Single-cell analysis of mosquito hemocytes identifies signatures of immune cell subtypes and cell differentiation

Hyeogsun Kwon[1], Mubasher Mohammed[2], Oscar Franzén[3], Johan Ankarklev[2,4], Ryan C. Smith[1]*

[1]Department of Entomology, Iowa State University, Ames, Iowa, USA

[2]Department of Molecular Biosciences, The Wenner-Gren Institute, Stockholm University, Stockholm, Sweden 10691

[3]Integrated Cardio Metabolic Centre, Department of Medicine, Karolinska Institutet, Novum, 14157 Huddinge, Sweden

[4]Microbial Single Cell Genomics facility, SciLifeLab, Biomedical Center (BMC) Uppsala University, SE-751 23, Uppsala, Sweden

*Corresponding author: smithr@iastate.edu

Abstract

Mosquito immune cells, known as hemocytes, are integral to cellular and humoral responses that limit pathogen survival and mediate immune priming. However, without reliable cell markers and genetic tools, studies of mosquito immune cells have been limited to morphological observations, leaving several aspects of their biology uncharacterized. Here, we use single-cell RNA sequencing (scRNA-seq) to characterize mosquito immune cells, demonstrating an increased complexity to previously defined prohemocyte, oenocytoid, and granulocyte subtypes. Through functional assays relying on phagocytosis, phagocyte depletion, and RNA-FISH experiments, we define markers to accurately distinguish immune cell subtypes and provide evidence for immune cell maturation and differentiation. In addition, gene-silencing experiments demonstrate the importance of lozenge in defining the mosquito oenocytoid cell fate. Together, our scRNA-seq analysis provides an important foundation for studies of mosquito immune cell biology and a valuable resource for comparative invertebrate immunology.

Introduction

Across metazoa, immune cells are vital to promoting wound healing, maintaining homeostasis, and providing anti-pathogen defenses (Chaplin, 2010). With immune cells mediating both innate and adaptive immune function in vertebrates, immune cell subtypes display highly specialized roles that have continually been resolved by technological advancements that enable their study (Papalexi and Satija, 2017; Proserpio and Mahata, 2016). Recently, the advent of single-cell sequencing (scRNA-seq) has continued to delineate and provide further resolution into new cell types and immune cell functions in mammals (Szabo et al., 2019; Villani et al., 2017). In lesser studied invertebrates lacking adaptive immunity, single-cell technologies have enhanced descriptions of previously described cell types and have redefined cell complexity (Cattenoz et al., 2020; Cho et al., 2020; Raddi et al., 2020; Severo et al., 2018; Tattikota et al., 2020).

In insects, hematopoiesis and immune cell function have predominantly been examined in Lepidoptera and *Drosophila* (Banerjee et al., 2019; Lavine and Strand, 2002), with the mosquito, *Anopheles gambiae*, recently serving as an emerging system of study (Kwon and Smith, 2019). Mosquito immune cells have proven integral to the cellular and humoral responses that limit invading pathogens in the mosquito host (Baton et al., 2009; Castillo et al., 2017; Hillyer et al., 2003; Kwon and Smith, 2019; Smith et al., 2016) and the establishment of immune memory (Ramirez et al., 2015; Rodrigues et al., 2010). Transcriptional (Baton et al., 2009; Pinto et al., 2009) and proteomic (Smith et al., 2016) analysis of mosquito hemocyte populations have yielded important information into the regulation of hemocyte function in response to blood-feeding and infection. However, the study of mosquito immune cells (or hemocytes) has been complicated by discrepancies in cell classification (Ribeiro and Brehélin, 2006), cell numbers (Hillyer and Strand, 2014), and methodologies to examine their function (Kwon and Smith, 2019). These constraints are magnified by the lack of genetic tools and markers that have limited studies of immune cells outside of *Drosophila* strictly to morphological properties of size and shape (Hillyer and Strand, 2014; Kwon and Smith, 2019).

