Microbiota, gut physiology and insect immunity

1 Gut structure and function

Adult *Drosophila* contains three distinct domains: foregut, midgut and hindgut. Foregut is at the anterior-most region and is originated from ectoderm. It includes pharynx and oesophagus for passage of ingested food and crop for food storage. Midgut is the region from cardia to midgut-hindgut junction where Malpighian tubules are attached. It is originated from endoderm and functions for food digestion and nutrient absorption. Cardia serves as a valve for food passage regulation. Hindgut is ectoderm-derived, extending to the rectum, and is responsible for absorption of water and ions.

Midgut is a single layer of epithelium and a visceral of muscle layer. The midgut epithelium contains four cell types: enterocytes (ECs), enteroendocrine cells (EECs), ISCs and enteroblasts (EBs). ECs are large polypoid cells secreting digestive enzymes and absorbing nutrients. They are the most abundant in midgut epithelium. EECs secrets hormones. ISCs are dividing progenitor cells. EBs are restricted progenitor cells produced by ISCs differentiation, and further differentiates into ECs or EECs. The lumenal side of midgut is covered by peritrophic matrix, a chitin polymer layer. A mucus layer fills between the epithelium and peritrophic matrix. The peritrophic matrix, mucus layer and epithelium act as physical barrier for immunity.

The midgut is further regionalized into anterior region, copper cell region (CCR) and posterior region. Anterior region functions for food breakdown by secreting enzymes. CCR is for further digestion with its low pH. Posterior region is for absorption of nutrients. When radius of gut is measured, midgut can be divided into six regions (R0-R5). R0 is cardia, R1-R2 is anterior region, R3 is CCR, and R4-R5 is posterior region.

2 Gut immunity

The primary immune systems in midgut of *Drosophila* are DUOX pathway (Ha *et al.*, 2009) and IMD pathway (Tzou *et al.*, 2000). Toll pathway is dispensable in gut epithelium.

2.1 IMD pathway

IMD pathway includes: (1) recognition of bacterial peptidoglycans; (2) intracellular cascade activating Relish, a member of NF-κB transcription factor family; (3) expression of antimicrobial peptides; (4) negative regulation of IMD pathway.

IMD pathway begins with PGRPs that recognize peptidoglycans. PGRP-LC is a transmembrane receptor recognizes DAP-type peptidoglycan characterized by meso-diaminopimelic acid in peptide chain (Choe *et al.*, 2002; Gottar *et al.*, 2002; Ramet *et al.*, 2002). PGRP-LE resides in cytoplasm and recognizes DAP-type peptidoglycan, and thus activates IMD pathway (Bosco-Drayon *et al.*, 2012).

After binding to peptidoglycan, PGRP-LC recruits IMD, Dredd and FADD to form a signaling complex (Georgel *et al.*, 2001; Naitza *et al.*, 2002). Dreed cleaves IMD and Relish for their activation. Ultimately, N-terminal cleaved Relish translocates into nucleus for target gene expression (Khush *et al.*, 2001).

Ubquitination and phosphorylation is required for IMD activation. Dredd activation requires K63-ubiquitination by IAP2, a E3-ligase (Meinander *et al.*, 2012). Dredd cleaves IMD, enabling its binding with IAP2. IAP2 generates K63-polyubiquitination, which is required for recruitment of TAK1/TAB2 complex (Vidal *et al.*, 2001). TAB2 binds to K63-polyubiquitination of IMD, and TAK1 is a MAPKKK kinase for activation of IKK complex. IKK complex is composed of IRD5 (catalytic activity) and Kenny (regulatory subunit). Activated IKK complex phosphorylates Relish on multiple sites, which activates its transcription factor activity (Erturk-Hasdemir *et al.*, 2009; Silverman *et al.*, 2000). Relish induces expression of genes involved in non-self recognition, signaling pathways, proteolysis and antimicrobial peptides.

IMD pathway is inhibited by several mechanisms. PGRP amidase (PGRP-LB, -SC1a, -SC1b, -SC2) degrades peptidoglycan and thus inhibits IMD pathway (Bischoff *et al.*, 2006; Guo *et al.*, 2014; Paredes *et al.*, 2011). PIRK is a transcriptional target of IMD pathway and inhibits IMD pathway (Aggarwal *et al.*, 2008; Kleino *et al.*, 2008; Lhocine *et al.*, 2008). It may disrupt IMD signaling as it interacts with PGRP-LC, PGRP-LE and IMD (Aggarwal *et al.*, 2008). Other inhibitors of IMD pathway including Dnr1 for Dredd inhibition (Foley and O'Farrell, 2004; Guntermann *et al.*, 2009); Caspar for Dredd-dependent Relish cleavage inhibition (Kim *et al.*, 2006); Trabid targeting TAK1 (Fernando *et al.*, 2014); CYLD, a deubiquitinating enzyme (Tsichritzis *et al.*, 2007); SkpA, a subunit of SCF-E3 ubquitin ligase targeting Relish (Khush *et al.*, 2002); and transcription inhibitors such as caudal (Ryu *et al.*, 2008) and Nubbin (Dantoft *et al.*, 2013).

IMD pathway is inhibited in gut, and its activation leads to pathologic symptoms including mocrobiota dysbiosis and dysplasia (Bosco-Drayon *et al.*, 2012; Guo *et al.*, 2014; Lhocine *et al.*, 2008; Ryu *et al.*, 2008). For instance, caudal is gut-specific inhibitor of IMD pathway and its knockdown causes gut cell

apoptosis, decreased survival rate and change of microbiome (Ryu *et al.*, 2008). Knockdown of PGRP-SC2, an inhibitor of IMD pathway, also leads to mocrobiota dysbiosis and dysplasia (Guo *et al.*, 2014).

