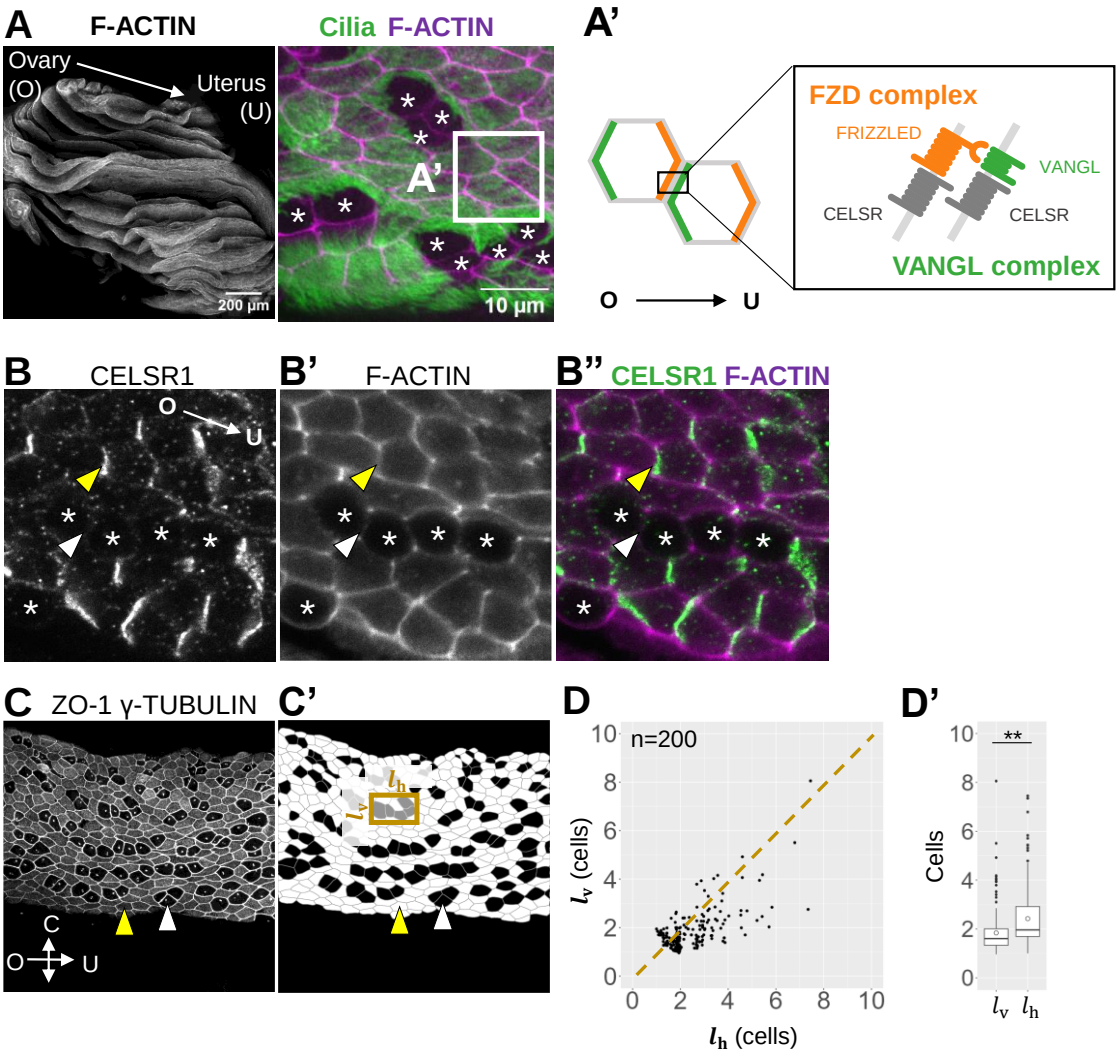

Figure 1. Distributions of SCCs in the oviduct epithelium

(**A**) The infundibulum of the oviduct was dissected, longitudinally opened along the ovary (O)-uterus (U) axis, and stained for F-actin (gray in the left panel and magenta in the right panel, a marker for cell edges) and acetylated-TUBULIN (green in the right panel, a marker for cilia). The oviduct lumen is lined with an epithelium composed of MCCs and SCCs (asterisks indicate SCCs in the right panel). Apical views of the epithelia are shown. (**A'**) A diagram illustrating the asymmetric distribution of the FZD and VANGL complexes on cell edges of the oviduct MCCs. (**B-B''**) Oviduct epithelium stained for CELSR1 (**B**, green in **B''**) and F-actin (**B'**, magenta in **B''**). Asterisks indicate SCCs. CELSR1 is enriched at cell edges perpendicular to the ovary-uterus axis. CELSR1 levels are higher at cell edges between MCCs (yellow arrowhead) than at those between SCCs (white

arrowhead). (**C-D'**) Analysis of SCC distribution in the oviduct epithelium of 11-week-old mice. (**C**) Oviduct stained for ZO1 (a marker for cell edges) and γ-TUBULIN (a marker for MCCs). White and yellow arrowheads point to a SCC and MCC, respectively. (**C'**) Gray lines: cell edges; white cells: MCCs; black cells: SCCs. The ovary-uterus (O-U) and circumferential (C) lengths of each SCC cluster (l_{h} and l_{v}, respectively) were measured using a bounding rectangle (gold box). (**D-D'**) l_{h} and l_{v} values were analyzed using a scatter plot (**D**) and a boxplot (**D'**). The dotted line in (**D**) represents $l_{\mathrm{h}} = l_{\mathrm{v}}$. Open circles in **D'** indicate average values. **: p<0.01 (Wilcoxon signed-rank test). Isolated SCCs were excluded from the analysis. N: 200 clusters; 3 mice.