Traditional classifications of mosquito hemocytes describe three cell types: prohemocyte precursors, phagocytic granulocytes, and oenocytoids that have primary roles in

melanization (Castillo et al., 2006). However, recent studies have begun to challenge these traditional immune cell classifications, demonstrating the existence of multiple types of phagocytic cells (Kwon and Smith, 2019; Severo et al., 2018) and that both granulocytes and oenocytoids contribute to prophenoloxidase expression (Bryant and Michel, 2016; Kwon et al., 2020; Kwon and Smith, 2019; Severo et al., 2018; Smith et al., 2016). Together, this suggests that there is additional complexity to mosquito immune cells that are not accurately represented by the traditional classification of mosquito hemocytes into three cell types.

For this reason, here we employ the use of scRNA-seq to characterize mosquito immune cells. Using a conservative approach, we identify seven hemocyte subtypes with distinct molecular signatures and validate these characterizations using a variety of bioinformatic and experimental molecular techniques. We define new markers that can accurately distinguish immune cell subtypes, improving upon the ambiguity of existing methodologies. Moreover, our data support a new model of immune cell differentiation and maturation that leads to a dynamic population of circulating immune cells in the adult female mosquito. In summary, these data represent a valuable resource to advance the study of mosquito immune cells, offering a dynamic data set for comparative immunology with other insect systems.

Results

Isolation of mosquito immune cells and scRNA-seq analysis

To examine mosquito immune cells by scRNA-seq, adult female *An. gambiae* were perfused as previously (Kwon et al., 2017; Kwon and Smith, 2019; Reynolds et al., 2020; Smith et al., 2016, 2015) from either naïve or blood-fed (24 h post-feeding) conditions to assess the physiological impacts of blood-feeding on hemocyte populations as previously suggested (Bryant and Michel, 2016, 2014; Castillo et al., 2011; Reynolds et al., 2020). Following perfusion, cells were stained with a live-dead viability stain to select for live cells, with mosquito immune cells distinguished by labeling with FITC-conjugated wheat germ agglutinin (WGA) as a general hemocyte marker and the far-red stain DRAQ5 to label DNA content as previously (Kwon and Smith, 2019). Based on consistent patterns of WGA/DRAQ5 signal intensity that were suggestive that these labeling properties could

distinguish distinct groups of immune cells (Figure 1A, Figure S1), we isolated individual cells by fluorescence-activated cell sorting (FACS) using three “gates” to enrich for defined cell populations using these WGA/DRAQ5 properties (Figure 1A, Figure S1). An additional, non-selective fourth gate isolated cells at random to achieve an unbiased cell population that would be influenced by overall cell abundance (Figure 1A). Based on these parameters, individual cells were isolated by FACS into a 384-well plate for further processing for scRNA-seq using the SMART-seq2 methodology (Picelli et al., 2014, 2013). A total of 262 cells passed the quality filtering threshold of 10,000 reads per cell (Figure S2), yielding 194 and 68 cells, respectively from naïve and blood-fed conditions (Table S1). Overall, we detected expression (>0.1 RPKM) from ~46% (6352/13,764) of the *An. gambiae* genome, with a median of 1646 genes expressed per cell (range 45 to 5215). In comparison to Severo *et al.* (Severo et al., 2018), we see a comparable number of detected genes in our analysis. However, our data display a higher number of genes per cell and larger variance in genes between cell types, patterns suggestive of a broader range of cell populations represented in our dataset. All immune cell data can be visualized and searched using the following database: https://alona.panglaodb.se/results.html?job=2c2r1NM5Zl2qcW44RSrjkHf3Oyv51y_5f09d74b770c9.