2.2 DUOX pathway

DUOX is a member of nicotinamide adenine dinucleotide phosphate oxidase (NOX) family and is responsibe for bacterial-induced reactive oxygen species (ROS) generation. ROS plays an important role in gut immunity, and is degraded by secretory immune-related catalase (IRC) (Ha *et al.*, 2005).

NOX/DUOX family proteins share a catalytic gp91phox domain, and DUOX contains an additional peroxidase homology domain (PHD). Generally, NOX generates superoxide anion in extracellular space by electron transfer from NADPH in cytoplasm to oxygen across the membrane. Superoxide anion is subsequently converted to H_2O_2, which can be further converted to HClO by myeloperoxidase.

In *Drosophila*, there are one NOX and one DUOX (Ha *et al.*, 2005). The produce of HClO is DUOX-dependent (Ha *et al.*, 2005). DUOX is essential for gut immunity (Ha *et al*, 2005; Ha *et al*, 2009). It is activated only with transient microorganisms, but not with commensals. Uracil acts as a ligand for DUOX activation (Lee *et al.*, 2013). It is secreted by several pathogens, but not by commensals.

DUOX enzymatic activity requires calcium released from ER, and thus is regulated by PLCβ and Gαq (Ha *et al.*, 2005; Ha *et al.*, 2009). At downstream of Gαq, PLCβ is required for generation of 1,4,5-triphosphate, which is recognized by corresponding receptor and enables release of calcium from ER. Transcription of DUOX is up-regulated by (Ha *et al.*, 2009) (1) peptidoglycan-dependent cascade composed of PGRP-LC, IMD, MEKK1, MKK3, p38 and ATF2; (2) uracil-dependent cascade including PLCβ, MEKK1, MKK3, p38 and ATF2. ATF2 is a transcription factor.

Negative regulation of DUOX transcription is mediated by inhibition of peptidoglycan-dependent p38 activation, which requires PLCβ, calcineurin B and MAP kinase phosphatase 3 (MKP3) (Ha *et al.*, 2009). It indicates that activation of DUOX requires certain amount of peptidoglycan, and therefore, DUOX remains inactivated under commensals.

3 Gut renewal

Mid gut is dynamic. Adult *Drosophila* intestinal epithelium is renewed every 1 week (Micchelli and Perrimon, 2006). This gut renewal is dependent on asymmetric division of ISC. The two daughter cells of ISC division, one becomes self-renewed ISC and another one differentiates into EB, which further differentiates into EC or EEC. The fate decision of ISCs after division is dependent on the antagonism of Delta-Notch signaling and BMP signaling (Tian and Jiang *et al.*, 2014). Delta-Notch signaling also plays an important role in differentiation into EC/EEC (Ohlstein and Spradling, 2007; Perdigoto *et al.*, 2011).

Proliferation of ISC is under tightly control to maintain gut homeostasis. Low rate of ISC proliferation leading to reduced replacement of damaged cells, destroying gut integrity and leads to originasm death. High rate of ISC proliferation leads to accumulation of unwanted cells, causing pathology (Biteau *et al.*, 2008). Several signaling pathways are involved in ISC proliferation activation, including JAK/STAT, EGFR, Hippo, JNK and Wingless (Biteau *et al.*, 2008; Cordero *et al.*, 2012; Jiang *et al.*, 2011; Karpowicz *et al.*, 2010; Lee *et al.*, 2009). Myc may be a common downstream of JAK/STAT, EGFR, Hippo and Wingless (Ren *et al.*, 2013). Insulin receptor signaling in ISC is required for ISC proliferation (Amcheslavsky *et al.*, 2009).

Ligands from nearby injured cells are responsible for activation of ISC proliferation. In stressed ECs, JAK/STAT ligand Upd3 and EGFR ligand Keren are expressed under control of JNK and Hippo signaling (Jiang *et al.*, 2009; Ren *et al.*, 2010; Jiang *et al.*, 2011). EGFR ligands vein and spitz are produced by visceral muscles and progenitors respectively (Jiang *et al.*, 2011). Stressed EBs produce Upd2 through activation of Hedgehog pathway (Tian *et al.*, 2015), and Wingless ligand under control of JNK signaling (Cordero *et al.*, 2012). Hippo signaling in ISC is controled by intracellular interactions of two cadherins, Fat in ISC and Dachsous (DS) in EC (Karpowicz *et al.*, 2010).

IMD pathway may regulate ISC proliferation via controling number of gut bacteria (Buchon *et al.*, 2009). ROS induced by DUOX signaling also accerlates ISC proliferation. ROS may induce ISC proliferation by tissue damaging (Buchon *et al.*, 2009; Karpowicz *et al.*, 2010; Ren *et al.*, 2010; Ren *et al.*, 2013; Shaw *et al.*, 2010; Staley and Irvine, 2010). ROS may also activates ISC proliferation directly by targeting redox-sensitive components of signalings. ROS activates JAK/STAT by redox-sensitive tyrosine phosphatase (Liu *et al.*, 2004), JNK by thioredoxin (Junn *et al.*, 2000) and Wnt by nucleoredoxin (Funato *et al.*, 2006).