Epithelial Cell Polarity During Drosophila Midgut Development

Jia Chen and Daniel St Johnston*

Gurdon Institute and the Department of Genetics, University of Cambridge, Cambridge, United Kingdom

The adult Drosophila midgut epithelium is derived from a group of stem cells called adult midgut precursors (AMPs) that are specified during the migration of the endoderm in early embryogenesis. AMPs are maintained and expanded in AMP nests that lie on the basal side of the larval midgut throughout the larval development. During metamorphosis, the larval midgut undergoes histolysis and programmed cell death, while the central cells in the AMP nests form the future adult midgut and the peripheral cells form the transient pupal midgut. Here we review what is known about how cells polarise in the embryonic, larval, pupal and adult midgut, and discuss the open questions about the mechanisms that control the changes in cell arrangements, cell shape and cell polarity during midgut development.

Keywords: Drosophila, midgut, polarity, apical, basal, junction

INTRODUCTION

The Drosophila intestine is composed of several different cell types, including epithelial cells, muscle cells, neurons, and trachea cells (Miguel-Aliaga et al., 2018). The gut tube is formed by single layer of polarised epithelial cells surrounded basally by the muscles, trachea and nerves. The fly intestine is anatomically organised into the foregut, midgut and hindgut regions with the crop and Malpighian tubules emanating at the foregut/midgut and hindgut/midgut boundaries. The midgut is the longest section of the intestine and forms the conduit between the foregut and hindgut. It can be further subdivided into the anterior, middle and posterior midgut, which are marked by tissue constrictions and differences in luminal pH (Shanbhag and Tripathi, 2005, 2009). The epithelium performs the major functions of the midgut: acting as a barrier between the gut lumen and the inside of the organism and absorbing nutrients. It is composed of two types of mature epithelial cells: enterocytes (EC) and enteroendocrine (ee) cells. Based on cell morphology, physiology and gene expression profiles, epithelial cells in the midgut can be further classified into at least 10 different subregions and 22 clusters (Buchon et al., 2013; Marianes and Spradling, 2013; Dutta et al., 2015; Hung et al., 2020).

Despite their diverse shapes and gene expression profiles, all epithelial cells share the same features of apical-basal polarity along the whole midgut (Figure 1). The major type of epithelial cells in midgut, the ECs, are absorptive and are usually of cuboidal/columnar shape, but in the middle midgut, specialised acid secreting ECs called the copper cells adopt a cup shape (Hoppler and Bienz, 1994; Strand and Michelli, 2011). The apical membrane of ECs is covered in a brush border of

Abbreviations: AJ, adherens junction; AMG, presumptive adult midgut; AMIS, apical membrane initiation site; AMP, adult midgut precursor; APF, after puparium formation; BM, basement membrane; EB, enteroblast; EC, enteroocyte; ee, enteroendocrine cell; ECM, extracellular matrix; EMT, epithelial to mesenchymal transition; IC, interstitial cell precursor; ISC, intestinal stem cell; MET, mesenchymal to epithelial transition; PAC, preformed apical compartment; pISC, presumptive intestinal stem cell; pLee, progenitor cell for larval enteroendocrine cell; PM, peritrophic membrane; PMEC, principal midgut epithelial cell; PMG, posterior midgut primordium; SJ, septate junction; tPMG, transient pupal midgut; VM, visceral mesoderm.
microvilli, while infoldings of the basal membrane generate the basal labyrinth. Each structure serves to maximize surface area, which is thought to facilitate nutrient absorption, although how the basal labyrinth forms and functions is not well studied (Shanbhag and Tripathi, 2009; Sauvanet et al., 2015). The ee cells are secretory cells that release neuropeptide hormones in response to gut contents (Beehler-Evans and Michelli, 2015; Hung et al., 2020). They exist as isolated diploid cells throughout the midgut epithelium, often have a bottle or fusiform shape and have little or no basal labyrinth. Recent work has shown that ee cells lack an apical brush border (Xu et al., 2018), as observed in midgut endocrine cells in other insect species (Billingsley and Lehane, 1996).

The brush border is enriched in actin and contains MyoIA, MyoIB and other actin crosslinking proteins (Table 1) (Morgan et al., 1995; Crawley et al., 2014). The region of the cortex at the base of the microvilli, which is called the “terminal web” in mammalian cells, contains MyoIA, Myo7a, aPKC and Par-6 (Table 1) (Morgan et al., 1995; Chen et al., 2018). The apical domain is supported by a submembraneous spectrin scaffold composed of $\beta_4$, spectrin-$\alpha$-spectrin heterotetramers that links the membrane to the actin cytoskeleton (Table 1) (Baumann, 2001; Chen and St Johnston, 2022). The apical surface of the epithelial layer in the adult is closely associated with, but does not necessarily contact, the peritrophic membrane (PM), which is composed of Type I PM produced by the entire midgut and type II PM that is secreted by the cardia/proventriculus at the most anterior region in midgut (Lehane, 1997). The PM serves a similar function to the mucous lining in mammalian gut as the outmost protective barrier (Zhang et al., 2017).