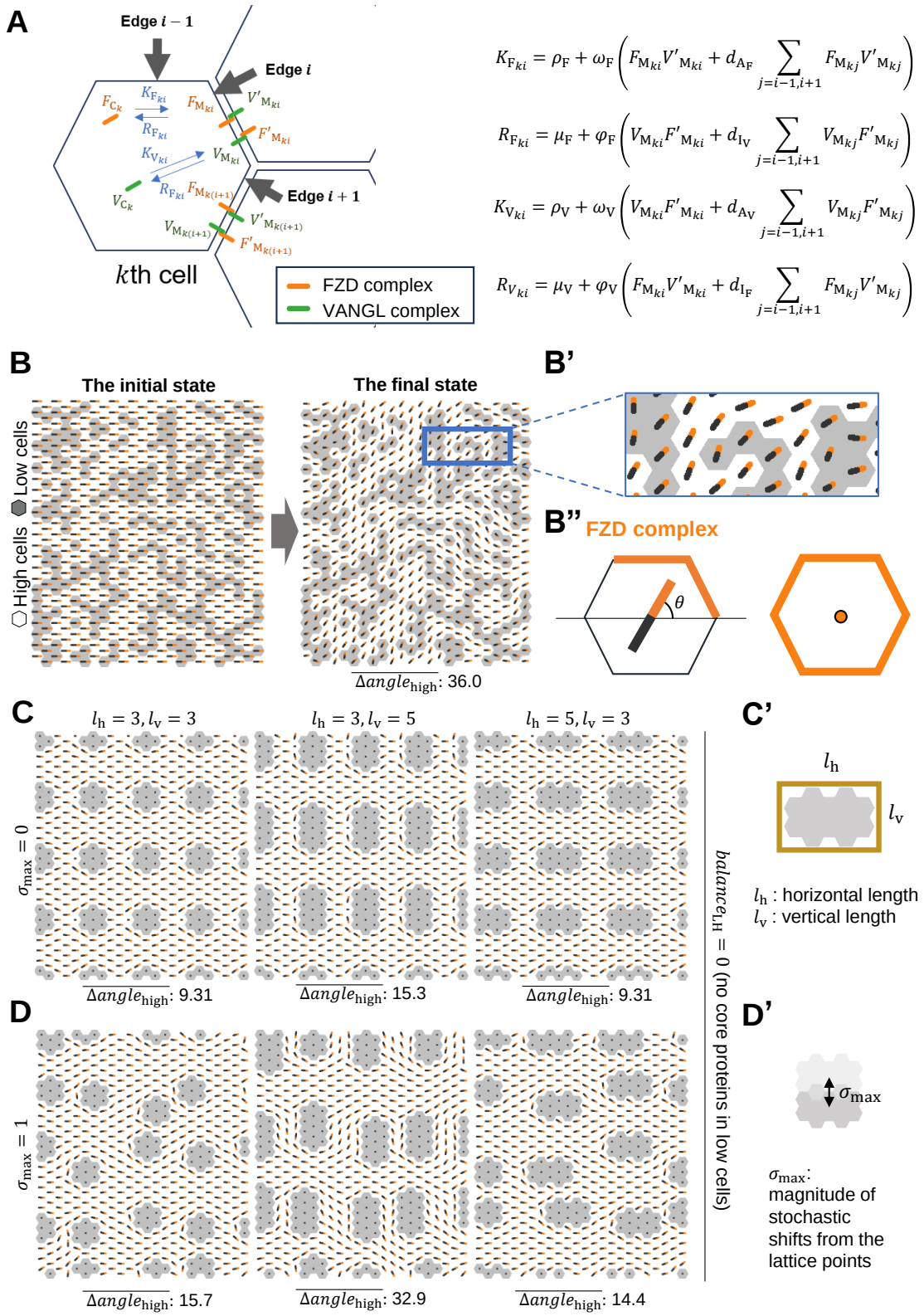
A
Edge $i-1$
Edge i
Edge $i+1$
kth cell
FZD complex
VANGL complex
$K_{F_{ki}} = \rho_F + \omega_F \left(F_{M_{ki}} V'_{M_{ki}} + d_{A_F} \sum_{j=i-1,i+1} F_{M_{kj}} V'_{M_{kj}} \right)$
$R_{F_{ki}} = \mu_F + \varphi_F \left(V_{M_{ki}} F'_{M_{ki}} + d_{I_V} \sum_{j=i-1,i+1} V_{M_{kj}} F'_{M_{kj}} \right)$
$K_{V_{ki}} = \rho_V + \omega_V \left(V_{M_{ki}} F'_{M_{ki}} + d_{A_V} \sum_{j=i-1,i+1} V_{M_{kj}} F'_{M_{kj}} \right)$
$R_{V_{ki}} = \mu_V + \varphi_V \left(F_{M_{ki}} V'_{M_{ki}} + d_{I_F} \sum_{j=i-1,i+1} F_{M_{kj}} V'_{M_{kj}} \right)$
B
The initial state
The final state
Low cells
High cells
$\overline{\Delta angle_{high}}$: 36.0
B'
B''
FZD complex
θ
C
$l_h = 3, l_v = 3$
$l_h = 3, l_v = 5$
$l_h = 5, l_v = 3$
$\sigma_{max} = 0$
$\overline{\Delta angle_{high}}$: 9.31
$\overline{\Delta angle_{high}}$: 15.3
$\overline{\Delta angle_{high}}$: 9.31
D
$\sigma_{max} = 1$
$\overline{\Delta angle_{high}}$: 15.7
$\overline{\Delta angle_{high}}$: 32.9
$\overline{\Delta angle_{high}}$: 14.4
$balance_{LH} = 0$ (no core proteins in low cells)
C'
l_h
l_v
l_h : horizontal length
l_v : vertical length
D'
σ_{max}
σ_{max}: magnitude of stochastic shifts from the lattice points