Using hierarchical clustering, we conservatively define eight distinct cell clusters or immune cell-subtypes (Figure 1B). These clusters are supported by the unique molecular profiles of each cell cluster when analyzed by tSNE (Figure 1C), as well as variability in the number of expressed genes (Figure 1D) that infers functional heterogeneity in these immune cell populations. When referenced to our FACS gating methodology based on WGA/DRAQ5 staining (Figure 1A, Figure S1), each cell cluster is represented in our non-selective gating conditions (Gate 4), while each of the specific gating conditions (Gates 1-3; Figure S1) provide enrichment for distinct cell types. Clusters 1, 7, and 8 are enriched in Gate 1, Clusters 5 and 6 in Gate 2, and Clusters 2-4 in Gate 3 (Figure S1), arguing that these cells have similar physical properties in WGA staining and DNA content. While at present, it is unclear what defines these properties on the molecular level, the enrichment achieved through our gating strategy provides further support for our FACS methodology. With the exception of Cluster 3, which was only identified in naïve mosquitoes, each of

the respective cell clusters were found under both naïve and blood-fed conditions (Figure S3). When paired with differential gene expression between naïve and blood-fed cells of each cluster, only Clusters 2 and 4 displayed significant changes in gene expression (Table S2). This suggests that blood-feeding may primarily influence the activation state and gene expression of specific immune cell subtypes as previously suggested (Bryant and Michel, 2016, 2014; Reynolds et al., 2020; Smith et al., 2016).

Characterization of *An. gambiae* immune cell clusters

To further characterize the cell clusters resulting from our scRNA-seq analysis, we used the Seurat package (Butler et al., 2018) to identify transcriptional markers significantly enriched for each cell cluster (Figure 2A, Table S3). Unlike other cell clusters, Cluster 1 did not display specific markers unique to the cluster and expressed several markers (such as *LRIM26* and *SCRB9*) at high levels that otherwise define specific cell clusters (Figure 2A). When paired with its localization as an outlier in the t-SNE analysis (Figure 1C) and high median number of expressed genes (Figure 1D), we believe that these cells may represent cell doublets of mixed cell origins (granulocyte and oenocytoid) resulting from errors in the FACS isolation. Alternatively, these cells may also represent a granulocyte undergoing transdifferentiation into an oenocytoid as has been similarly suggested in *Drosophila* (Leitao and Sucena, 2015), yet in the absence of increased expression of cell cycle genes (Figure 2B) that have been implicated in dividing hemocytes (Raddi et al., 2020), this becomes a less likely scenario. Additional considerations that these cells are pluripotent precursors or represent recently described megacyte populations (Raddi et al., 2020) also seem unlikely given that pluripotent precursors have not been described in other insect single-cell studies (Cattenoz et al., 2020; Cho et al., 2020; Raddi et al., 2020; Severo et al., 2018; Tattikota et al., 2020) and that Cluster 1 cells do not display enriched expression of *TM7318* and *LL3* (Table S4) that are indicative of megacytes (Raddi et al., 2020). Without a defined expression pattern and the potential that these cells may be experimental artifacts, we have focused our efforts on defining the other cell clusters.

When we examine the expression of mitotic markers in other cell clusters, only cells of Cluster 2 display increased expression (Figure 2B), supporting that these cells may have

some capacity for proliferation as previously suggested (King and Hillyer, 2013). To more closely evaluate the molecular profiles of each cell cluster, transcripts identified in more than 80% of cells in each cluster (Table S5) were used to perform gene ontology (GO) analysis (Figure 2C). Comparisons across cell clusters provide further support that Clusters 2-4 are highly analogous in their core machinery, with Cluster 6 displaying a related, yet divergent cellular composition represented by an increased representation of transcripts involved in translation (Figure 2C). Correlations with a previous proteomics study of phagocytic granulocytes in *An. gambiae* (Smith et al., 2016) demonstrate that transcripts of Clusters 2-4 have the strongest associations with phagocytic immune cells (Figure S4), providing support that these clusters represent populations of phagocytic granulocytes. Cells in Cluster 5 display a unique profile predominantly comprised of genes implicated in redox metabolism/stress responses (Figure 2C), while Clusters 7 and 8 display marked differences in composition (Figure 2C), despite sharing similar markers to delineate these cell types (Figure 2A).