As mentioned above, the basal sides of the epithelial cells contact the extracellular matrix (ECM), except for the invaginations of the basal labyrinth, which do not appear to have any ECM in their lumens (Baumann, 2001; Shanbhag and Tripathi, 2009). Integrin associated proteins, such as Integrin linked kinase (Ilk), Rhea and Fit localise to the basal cortex (Table 1) (Chen et al., 2018). The ECM is assembled into a sheet-like basement membrane (BM) between the epithelium and the visceral muscle layers (Shanbhag and Tripathi, 2009). All four main types of the basement membrane components are present: type IV collagen ($\alpha_1;\alpha_2$ heterotrimers with Col4a1 as the $\alpha_1$ subunit and Vkg as the $\alpha_2$ subunit), Laminins (a$\beta$y-heterotrimers with LanA and Wb as $\alpha$ subunits, LanB1 as the $\beta$ subunit and LanB2 as the $\gamma$ subunit), Nidogen and Perlecan (Table 1) (Broadie et al., 2011; Davis et al., 2019; Töpfer and Holz, 2020). Laminins and Type IV collagen form independent mesh-like structures with the Laminins closer to the epithelial cells. In addition, the gut BM contains Netrins (Fert et al., 2015), Secreted protein, acidic, cysteine-rich (SPARC) (Martinek et al., 2002, 2008), Macrophage derived proteoglycan-1 (MDP-1) (Kramerova et al., 2003), Glutactin (Olson et al., 1990) and Peroxidasin (Table 1) (Nelson et al., 1994). Unlike most other fly epithelia, the *Drosophila* midgut epithelium is derived from the endoderm and the intercellular junctions in both the EC and ee cells have a different morphology and arrangement from the junctions in non-endodermal epithelia (Figure 1). Endodermal epithemia form smooth septate junctions (sSJs), analogous to tight junction in mammals, which lie apical to the adherens junctions (AJs), whereas in other epithelial cells, the electron-dense AJs lie above the septate junctions, which are pleated not smooth (Figure 1) (Lane and Skaer, 1980; Tepass and Hartenstein, 1994b; Baumann, 2001). Recent studies reveal that the smooth SJs are organised by the endoderm-specific proteins, Mesh, Snakeskin, Tsp2a and Hoka, which form a transmembrane protein complex (Table 1) (Izumi et al., 2012, 2016, 2021; Furuse and Izumi, 2017). The SJs at the vertices where three cells meet contain additional components, including Bark, Gli and M6, which are also found in the tri-cellular junctions in epithelia with pleated SJs (Table 1) (Schulte et al., 2003; Byri et al., 2015; Hildebrandt et al., 2015; Bosveld et al., 2018; Esmangart de Bouronville and le Borgne, 2020; Wittek et al., 2020). Loss of these tri-cellular SJs proteins during ageing leads to defects in the function of the intestinal barrier in older flies (Resnik-Docampo et al., 2017). In ectodermally-derived epithelia, Sidekick localises to the tri-cellular AJs and modulates apical adhesion and tension during the active junctional remodelling during embryo morphogenesis (Finegan et al., 2019; Letizia et al., 2019; Uechi and Kuranaga, 2019). It is not known whether bi- or tri-cellular AJ in the midgut also contain specific components since ECad, Arm and $\alpha$-Cat are the only known components of AJs in the midgut (Choi et al., 2011; Campbell and Casanova, 2015; Liang et al., 2017).

During the past 20 years, *Drosophila* midgut has proven an exciting model system to study epithelial homeostasis, since...
### TABLE 1 | Drosophila genes and their encoded protein’s localization during midgut development.

| **Drosophila gene (Abbreviation)/Alias** | **Human ortholog** | **Protein type** | **Protein localisation in the Drosophila midgut epithelial cell** | **References** |
|----------------------------------------|-------------------|-----------------|---------------------------------------------------------------|--------------|
| Myosin 31DF (Myo31DF)/MyoA             | MYOID             | Myosin          | Apical brush border and terminal web in stage 17 E\textsuperscript{1}, L\textsuperscript{2} and A\textsuperscript{3} | Morgan et al. (1995); Crawley et al. (2014) |
| Myosin 61F (Myo61F)/MyoB               | MYOIC             |                 | Relocates from basolateral domain to apical brush border in stage 17 E; apical brush border in L and A | Chen et al. (2018) |
| Crinkled (ck)/myosin VIIA (myo7a)      | MYO7A             | Kinase          | Apical in A                                                  | Chen et al. (2018) |
| Par-6                                  | PRKCZ             |                 | Apical side of the lateral junction in stage 9 E            | Campbell et al. (2011) |
| Bazooka (baz)/par-3                    | SPTBN6            | Spectrin        | Apical domain in L and A                                     | Baumann, (2001); Chen et al. (2018) |
| Karst (kst)/β-spectrin                 | SPTBN1            |                 | Basolateral domain in L and A                                | Chen and St Johnston, (2022); Devenport and Brown, (2004) |
| βSpectrin (β-Spec)                     | SPTAN1            |                 | Cell cortex in L and A                                       | Chen and St Johnston, (2022); Devenport and Brown, (2004) |
| Myospheroid (mys)/βPS                  | ITGA6/7           | TM\textsuperscript{5} | Apical in stage 9 E                                          | Campbell et al. (2011) |
| Inflated (if)/αPS2                     | ITGA8             | PDZ, TM         | Apical in stage 12–15 E; basal in A                         | Campbell et al. (2011) |
| Myospheroid (mys)/βPS                  | ITGB1             | TM, TM          | Muscle layer                                                | Yee and Hynes, 1993; Martin-Bermudo et al., 1999; Lin et al., 2013; Okumura et al., 2014; Pitsidianaki et al., 2021 |
| Integrin linked kinase (ilk)           | NEO1              | FERM\textsuperscript{46} | Basal domain in stage 12 E and A                             | Chen et al. (2018) |
| Multiple edematous wings (mew)/αPS1    | ITGA8             | PDZ, TM         | Basal domain in stage 12 E and A                             | Chen et al. (2018) |
| Rhead/talin                            | TLN               | TM              | Basal domain in from stage 12 E                             | Pert et al. (2015) |
| Frazzled (fra)/DCC                     | DAG1              | FERM            | Tissue constriction region in stage 16 E                    | Schneider and Baumgartner, (2008) |
| Division abnormally delayed (daily)    | GPC5              | Glypican TM     | Mainly basal in stage 12–15 E; basal in A                   | (Yee and Hynes, 1993; Martin-Bermudo et al., 1999; Lin et al., 2013; Okumura et al., 2014; Pitsidianaki et al., 2021) |
| Daily-like (dlp)                       | GPC4              |                 | Muscle layer                                                | Yee and Hynes, 1993; Martin-Bermudo et al., 1999; Lin et al., 2013; Okumura et al., 2014; Pitsidianaki et al., 2021 |
| Syndecan (Sdc)                         | SDC               | Proteoglycan TM | Mainly basal in stage 12–15 E; basal in A                   | Pert et al. (2015) |
| Laminin A (LanA)                       | LAM5              | ECM             | LanA heterotrimer is mainly basal between the endoderm and mesoderm, also surrounding ICP cells and weakly at apical side in E; basal in L and A | Wang and Holz, 2012; Lin et al., 2013; You et al., 2014; Pert et al., 2015; Töpfer and Holz, 2020; Pitsidianaki et al., 2021 |
| Wing blister (wb)                      | LAM4              | ECM             | Basal ECM in E, L and A                                     | (Yee and Hynes, 1993; Martin-Bermudo et al., 1999; Lin et al., 2013; Okumura et al., 2014; Pitsidianaki et al., 2021) |
| LamB1/LamininB1                        | LAM2A             | ECM             | Basal ECM in stage 16 E, L and A                             | (Yee and Hynes, 1993; Martin-Bermudo et al., 1999; Lin et al., 2013; Okumura et al., 2014; Pitsidianaki et al., 2021) |
| Collagen type IV alpha 1 (Col4a1)/Cg25C| COL4A1            | ECM             | Basal ECM in stage 16 E, L and A                             | (Yee and Hynes, 1993; Martin-Bermudo et al., 1999; Lin et al., 2013; Okumura et al., 2014; Pitsidianaki et al., 2021) |
| Viking (Vkg)                           | COL4A1            | ECM             | Basal ECM in stage 16 E, L and A                             | (Yee and Hynes, 1993; Martin-Bermudo et al., 1999; Lin et al., 2013; Okumura et al., 2014; Pitsidianaki et al., 2021) |
| Snakeskin (Ssk)                        | SUSD2             | ECM             | Basal ECM in stage 16 E, L and A                             | (Yee and Hynes, 1993; Martin-Bermudo et al., 1999; Lin et al., 2013; Okumura et al., 2014; Pitsidianaki et al., 2021) |
| Tetraspanin 2A (Tsp2A)                 | TSPAN8            | TM              | SJs from stage 16 E, L and A                                | Olson et al. (1990) |
| Peroxidasin (Pxn)                      | PXDN              | ECM             | Basal ECM in E                                              | Nelson et al. (1994) |