Figure 2. Evaluating the effects of directions of low-cell alignment on PCP

(**A**) Diagram illustrating an overview of the mathematical model. See the main text and Supporting information for details. (**B-B”**) Example of the analysis of the effects of cell-type distribution on PCP ($balance_{\mathrm{LH}} = 0.5, freq_{\mathrm{low}} = 0.4$). White hexagons: high cells; gray hexagons; low cells. (**B**) The unidirectionally biased distribution of the FZD and VANGL complexes was reorganized due to the core protein imbalance (compare the initial and final states). A boxed region in (**B**) is enlarged and shown in (**B’**). (**B”**) A bar centered in each cell represents polarity of cell-edge distribution of the FZD complexes. The bar’s length and angle indicate magnitude and orientation of complex enrichment, respectively. (**C-D**) Evaluation of the effects of low-cell cluster lengths (C’) on PCP. The resulting $\overline{\Delta angle_{\mathrm{high}}}$ values are shown at the bottom of each panel. Clusters were arranged in a lattice (**C**) and were randomly shifted up to $\sigma_{\max}$ cell(s) along the X and Y axis of each cell sheet (**D-D’**).

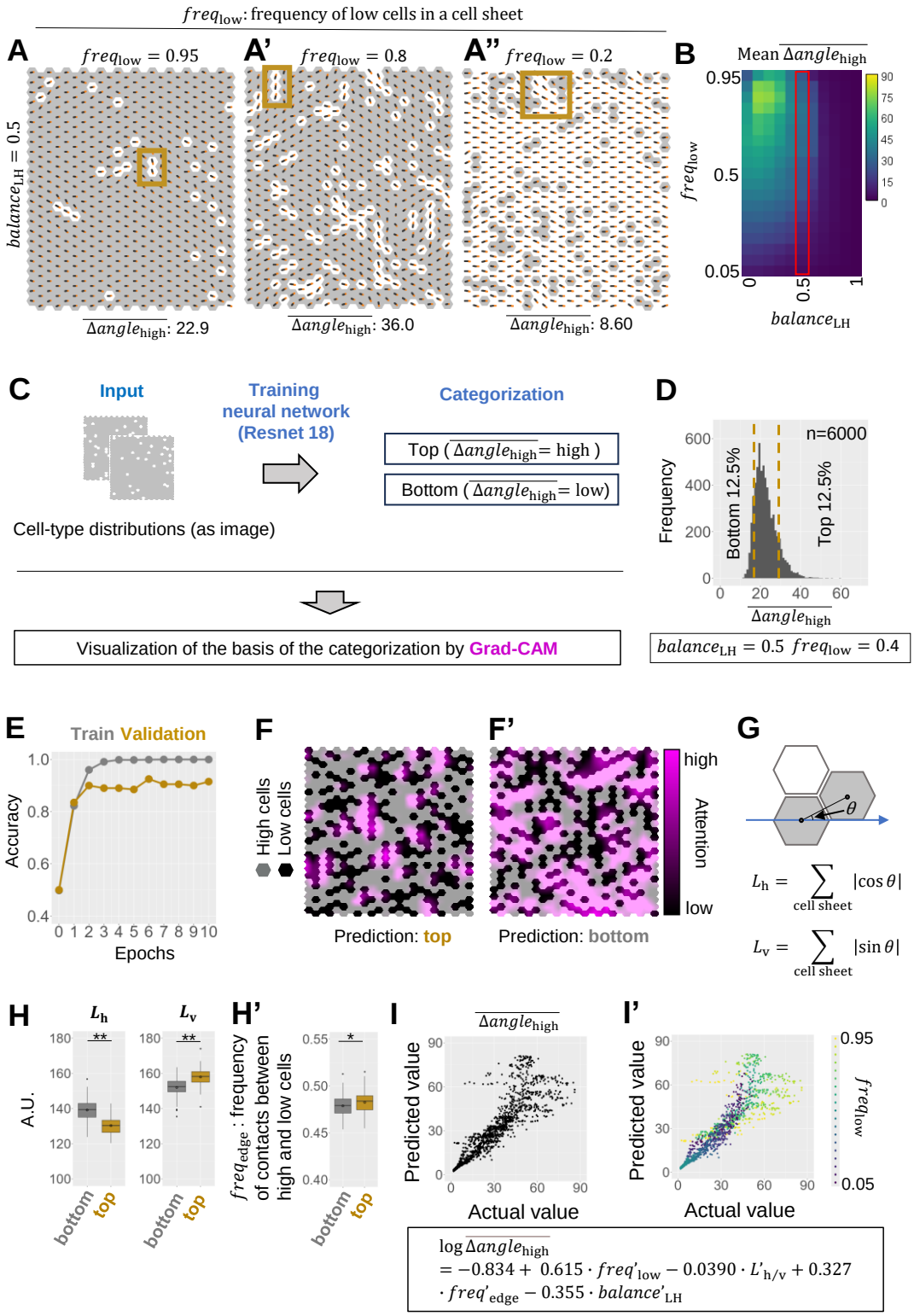

Figure 3. Exploring features of cell-type distribution affecting PCP

(**A-A”**) Examples showing the effects of random low-cell distributions on PCP. $balance_{\mathrm{LH}} = 0.5$. $freq_{\mathrm{low}} = 0.95$ (**A**), 0.8 (**A’**), and 0.95 (**A”**). The resulting values of $\overline{\Delta angle_{\mathrm{high}}}$ are shown at the bottom of each panel. Boxes in **A** and **A’** highlight alignment of high cells along the Y-axis of the cell sheet. A box in **A”** highlight the misoriented polarity vectors of high cells in the vicinity of low cells. (**B**) For each $freq_{\mathrm{low}}$ value, 10 random distributions of low cells were generated and their effects on PCP were evaluated by simulation. Heatmap colors represent the average $\overline{\Delta angle_{\mathrm{high}}}$ values over 10 random distributions for each combination of $balance_{\mathrm{LH}}$ and $freq_{\mathrm{low}}$ (unit: degrees). The boxed column represents data at $balance_{\mathrm{LH}} = 0.5$. (**C-H**) Exploring features of cell-type distribution affecting PCP using deep learning. (**C**) Diagram illustrating the strategy for deep