In order to further determine the classifications of our immune cell clusters, we examined the expression of well-characterized *Drosophila* hemocyte gene orthologs (Dudzic et al., 2015; Evans et al., 2014; Fossett et al., 2003; Franc et al., 1996; Kocks et al., 2005; Manaka et al., 2004; Martinek et al., 2011; Stofanko et al., 2008; Waltzer et al., 2003) in our dataset (Figure 2D). SPARC and Cg25C were expressed at high levels across cell clusters, suggesting that these could be universally expressed markers of mosquito immune cells (Figure 2D). Clusters 2-4 and 6 express the *Drosophila* plasmatocyte markers (equivalent to mosquito granulocytes) *peroxidasin* (pxn) and *eater* (Figure 2D) suggestive of phagocytic cell function. In contrast, the expression of *lozenge* (lz) and *PPO1* indicative of *Drosophila* crystal cells (Cattenoz et al., 2020; Dudzic et al., 2015; Evans et al., 2014; Fossett et al., 2003; Tattikota et al., 2020) were most prevalent in Clusters 7 and 8 (Figure 2D), suggesting that these cell clusters are representative of mosquito oenocytoid populations (equivalent to crystal cells). However, little resolution into the role of Cluster 5 was provided through these comparisons to known *Drosophila* markers (Figure 2D).

Similar classifications were performed for described mosquito hemocyte genes (Bryant and Michel, 2016; Castillo et al., 2006; Danielli et al., 2000; Estévez-Lao and Hillyer, 2014; Kwon and Smith, 2019; Midega et al., 2013; Pinto et al., 2009; Raddi et al., 2020; Severo et al., 2018; Smith et al., 2015, 2016) to provide additional resolution into our mosquito immune cell clusters (Figure 2E). As previously suggested (Kwon and Smith, 2019), the expression of NimB2 and PPO6 are universal markers of mosquito immune cells (Figure 2E). Through the use of described granulocyte markers (Danielli et al., 2003; Raddi et al., 2020; Smith et al., 2016), we are able to describe two discrete phagocytic cell types (Figure 2E). Those cells (Clusters 2-4, 6) that express LRIM16A, LRR8, and LYS I (Raddi et al., 2020; Severo et al., 2018; Smith et al., 2016), distinguished by more specialized granulocyte populations (Clusters 2-4) that generally express SCRASP1, SRPN10, and LRIM15 (Danielli et al., 2003; Smith et al., 2016). Other markers such as *Ninjirin* (Pinto et al., 2009) and *DOX-A2* (Castillo et al., 2006) expressed in Clusters 2 and 4 further delineate these presumed phagocytic cell types (Figure 2E). In addition, *AGAP007314* grouped strongly with the presumed oenocytoid cell population of Cluster 8, supporting its previously described roles in melanization (Pinto et al., 2009).

The patterns of *LysI* (AGAP011119) and *FBN10* (AGAP011230) used to respectively define PPO6^{low} and PPO6^{high} immune cell populations in previous scRNA-seq studies (Severo et al., 2018) also provide significant comparative insight into the immune cell populations defined in our study. We demonstrate that *LysI* is predominantly expressed in the phagocytic granulocyte populations of Clusters 2-4, while *FBN10* can be found in both tentative granulocyte and oenocytoid populations (Figure 2E, Figure S5). There is a significant correlation of *LysI* and *FBN10* with *PPO6* expression (Figure S5), mirroring the PPO6^{low} and PPO6^{high} phenotypes as previously described (Severo et al., 2018), yet our data argue that these cell markers do not accurately account for the added complexity of mosquito immune cell populations identified in our study.