(Continued on following page)
TABLE 1 (Continued) Drosophila genes and their encoded protein’s localization during midgut development.

| Drosophila gene (Abbreviation)/Alias | Human ortholog | Protein type | Protein localization in the Drosophila midgut epithelial cell | References |
|------------------------------------|----------------|--------------|-------------------------------------------------------------|------------|
| Bark beetle (bark/anakonda) (aka) | -              | TM           | Tri-cellular junctions in E                                | Byri et al. (2015); Wittek et al. (2020) |
| Gliotactin (Gil)                   | GPM6A          | TM Cadherin  | Apical side of the lateral junction in stage 9 E; AJ in A   | Campbell et al. (2011); Chen et al. (2018) |
| M6                                 | CDH20          |              |                                                             | Chen et al. (2018) |
| Shotgun (shg)/DECad                | CTNNB1         | Armadillo repeat Catenin | AJ in A                                             | Takashima et al. (2011b) |
| armadillo (arm)/β-catenin           | CTNNB1         | Armadillo repeat Catenin | AJ in A                                             | Takashima et al. (2011b) |
| α Catenin (α-Cat)                   | CTNNA          | PDZ          | Apical side of the lateral domain in the developing adult midgut at pupal stage | Chen et al. (2018) |
| Discs large 1 (dlg1)               | DLG1           | PDZ          |                                                             | Chen et al. (2018) |
| Fasciclin 3 (Fas3)                 | NECTIN3        | TM           |                                                             | Chen et al. (2018) |

- Not found.

#1,2,3 E, L and A denote the embryonic, larval and adult midgut epithelium separately.

#4, PDZ domain containing scaffolding protein.

#5, TM denotes transmembrane protein.

#6, FERM domain containing protein.