learning analysis. Cell-type distributions were provided as images for the input into the neural network. No information about polarity vectors was included in the input images. The neural network (ResNet18) was trained to predict whether a given cell-type distribution corresponds to a “top” ($\overline{\Delta angle_{\mathrm{high}}}$ = high) or “bottom” ($\overline{\Delta angle_{\mathrm{high}}}$ = low) classification. Finally, Grad-CAM was applied to visualize the regions in the input images that contributed to the prediction. (**D**) A total of 6,000 random distributions were generated at $freq_{\mathrm{low}} = 0.4$, and their effects on PCP were simulated at $balance_{\mathrm{LH}} = 0.5$. The variation in the resulting $\overline{\Delta angle_{\mathrm{high}}}$ values is shown as a histogram. Distributions within the top 12.5% and bottom 12.5% $\overline{\Delta angle_{\mathrm{high}}}$ values were labeled “top” and “bottom”, respectively. (**E**) Prediction accuracy of the neural network during training. The gray and gold lines show changes in prediction accuracy for the training and validation data, respectively. (**F-F’**) Grad-CAM attention maps highlighting regions of input distributions where the network focused during predictions. Black and gray hexagons represent low and high cells, respectively. Magenta brightness indicates attention magnitude. (**F**) A distribution predicted as “top”. Attention was concentrated at vertical edges between low cells and high cells. (**F’**) A distribution predicted as “bottom”. Attention was concentrated at horizontal edges between low and high cells. (**G**) Diagram illustrating the definition of L_{h} and L_{V}. Gray and white cells represent low and high cells, respectively. Blue arrow indicates the X-axis of the cell sheet. (**H**) Boxplots showing L_{h} and L_{V} values for top and bottom validation data (**: p<0.01; Wilcoxon rank-sum test). (**H’**) Boxplots showing frequencies of edges between low cells and high cells for the top and bottom validation data (*: p<0.05; Wilcoxon rank-sum test). (**I-I’**) Scatter plots showing the correlation between actual $\overline{\Delta angle_{\mathrm{high}}}$ values and those predicted by the GLM. Each dot represents a data point shown in (**B**). $freq_{\mathrm{low}}$ values are color-coded in (**I’**).