When we examine the transcriptional profiles of prophenoloxidases (PPOs), a family of enzymes that catalyze the production of melanin in response to infection (Dudzic et al., 2015), we demonstrate that the eight PPOs detected in our analysis are expressed in each of the major immune cell subtypes (Figure 2F). As previously suggested (Kwon and

Smith, 2019; Severo et al., 2018), *PPO6* is universally expressed in all hemocytes (Figure 2E and 2F). *PPO2*, *PPO4*, *PPO5*, and *PPO9* are most abundant in putative granulocyte populations, while *PPO1*, *PPO3*, and *PPO8* are enriched in putative oenocytoids (Figure 2F). This directly contrasts previous suggestions that mosquito PPOs are only constitutively expressed in oenocytoid populations (Castillo et al., 2006; Hillyer and Strand, 2014; Strand, 2008), yet is supported by recent evidence that phagocytic granulocyte populations in mosquitoes significantly contribute to *PPO* production (Kwon and Smith, 2019; Smith et al., 2016). Furthermore, the enriched expression of *PPO1*, *PPO8*, and (to lesser extent) *PPO3* in oenocytoids is supported by recent studies examining prostaglandin signaling on *PPO* expression in mosquito oenocytoid populations (Kwon et al., 2020).

Additional characterizations of immune signaling pathways (Figure S6), SRPNs and CLIPs (Figure S7), chemosensory receptors/proteins (Figure S8), and tRNA expression (Figure S9) across cell clusters provide further detail into the functions of our tentative immune cell clusters. We demonstrate that known anti-microbial genes and signaling components of the Toll, IMD, and JAK-STAT pathways (Cirimotich et al., 2010) display the highest expression in Clusters 2 and 8 (Figure S6), similar to the expression patterns of SRPNs and CLIPs (Figure S7) that mediate immune activation (Gulley et al., 2013; Kanost and Jiang, 2015). This suggests that mosquito granulocyte and oenocytoid populations each contribute to the expression of a distinct subset of immune signaling processes. However, at present, it is unclear if this corresponds to pathogen-specific defenses or immune responses unique to each particular cell type. Interestingly, receptors involved in chemosensory recognition (ionotropic, gustatory, and odorant receptors; odorant binding proteins) are highly expressed in Clusters 7 and 8 (Figure S8). Although their function has not been described in mosquitoes, the role of odorant binding proteins on immune system development has been described previously in other insect systems (Benoit et al., 2017). Moreover, the differential expression of transfer RNA (tRNA) genes across cell populations provided useful measures to tease apart Clusters 5-8 from other cell clusters (Figure S9), potentially representing different activation states or stages of immune cell development as previously defined in mammalian systems (Krishna et al., 2019; Rak et al., 2020; Torrent et al., 2018).

Based on these characterizations, our data support the identification of NimB2, SPARC, and PPO6 as universal markers of mosquito immune cell populations that can be found in each of our cell clusters (Figure 2G, Figure S10). Given the low number of expressed genes (Figure 1), the lack of discernable markers (Figure 2, Figure S10), and low levels of cyclin G2 (Table S4) that define differentiated cell populations (Horne et al., 1997; Martínez-Gac et al., 2004), we believe that Cluster 5 represents a progenitor population of prohemocytes (Figure 2G). Clusters 2-4 and 6 can be described as granulocytes, distinguished by LRIM16A and LRR8 (Raddi et al., 2020; Smith et al., 2016), and further delineated as "mature" granulocyte populations in Clusters 2-4 marked by the expression of LRIM15 and SCRASP1 (Figure 2G, Figure S10). In the absence of these additional phagocytic markers, we believe that the less defined populations of Cluster 6 likely represent immature granulocytes. Clusters 7 and 8 represent populations of oenocytoids that can easily be denoted by the expression of two scavenger receptors, SCRB9 and SCRB3 (Figure 2G, Figure S10).

Confirmation of mosquito immune cell subtypes

To confirm the identification of our immune cell clusters and to establish a reliable set of markers to distinguish immune cell subtypes, we performed RNA-FISH on fixed immune cell populations and paired these observations with the phagocytic properties of each of the respective cell populations (Figure 3). Supported by our expression data, the "universal" marker *Nimrod B2* (NimB2) labeled all hemocytes (Figure 3A). Serving as a marker for phagocytic cells, we demonstrate that *LRIM15* effectively labels phagocytic cell populations (Figure 3B), while the labeling of *SCRB9* (Figure 3C, Figure S11) and *SCRB3* denote mosquito oenocytoid populations (Figure S12). As further validation, RNA-FISH experiments performed with both *LRIM15* and *SCRB9* probes identify distinct populations of *LRIM15*$^{+}$ or *SCRB9*$^{+}$ cells, confirming that these markers label unique cell populations (Figure S13).