Basally-localised intestinal stem cells (ISC) can divide and differentiate into both ECs and ee cells in the adult midgut. The signals and mechanical cues that regulate ISC division and differentiation have been extensively characterised and have been summarised in many excellent reviews of this topic (Micchelli and Perrimon, 2006; Ohlstein and Spradling, 2006; Jiang and Edgar, 2011; Lucchetta and Ohlstein, 2012; Zeng et al., 2013; Antonello et al., 2013; He et al., 2018; Miguel-Aliaga et al., 2018; Billingsley and Lehane, 1996; Caccia et al., 2019). The exact mechanism that polarises ECs and ee cells in the adult midgut is not understood, but this does not require any of the canonical epithelial polarity factors that polarise non-endodermal epithelia, including Bazooka (Par-3), Par-6, atypical protein kinase C (aPKC), Crumbs (Crb), Stardust (Sdt), Discs large (Dlg), Lethal (2) giant larvae (Lgl) or Scribble (Chen et al., 2018). Instead, the basally localised integrin associated proteins, Rhea and Fit1 are required for all steps in EC polarisation and sJ components are required for the formation of the PAC during EB integration (Chen et al., 2018; Chen and St Johnston, 2022). The progenitor and precursor cells, ISCs and EBs, lie at the basal side of the epithelium without any access to the apical lumen and do not form sJs with neighbouring cells. This indicates that the midgut epithelial cells require sustained basal signalling from the contact with the ECM to polarise in a basal to apical fashion. Their further polarisation, including the formation of sJ and the apical brush border, requires positional cues from the SJs and the gradual growth of the apical domain via polarised membrane trafficking. The Drosophila midgut epithelium provides an excellent model for mammalian epithelia, which have a similar junctional arrangement and also require ECM contacts for polarity (Yu et al., 2005). The adult Drosophila midgut epithelium is derived from AMPs, which are specified during early embryogenesis and segregated from the cells of the larval and pupal midgut during development (Campos-Ortega and Hartenstein, 1985; Teppas and Hartenstein, 1994a, 1995; Takashima et al., 2011b, 2011a, 2016a). Developmentally, the midgut epithelium is categorised as “a secondary epithelium”, since it goes through an epithelial-to-mesenchymal-transition (EMT) during early endoderm formation and later undergoes a mesenchymal-to-epithelial-transition (MET) to repolarise. In embryos, both the migration of the midgut primordia and repolarisation require
basal contact with the mesoderm and ECM components surrounding the endoderm layer (Tepass and Hartenstein, 1994a; Yarnitzky and Volk, 1995). It has been suggested that a similar mechanism is deployed during EB polarisation and differentiation, when EBs acquire a migratory potential before repolarising and integrating into the epithelial layer (Micchelli, 2012; Antonello et al., 2015). In this review, we will describe what is known about cell polarity changes during embryonic, larval and pupal midgut development and discuss what this suggests about the mechanisms of apical-basal polarisation in endodermal tissues.

**Cell Polarity During Embryonic Midgut Development**

The *Drosophila* midgut primordium forms from the endoderm during gastrulation (Campos-Ortega and Hartenstein, 1985). Under the coordinated action of the GATA transcription factor Serpent and the winged-helix transcription factor Forkhead, the posterior midgut primordium (PMG) together with the ectodermally-derived hindgut primordium are internalised into the embryo (Weigel et al., 1989; Reuter, 1994; Nakagoshi, 2005). The PMG cells initially have the same apical-basal polarity as all ectodermal cells, which is established during the process of cellularisation (Tepass and Hartenstein, 1994b). Stranded-at-second (Sas), and the canonical apical polarity factors, Crb and Sdt, localise to the apical surface and Baz and ECad are localised to the apical AJ (Table 1) (Figure 2A) (Campbell et al., 2011). During stage 10 of embryogenesis, Serpent induces the PMG to undergo an EMT and become migratory by repressing the expression of Crb, Sdt, Sas, and pleated SI genes (Tepass and Hartenstein, 1994b; Campbell et al., 2011). As a result, the apical AJ dissolve and ECad and Baz relocalise from the AJ to dynamic puncta at cell-cell contacts. By stage 11, the PMG has established contact with the visceral muscle primordium and uses it as a substrate for its migration (Tepass and Hartenstein, 1994a). Three different cell types can be distinguished transcriptionally and morphologically among the migrating midgut mass. Most cells are principal midgut epithelial cells (PMECs), which will give rise to the larval midgut ECs and always contact the muscle primordium. The other two populations of mesenchymal cells, interstitial cell precursors (ICPs) and AMPs, are attached to the apical surface of PMECs and are carried along by the latter. ICPs express Insuteable and Asense from late stage 10 to mid-stage 11, and AMPs, which will
give rise to the future adult midgut, are Asen-tive-positive from early stage 11 to late stage 12 (Figure 2B) (Tepass and Hartenstein, 1994a, 1995; Campbell and Casanova, 2015). ICPs and AMPs delaminate sequentially from the outer layer of PMECs between stage 10–11 during their posterior migration (Tepass and Hartenstein, 1995). Some AMPs at this stage can also be marked with anti-Pros antibody staining, suggesting that they may be progenitor cells for future larval ee cells (pLee), although there is no lineage tracing data to support this. Like AMPs, pLees remain in the mesenchymal inner mass during migration, but become esg- and segregate from the AMPs by stage 14 (Jiang and Edgar, 2009; Takashima et al., 2011a). The cohesive and ordered migration of these three/four types of cells along the visceral mesoderm is coordinated through ECad-mediated cell adhesion and relies on the Integrin/Laminin and Frazzled/Netrin signalling pathways (Martin-Bermudo et al., 1999; Devenport and Brown, 2004; Campbell and Casanova, 2015; Pert et al., 2015; Pitsidianaki et al., 2021). Between late stage 11 and stage 12, shortly before and during germ band retraction, the PMECs reorganize and go through MET to form the midgut epithelium. By the end of germ band retraction at stage 13, the anterior and posterior midgut rudiments approach each other and finally fuse, the PMECs assume a columnar shape and the ICPs form two clusters in the middle of the developing midgut (Tepass and Hartenstein, 1994a). MET coincides with the downregulation of Fkh and Srp (Weigel et al., 1989; Campbell et al., 2011). However, Srp down-regulation is not sufficient to trigger MET, which instead depends on basal cues from Laminin and Netrins produced by the visceral mesoderm acting through Integrins and Fra respectively (Pert et al., 2015; Pitsidianaki et al., 2021) (Figure 2B and discussed later).