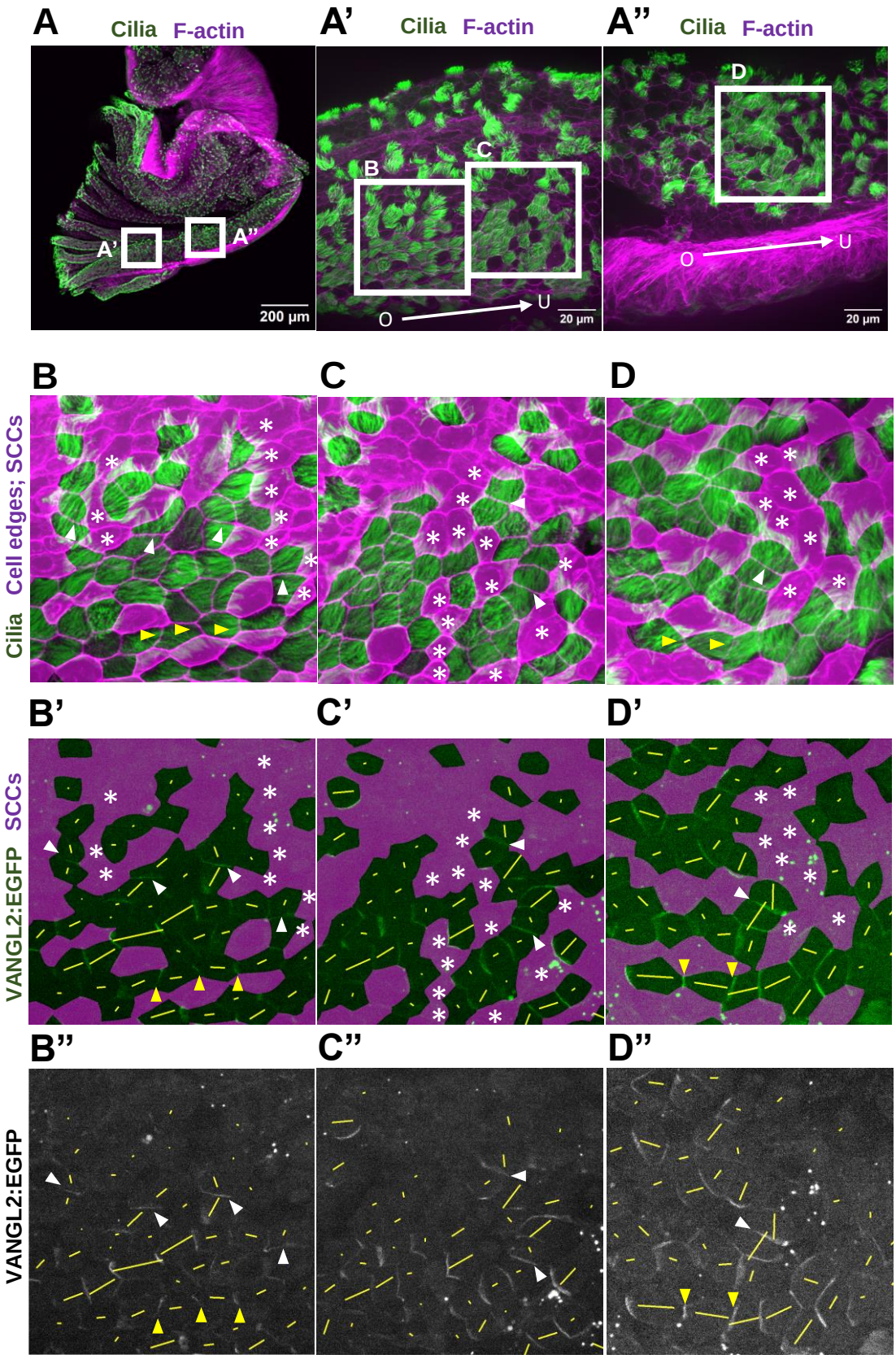
A
Cilia F-actin
A'
A"
200 µm
A'
Cilia F-actin
B
C
O
U
20 µm
A"
Cilia F-actin
D
O
U
20 µm
B
C
D
Cilia Cell edges; SCCs
B'
C'
D'
VANGL2:EGFP SCCs
B"
C"
D"
VANGL2:EGFP