When these respective RNA-FISH markers are used to examine cell abundance, >90% of fixed cells are *NimB2*$^{+}$ (Figure 3D), further demonstrating its role as a reliable general cell marker. *LRIM15*$^{+}$ phagocytic granulocyte populations represent ~60% of fixed cells, while only ~5% of cells are *SCRB9*$^{+}$ (Figure 3D). As expected, *LRIM15*$^{+}$ cells are

phagocytic, while *SCRB9*$^{+}$ cells do not display phagocytic activity (Figure 3E), agreeing with the respective phagocytic and non-phagocytic roles of mosquito granulocytes and oenocytoids. Moreover, when phagocytic activity was compared between *LRIM15*$^{+}$ and *LRIM15*$^{-}$ phagocytes, *LRIM15*$^{+}$ cells displayed significantly higher phagocytic activity (Figure 3F). These data also indirectly support that the *LRIM15*$^{-}$ phagocytic cells are likely those of the "immature" granulocytes of Cluster 6. Additional experiments using clodronate liposomes to deplete phagocytic cell populations (Kwon and Smith, 2019) demonstrate that *NimB2*$^{+}$/*LRIM15*$^{+}$ cells are highly susceptible to clodronate treatment (Figure 3G), providing further confirmation of their phagocytic cell function. The specificity of clodronate treatment was further validated by qRT-PCR, demonstrating that the expression of "universal" or "granulocyte" transcripts were reduced following phagocyte depletion, while "oenocytoid" markers remain unaffected (Figure 3H). Together, these data confirm the identification of our mosquito immune cell clusters and define the use of specific cell markers to delineate granulocyte and oenocytoid populations in *An. gambiae*.

However, ~30% of cells (most are *NimB2*$^{+}$) cannot be fully resolved by the expression of *LRIM15* and *SCRB9* alone (Figures 3B and 3C). This is evidenced by the *NimB2*$^{+}$/*LRIM15*$^{-}$ cells displayed in Figure 3F, which may represent other adherent cell populations with "granulocyte-like" morphology, potentially belonging to cells of Clusters 5 or 6 that are *LRIM15*$^{-}$ (Figure 3B).

Further defining mosquito granulocyte and oenocytoid sub-populations

Based on the initial comparisons of known phagocytic cell markers (Figure 2) and the confirmation of phagocytic activity in *LRIM15*$^{+}$ cells (Figure 3), our data conservatively support the presence of four granulocyte subtypes (Clusters 2, 3, 4, and 6; Figure 2, Figure S10). To better define these subtypes, we more closely investigated the putative functional roles of each of these cell populations. Cells of Cluster 2 display high levels of immune gene expression of anti-microbial peptides (AMPs), components of the Toll pathway, *TEP1*, *MMP1*, and *LRIM26* (Figure S14) that may represent a class of specialized immune cells similar to other recent studies (Cattenoz et al., 2020; Raddi et al., 2020; Tattikota et al., 2020). However, tentative cell functions for the other cell clusters are less transparent. Cells of Cluster 3 have little immune gene expression and are

distinguished by the increased production of *LRIM6*, *cathepsin L*, and *cathepsin F*, while cells in Cluster 4 display high levels of *FBN 8*, *FBN 10*, and multiple PPO genes (Figure S14). By contrast, cells in Cluster 6 display reduced expression of several phagocytic markers (Figure S14), suggesting that cells of this subtype lack the specialized phagocytic function of fully differentiated granulocytes. The reduced expression of *Cyclin G2* (Figure S14), a marker of differentiated cells (Horne et al., 1997; Martínez-Gac et al., 2004), supports this hypothesis. Together, these data support that Cluster 6 likely represents a granulocyte precursor, whereas Clusters 2-4 are differentiated subtypes with unique cell functions.