During endoderm migration, the ECM between the endoderm and the mesoderm is not yet fully organised, since early electron microscopy studies demonstrated that PEC migration is mediated through direct mesoderm/endoderm contact without any detectable ECM or junctional specialisations (Tepass and Hartenstein, 1994a). However, Srp activates *LanB1* and *LanB2* RNA expression in stage 11 midgut primordium cells (Wolfstetter and Holz, 2012; Töpfer et al., 2019). Moreover, the laminin matrix secreted by the visceral muscle primordium contains Wb, which is thought to induce MET, whereas that secreted by endodermal cells contain LanA, and both LanA and Wb play crucial roles in controlling the speed of migration (Urbano et al., 2009; Wolfstetter and Holz, 2012; Pitsidianaki et al., 2021). At this stage, haemocytes (migrating macrophages) are the only source of secreted type IV collagen and Perlecan (Matsubayashi et al., 2017) and they do not reach the endoderm until after the migration is complete (Urbano et al., 2011; Pitsidianaki et al., 2021). Nidogen is reported to have similar expression pattern to *LanB1* during embryogenesis but is not required for endoderm migration or formation (Urbano et al., 2009; Dai et al., 2018; Töpfer and Holz, 2020). At stage 16, Laminins, Collagens, Nidogen, and Perlecan, as well as other mature ECM components, such as MDP-1 and SPARC are all found in between the endoderm and mesoderm, forming a more complex ECM network (Wolfstetter and Holz, 2012).

Cells rely on ECM receptors to receive migratory/adhesive cues from the ECM, including Integrins, Fra, Dystroglycan (Dg), the Glycips Daily and Dally-like and Syndecan (Sdc) (Table 1). Integrins function as heterodimers of α and β subunits and are required for both midgut migration and visceral muscle formation (Devenport and Brown, 2004). Flies have five α integrin subunits, αPS1-5 and two β subunits, Mys and βν (Table 1). Embryonic midgut migration requires the expression of both αPS1 in the endoderm and αPS2 in the visceral muscle, while αPS3 cooperates with αPS1 in the endoderm layer but is not required (Brown, 1994; Stark et al., 1997; Martin-Bermudo and Brown, 1999; Martin-Bermudo et al., 1999). Phylogenetic studies show that the αPS3-5 subunits are closely related and the result of gene duplication events (Hughes, 2001). αPS3 and αPS4 are expressed in adult midgut ECs, whereas αPS5 is not (Lin et al., 2013; Patel et al., 2015). Mys is widely expressed and is essential for viability, whereas βν is specifically expressed in the developing endoderm and the larval and adult midgut, but is not required for viability or fertility (Yee and Hynes, 1993). Integrins must form heterodimers in the endoplasmic reticulum to be trafficked to the cell surface and flies without both β subunits have no integrin function at all (Leptin et al., 1989; Devenport and Brown, 2004). Both αPS1/Mys and αPS3/βν pairs of integrins can be found in the migrating endoderm at late stage 11, with αPS1/Mys localising to basal side and αPS3/βν localising mainly apically at the end of migration (Figure 2B) (Devenport and Brown, 2004; Pitsidianaki et al., 2021). Two of the three Dystroglycan splicing isoforms are expressed in the midgut at stage 16, but their functions have not yet been characterised (Schneider and Baumgartner, 2008). Fra localises to the basal side of the PMECs at stage 12 and to the basal and junctional domain of the migrating midgut cells at stage 13 (Figures 2B,C). Interestingly, AMPs, which normally remain apical to the migrating PMECs at stage 12, are mis-localised and contact the visceral muscle in *netrin* mutant embryos. This phenotype has been attributed to the dis-organisation and loose adhesion of the PMG epithelium, rather than loss of direct signalling to AMPs, (Pert et al., 2015).

The PMECs are the first cell-type in the midgut primordia to go through MET, with AMPs, pLees, and ICPs remaining mesenchymal in the apical lumen until later. Although the exact time at which AMPs invade and translocate across the epithelium is not defined, they are located at the basal side of the gut in newly hatched larvae while the pLees have polarised and integrated into the epithelium (Hartenstein and Nung Jan 1992; Micchelli, 2012). This raises the question of how AMPs translocate to the basal side of the epithelium, since the apical junctions between the PMECs, which are marked by ECad, start to develop during migration in the outermost trailing region of the posterior midgut and sSJs start to develop in midgut from stage 15. Moreover, it is not clear whether the pLees become polarised and integrate into the epithelium during translocation or repolarise/integrate after translocating to the basal side (Takahima et al., 2011a). It has been hypothesized that the early delamination and late segregation and translocation of the AMPs and ICPs are due to differences in cell-cell affinity (Tepass and Hartenstein, 1995). However, there are no defects in the apical location of AMPs in *Ecad/shg* mutant embryos, it is therefore unclear whether their delamination and translocation is...
a passive cell-sorting event or an active migration process (Tepass and Hartenstein, 1994a). Furthermore, it will be important to determine the relationship between cell fate determination and the corresponding EMT–MET processes.

By stage 15, the visceral mesoderm (VM) expands ventrally and dorsally to form the circular mesoderm and the endodermal layer follows this movement to form a closed chamber. Although the early specification of the endoderm into distinct PMEC, ICP, pLee and AMP cell types does not depend on interaction with the mesoderm, VM induces the further specification and development of future larval midgut epithelium after the midgut rudiments fuse, including the formation of the three midgut constrictions during stages 14–16 and the specification of the middle midgut region and proventriculus (Nakagoshi, 2005).

Between stage 16 and 17, the future larval ECs change their morphology from short cuboidal cells to tall columnar cells and develop elaborate cellular junctions and an apical brush border (Morgan et al., 1995). Smooth SJ components start to express during stage 12 and become localised at stage 16, but mature sSJs only become visible at late stage 17 (Tepass and Hartenstein, 1994a; Izumi et al., 2012). Myo61F relocates from the basal-lateral region to the apical microvilli, coincident with the disappearance of the yolk mass which indicates the start of digestive function (Morgan et al., 1995).