Figure 4. Distribution of VANGL2:EGFP in the ampulla of the oviduct

(**A-A”**) The ampulla of the oviduct was opened longitudinally and stained for acetylated-TUBULIN (green, a marker for cilia) and F-actin (magenta, a marker for cell edges). The boxed regions in **A** are shown in **A’** and **A”.** The direction of the ovary-uterus axis is indicated by an arrow in each image. Regions of the epithelium where SCCs were aligned circumferentially are highlighted by white boxes in **A’** and **A”**, and are shown in **B-D”**. (**B-D”**) Regions of the epithelium where circumferential alignment of SCCs was observed. Three examples are shown (**B-B”**, **C-C”**, and **D-D”**). (**B-D’**) SCCs are filled with magenta (**B-D’**), and cell edges are also outlined in magenta in **B–D**. Asterisks in (B’-D”) mark examples of circumferential SCC alignment. The distribution of VANGL2:EGFP signals is shown in **B’-D”** (green in **B’-D’**). White arrowheads mark VANGL2:EGFP signals enriched at cell edges running along the ovary-uterus axis. These enrichments were observed in MCCs adjacent to SCCs aligned circumferentially. In the same view field, enrichments of VANGL2:EGFP signals at cell edges running circumferentially were also observed (yellow arrowheads). These enrichments were found near low cells aligned along the ovary-uterus axis (e.g., compare yellow arrowheads with white arrowheads in **B-B”**). Yellow bars in B’-B” represent polarity of VANGL2:EGFP distributions at cell edges. The length and orientation of each bar indicate the magnitude and direction of polarity at VANGL2:EGFP-enriched cell edges, respectively. Compare the orientations of bars near the white arrowheads with those near the yellow arrowheads.