*Drosophila* embryonic midgut formation takes less than 9 h, between stage 10 when PMG starts EMT and stage 15 when the midgut migration finishes. Both EMT and MET happen gradually, whereas cell polarity changes dramatically and rapidly during this process. The apical polarity factor Grb disappears early on, the original apical-lateral junctions dissolve, giving rise to a group of mesenchymal migratory cells connected by limited spot AJs. These cells later re-polarise forming smooth SJJs rather than pleated SJJs at the apical/lateral side of the cell-cell junctions. Embryonic midgut development also demonstrates the importance of the sustained basal signalling from the mesoderm much like the polarisation of the EBs in the adult midgut epithelium. It is still unclear, however, what lies downstream of basal ECM and their receptors to induce epithelial polarisation, and whether apical extracellular LanA and apically localised αPS3/βv integrins plays any role in polarising the embryonic midgut epithelium. Past work has focused on the morphological development of the midgut and the genetic control of endoderm formation and differentiation (Bilder and Scott, 1995; Harbecke and Lengyel, 1995). Much less is known about the genetic control of AMP and ICP delamination and translocation, which also involves the loss and gain of cell polarity. These processes are challenging to study, however, because they occur over short time periods in the centre of the embryo.

**Larval Midgut Epithelial Cells**

The larval midgut is composed of anterior, middle and posterior regions, each maintaining a different pH, and is anatomically similar to the adult midgut, although the constriction around the middle midgut is less obvious (Shanbhag and Tripathi, 2005; Overend et al., 2016). The larval midgut contains four gastric caeca, which are blind sacs that emerge from the anterior midgut just posterior to the proventriculus. They persist in the larva but are lost during pupation and are not present in the adult fly (Skaer, 1993). Larval ECs are polyploid and derive from PMECs, whereas larval ee (lee) cells are diploid and derive from pLees (Takashima et al., 2011a). Cell specification has been well-studied in the larval middle midgut (Hoppler and Bienz, 1994, 1995). Large cells in this region were first called calycocytes, and were later named cuprophilic or copper cells, since they accumulate copper and display orange fluorescence when the larvae are fed with copper-enriched food. This property is attributed to the binding of copper ions to metallothionein, which is constitutively expressed in the cytoplasm in the middle midgut region (Skaer, 1993; Durliat et al., 1995; McNulty et al., 2001). It has been proposed that copper cells derive from the ICPs, although this is at odds with the observation that the ICPs disseminate over the whole embryonic midgut after stage 15 (Poulson and Waterhouse, 1960; Skaer, 1993). The copper cells are cup-shaped, with an invaginated apical domain containing long microvilli. They are surrounded by columnar interstitial cells with a normal apical domain, short microvilli and a more extensive basal labyrinth (Filshie et al., 1971). It is thought that the copper cells are the acid secreting cells, based on the correlation between the number of residual copper cells in *labial* mutant larvae and the number of remaining acid-retaining cells (Hoppler and Bienz, 1994; Dubreuil et al., 1998; Dubreuil, 2004). Several V-ATPase and other ion transporters are required for the acidic pH generation (Overend et al., 2016; Tian et al., 2022). Interestingly, copper absorption from the food can inhibit acid secretion and the acid secretion defective α-spec mutant copper cells are not able to accumulate copper, which raises the question of how copper absorption and acid secretion are linked (Dubreuil et al., 1998; McNulty et al., 2001). Furthermore, we still do not know how and why the apical domain in the copper cells invaginates nor how the interdigitated arrangement of copper cells and interstitial cells arises. One clue comes from the stage 15 embryonic midgut, when the inner ICPs interdigitate between the outer *labial*-positive ICPs (Skaer, 1993), which means that the arrangement of copper cells and interstitial cells are probably also under the control of *labial*.

The larval midgut is remarkably similar to the adult midgut at the level of cellular structure, with an apical brush border facing the gut lumen, a basal side in contact with the visceral muscle and a basal labyrinth of invaginations from the basal membrane (Figure 3) (Shanbhag and Tripathi, 2005, 2009). The larval midgut also forms sSJs (Izumi et al., 2012, 2016, 2021) and the apical domain is enriched for actin and β₁H-spectrin α-spectrin, while β-spectrin/α-spectrin heterotetramers label the basolateral domain (Dubreuil et al., 1998). Spectrins are not required for copper cell polarity, but loss of β₁H-spectrin leads to loss of the apical proton pump, the H⁺V-ATPase which probably causes the defect in acid secretion seen in α-spec mutant larvae (Phillips and Thomas, 2006). The two class I myosin family proteins, Myo31DF and Myo61DF can also be found in the apical terminal web and brush border microvilli in the larval ECs, but neither is required for cell polarity or brush border organisation (Morgan et al., 1995; Okumura et al., 2015). Interestingly, the AJs marked by ECad and Baz localise apical to the sSJ before the embryo hatches, whereas, AJs localise to the basal side of the sSJ in the first instar larva and adult ECs (Tepass
et al., 2011). 1-3 of the cells in an AMP nest differentiate into STAT92E
midgut (Takashima et al., 2011b). The AMP nests can reach 2/3 of
cells and become an integral part of the future transient pupal
(Houtz et al., 2019). Some AMP cells also differentiate into Pros +
regenerative when the larval midgut is challenged with infection
However, the central cells can differentiate and become partially
central AMP cells until metamorphosis (Mathur et al., 2010).

Peripheral cells are post-mitotic with bigger nuclei than the central
elongate and surround the inner mass of small, round central cells.
Peripheral cells are post-mitotic with bigger nuclei than the central
they function as a niche to maintain the stem-cell state of the
cells. They function as a niche to maintain the stem-cell state of the
central AMP cells until metamorphosis (Mathur et al., 2010).
However, the central cells can differentiate and become partially
regenerative when the larval midgut is challenged with infection
(Houtz et al., 2019). Some AMP cells also differentiate into Pros +
cells and become an integral part of the future transient pupal
midgut (Takashima et al., 2011b). The AMP nests can reach 2/3 of
the height of the larval ECs, but do not reach the apical lumen,
presumably because they cannot pass the sSJs between the larval
ECs (Figure 3B). Both peripheral cells and central cells appear to
maintain the contact with ECM and the peripheral cells contact the
larval ECs. It is not known how the peripheral cells adopt a sheath-
like shape and encase the central cells, nor how they provide a niche
for the central AMP cells, except that the Dpp signalling is required
(Mathur et al., 2010).

Pupal Midgut Epithelial Layer and the Formation of Adult Midgut Epithelium

Shortly after puparium formation (APF), the larval midgut
shortens, bringing the scattered AMP nests together. The
outer peripheral cells contact each other first and by 6h APF
become squamous and join together to form a multi-layered
sheet called the transient pupal midgut (tPMG). Central cells
also change their shape, flattening longitudinally and expanding
laterally, to form a continuous layer of presumptive adult
midgut (AMG) in a process that is thought to be MET
(Takashima et al., 2011b). At this stage, the larval midgut,
tPMG and AMG are all connected via spot AJ. At 8 h APF,
the AMG starts to show polarised features, with aPKC localising apically, Fas3 at the apical and lateral domains and Arm along the lateral and basal domains. Precursors of the adult ISCs, the presumptive intestinal stem cells (pISCs), remain at the basal side of the AMG layer (Figure 4A). At the same time, the tPMG also differentiates to a certain degree, forming microvilli and containing ec cells that have a spindle shape and remain detectable until 24 h APF (Figure 4A, 3B) (Takashima et al., 2011b; Takashima et al., 2016). At 6 h APF, the ECM layer surrounding the midgut and visceral muscle starts to break down and disappears by 24 h APF. By 36 h APF, myo fibrils disappear since the visceral muscle fibres surrounding midgut de-differentiate into secondary myoblasts (Aghajanian et al., 2016). During this time, both the tPMG and the AMG keep differentiating. The tPMG develops pleated SJs, whereas the AMG develops apical microvilli and smooth SJs at the apical side of the lateral domain. The AMG is in direct contact with the myoblasts since no ECM is observed in between (Aghajanian et al., 2016). Interestingly, an electron dense liquid has been observed separating the larval and tPMG from the AMG at this stage (Figure 4B) (Takashima et al., 2011b). Both the myo fibrils and ECM reorganise and reappear by 48 h APF (Aghajanian et al., 2016). Between 48 and 72 h APF, some esg-positive pISCs express Pros and divide asymmetrically to give rise to the adult ec cells (Guo and Ohlstein, 2015). The re-emergence of ECM is thought to be important for the pISC division and specification at this stage (Aghajanian et al., 2016). During later stages of metamorphosis, the larval and transient pupal midguts remain closely associated and further contract and become the “yellow body” in the lumen of the developing adult midgut. They are eventually discharged from the intestinal tract after eclosion.

The separation between larval midgut/tPMG and the AMG is essentially the delamination of larval midgut epithelial cells and the detachment between peripheral cells and central cells of the AMP nests. This results in the reorganisation of the tissue into three layers with spot AJ still present among them (Takashima et al., 2011b). The reorganisation happens within the first 12 h during pupal development, while the visceral muscle and ECM are still present. Both the tPMG and AMG keep differentiating, but only the AMG remains attached to the ECM, which means the separation cannot be simply explained by apoptosis-induced cell extrusion. It would be interesting to find out whether basal integrin adhesion is weakened in the larval epithelium and tPMG but retained in the AMG. Many other questions still remain about the adult midgut formation during pupal metamorphosis. First of all, before metamorphosis begins, there is direct signalling between the peripheral cells and central cells in the seemingly compact AMP nests, but almost nothing is known about the molecules that mediate adhesion between them or the molecular mechanisms that control the separation and reorganisation of the tPMG and AMG. Secondly, although the tPMG loses contact with the ECM and muscle layer, it still manages to differentiate to form pleated SJ. The functional significance of this junction and how it is formed are unclear. Thirdly, the AMG cells are believed to go through MET as they polarise, while the pISCs remain basal and in contact with the re-formed ECM and muscle layer. Based on what we know about the formation of the embryonic and adult midgut, it will be interesting to determine whether pISC specification requires a similar translocation process to AMPs in the embryo and if the AMG cells polarise in the same way as adult EBs and form a PAC as they integrate into the epithelium, and if their polarisation requires basal integrin signalling and SJ components.

CONCLUDING REMARKS

AMPs are specified at an early embryonic stage, delaminate apically but remain attached to the PMECs via spot AJ s and stay in the apical lumen of the migrating midgut primordium. They then translocate across the newly formed epithelium at the end of midgut development and remain basally after the embryo hatches. The mesoderm is not involved in the AMP specification and delamination, whereas cell-cell adhesion is proposed to play an important role in both the delamination and translocation. These processes are accompanied by the migration and repolarisation of the midgut primordium to form the future larval midgut epithelium. The migration and repolarisation require secreted Laminins from both germ layers, LanW at the basal side from the mesoderm and LanA at the apical side from the endoderm. The LanW from the basal side interacts with integrins receptors to activate downstream signaling pathways that are proposed to provide the cue that polarises the midgut epithelium and induce further polarised trafficking. The smooth SJs and the apical brush border microvilli form as the last step of polarisation in the epithelium. The polarised membrane features in the embryonic midgut epithelium are different from the steady state adult midgut epithelium, where integrin signalling components are only found basally and the lateral cell-cell junctions are clearly separated into apical-lateral sSJs and basal-lateral AJs. However, similar transcription factors control adult ISC maintenance and differentiation and embryonic midgut morphogenesis (Okumura et al., 2016). Moreover, EBs also go through a migratory stage before repolarising into ECs (Antonello et al., 2015). This means that the molecular mechanisms governing cell migration, cell translocation and MET-EMT could be the same in the embryo and adult.