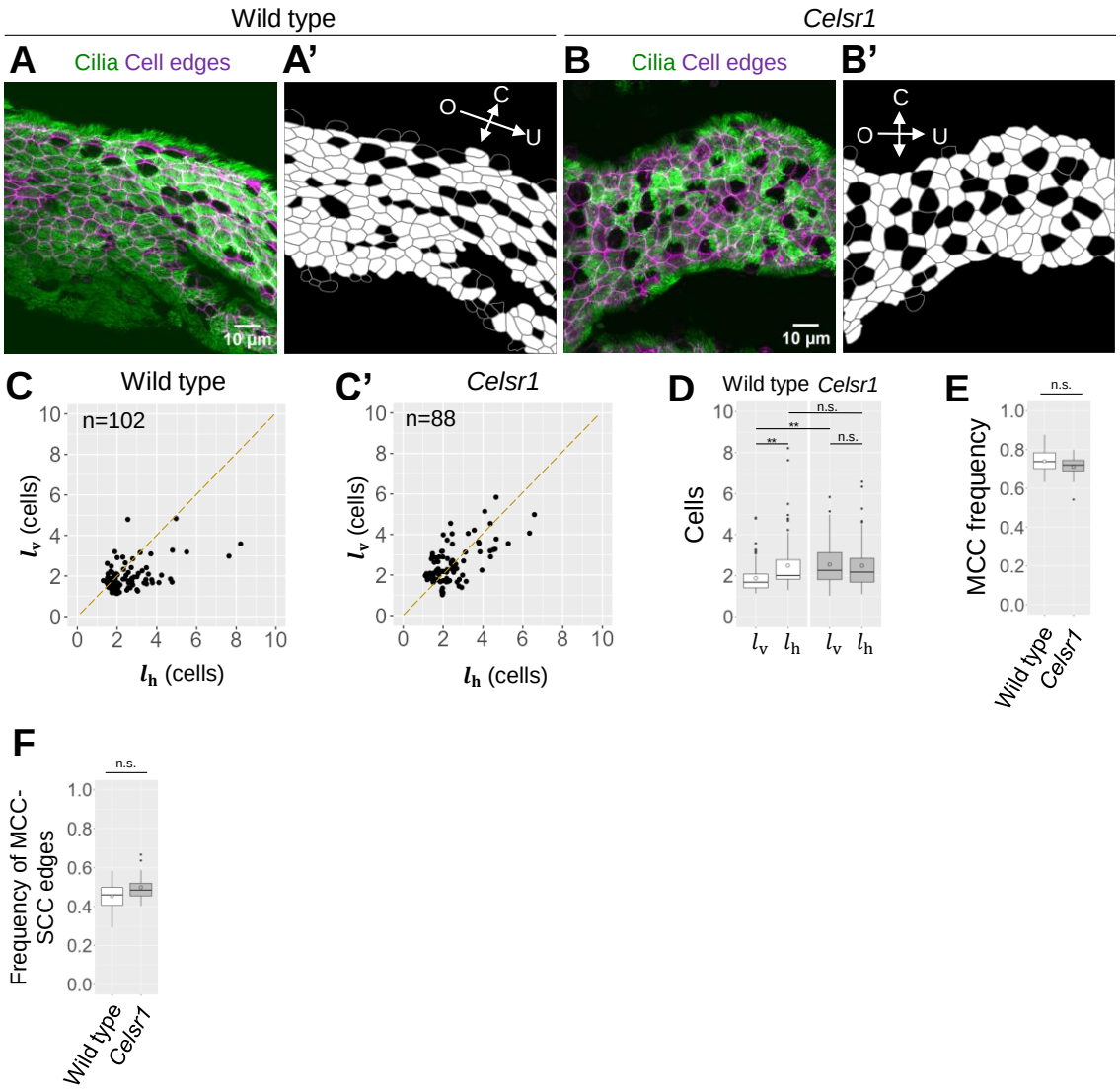

Figure 5. CELSR1 aligns SCCs along the ovary-uterus axis

(**A-B**) Oviduct epithelium of 11-week-old mice stained for F-actin (a cell edge marker, magenta) and acetylated-TUBULIN (a cilia marker, green). Images of wild-type (**A-A'**) and *Celsr1* mutant (**B-B'**) tissues are shown. (**A'** and **B'**) Black cells: SCCs; white cells: MCCs; gray lines: cell edges. (**C-D**) l_h and l_v were analyzed by a scatter plot (**C-C'**) and boxplot (**D**). **: p<0.01; n.s.: not significant (Steel-Dwass test). n: number of clusters. Isolated SCCs were excluded from the plots. Each dot represents an SCC cluster. (**E**) MCC frequencies in wild type and *Celsr1* mutant oviducts. n.s., not significant; Wilcoxon rank-sum test; n=20 view fields, 5 wild type oviducts; n=20 view fields, 5 *Celsr1* mutant oviducts. (**F**) Frequencies of edges between MCCs and SCCs in wild type and *Celsr1*

mutant oviducts. n.s.: not significant; Wilcoxon rank-sum test. n=20 view fields, 5 wild type mice; n=20 view fields, 5 *Celsr1* mutant oviducts. Open circles in (**D-F**) represent average values.