During larval development, the AMPs expand, differentiate and form a nest containing peripheral cells and central cells. The peripheral cells are polarised to form sheath that surrounds and presumably isolates the central cells from the larval epithelial cells. It is not clear what type of cell-cell junctions form in the AMP nest and between the nest and larval epithelial cells. The peripheral cells later separate from the central cells to form the tPMG and delaminate with the larval epithelium at the start of pupation. By contrast, the central cells remain in contact with the basement membrane while adhering with each other to form the future AMG. Although both peripheral cells and central cells originate from AMPs, the peripheral cell-derived tPMG will develop pleated
SJs instead of smooth SJs. During the separation and reorganisation, spot AJSs are found connecting the larval midgut, tPMG and AMG. This raises the possibility that cell-cell adhesion dynamics regulate the separation. After the visceral muscle and ECM layer reform at the basal side, the central cells start to polarise. Little is known about how polarised domains form in the AMG, except that aPKC localises to the apical domain, Dlg and FasIII occupy the apical-lateral junction and Arm/Ecad are localised at the basal-lateral domain (Takekawa et al., 2011b). Since both embryonic midgut formation and EB polarisation in the adult midgut require sustained basal signalling, it seems likely that the AMG requires basal signalling from the newly-formed ECM to polarise, but the molecular mechanisms remain to be discovered.

In summary, studies on the behaviour of stem cells and the stem cell niche in the Drosophila midgut during embryonic, larval, pupal and adult development have paved the way for investigations into how cells are specified at each stage and how their polarity is controlled. Elucidating the roles of cell-cell interactions and signals from the ECM in the control of cell fate, cell shape and cell polarisation, will advance our understanding of how the gut epithelium develops and functions under healthy conditions, and how this is perturbed in diseased states such as cancer.

REFERENCES

Aghajanian, P., Takashima, S., Paul, M., Younossi-Hartenstein, A., and Hartenstein, V. (2016). Metamorphosis of the Drosophila Visceral Musculature and Its Role in Intestinal Morphogenesis and Stem Cell Formation. Dev. Biol. 420, 43–59. doi:10.1016/j.ydbio.2016.10.011

Antonello, Z. A., Reiff, T., Ballesta-Illan, E., and Dominguez, M. (2015). Robust Intestinal Homeostasis Relies on Cellular Plasticity in Enteroblasts Mediated by miR-8-Escargot Switch. Embryo J. 34, 2025–2041. doi:10.1252/embj.201591517

Baumann, O. (2001). Posterior Midgut Epithelial Cells Differ in Their Organization of the Membrane Skeleton from Other drosophila Epithelia. Exp. Cell. Res. 270, 176–187. doi:10.1006/excr.2001.5343

Beehler-Evans, R., and Micchelli, C. A. (2015). Generation of Enteroendocrine Cell Diversity in Midgut Stem Cell Lineages. Dev. Camb. 142, 654–664. doi:10.1242/dev.114959

Bilder, D., and Scott, M. P. (1995). Genomic Regions Required for Morphogenesis of the Drosophila Embryonic Midgut. Genetics 141, 1087–1100. doi:10.1093/ genetics/141.c.1087

Billingesley, P. F., and Lehane, M. J. (1996). "Structure and Ultrastructure of the Insect Midgut," in Biology of the Insect Midgut (Dordrecht: Springer Netherlands), 3–30. doi:10.1007/978-94-009-1519-0_1

Bosveld, F., Wang, Z., and Bellaiçi, Y. (2018). Tricellular Junctions: a Hot Corner of Drosophila Epithelial Cell Diversity in Midgut Stem Cell Lineages. Dev. Camb. 142, 654–664. doi:10.1242/dev.114959

NOTE#1

By the time of submitting this review paper, these two research papers (Chen and St Johnston, 2022; Moreno-Roman et al., 2022) are still in the peer-reviewed stage for publishing. The citations are referring to the versions published on bioRxiv.org.

AUTHOR CONTRIBUTIONS

JC made the figures, JC and DStJ wrote the manuscript.

FUNDING

This work was supported by a Wellcome Principal Fellowship (207496) to DStJ and core funding from the Wellcome Trust (203144) and Cancer Research United Kingdom (A24843). JC was supported by a Royal Society K.C. Wong Postdoctoral Fellowship.

ACKNOWLEDGMENTS

We would like to thank all past and present members of the St Johnston laboratory for their advice and support.
Yuan, W., Datta, A., Leroy, P., O’Brien, L. E., Mak, G., Jou, T.-S., et al. (2005). β1-Integrin Orient Epithelial Polarity via Rac1 and Laminin. *MboC* 16, 433–445. doi:10.1091/mbc.E04-05-0435

Zeng, X., Chauban, C., and Hou, S. X. (2013). Stem Cells in the Drosophila Digestive System. *Adv. Exp. Med. Biol.* 63, 74–83. doi:10.1007/978-94-007-6621-1_5

Zhai, Z., Boquete, J.-P., and Lemaître, B. (2017). A Genetic Framework Controlling the Differentiation of Intestinal Stem Cells during Replication in Drosophila. *PLoS Genet.* 13, e1006854. doi:10.1371/journal.pgen.1006854

Zhai, Z., Kondo, S., Ha, N., Boquete, J.-P., Brunner, M., Ueda, R., et al. (2015). Accumulation of Differentiating Intestinal Stem Cells Drives Tumorogenesis. *Nat. Commun.* 6, 10219. doi:10.1038/ncomms10219

Zhang, L., Turner, B., Ribbeck, K., and ten Hagen, K. G. (2017). Loss of the Mucosal Barrier Alters the Progenitor Cell Niche via Janus Kinase/signal Transducer and Activator of Transcription (JAK/STAT) Signaling. *J. Biol. Chem.* 292, 21231–21242. doi:10.1074/jbc.M117.809848

**Conflict of Interest:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

**Publisher’s Note:** All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

**Copyright © 2022 Chen and St Johnston. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.**