Similarly, differences in gene expression were also used to distinguish between the two oenocytoid subtypes. While both cell clusters express a subset of genes unique to the oenocytoid lineage (Figure 2), expression of most oenocytoid markers, such as *PPO1*, *SCRB3* and *SCRB9*, are higher in Cluster 8 (Figure S14). Moreover, Cluster 8 also expresses high levels of *hnt/peb*, *DnaJ-1*, *Mlf*, *klu*, and *lozenge* (Figure S14) that are indicative of mature *Drosophila* crystal cells (Koranteng et al., 2020; Miller et al., 2017; Tattikota et al., 2020; Terriente-Felix et al., 2013), suggesting that Clusters 7 and 8 respectively represent populations of immature and mature oenocytoids. This is further supported by reduced levels of *Cyclin G2* in cells of Cluster 7 (Figure S14), characteristic of a less differentiated cell type (Horne et al., 1997; Martínez-Gac et al., 2004).

Differentiation of mosquito immune cell populations using lineage analysis

Previous studies in mosquitoes have suggested that prohemocyte precursors give rise to differentiated granulocyte and oenocytoid populations (Ramirez et al., 2014; Rodrigues et al., 2010; Smith et al., 2015). Additional evidence supports that granulocytes undergo mitosis to proliferate in response to infection (King and Hillyer, 2013). However, these observations have been based on morphological characterization, providing only speculation to the source of these immune cell populations. To better understand the origins of our identified immune cell clusters, we performed lineage analysis to determine relationships between the transcriptional profiles of individual cells to construct cell lineages in pseudotime using Monocle3 (Cao et al., 2019; Packer et al., 2019; Trapnell et al., 2014). Individual cells were analyzed in pseudotime to reveal two distinct lineages

from a shared precursor population (Figure 4A). When visualized by cell cluster (Figure 4B), these patterns support that the presumed prohemocyte precursors of Cluster 5 serve as the initial branching point for our cell lineages delineate into either granulocyte (Clusters 2, 3, 4, 6) or oenocytoid lineages (Clusters 7, 8). For the granulocyte lineage, precursor cells progress into cells of Cluster 6 before final maturation in the specialized granulocyte populations of Cluster 2-4 (Figure 4B), where at present, we are unable to provide further resolution into the differentiation of these cells without further detailed experiments. In contrast, for the oenocytoid lineage, Monocle3 analysis supports that precursor cells first differentiate into cells from Clusters 7 before maturation into the mature oenocytoid cells of Cluster 8 (Figure 4B). Based on these cell trajectories (Figures 4A and 4B) as well as transcriptional differences that likely define immature cell types (Figure S13), our results support a model for immune cell differentiation and the progression of cells within each lineage (Figure 4C). These data corroborate the differentiation of cells from a prohemocyte precursor as previously proposed (Ramirez et al., 2014; Rodrigues et al., 2010; Smith et al., 2015), while providing insight into the potential role of cell intermediates undergoing maturation before terminal differentiation of mosquito immune cell subtypes (Figure 4C) similar to those recently described in *Drosophila* (Tattikota et al., 2020).

Lozenge promotes differentiation of the oenocytoid lineage

While several genes have been described that promote *Drosophila* immune cell lineages (Evans et al., 2014), the factors that define mosquito immune cell lineages have not been described beyond the role of multiple immune signaling pathways that influence hemocyte differentiation in response to malaria parasite infection (Ramirez et al., 2014; Smith et al., 2015). To further explore the factors that determine mosquito immune cell lineages, we focused on oenocytoid differentiation and the role of lozenge. In *Drosophila*, *lozenge* (lz) expression is integral to defining crystal cell fate (Fossett et al., 2003; Waltzer et al., 2003), the equivalent of mosquito oenocytoids. To similarly examine the role of lozenge in mosquito oenocytoid development, we used RNA-FISH to demonstrate and confirm the expression of *lozenge* in mosquito immune cells (Figure 5A). *Lozenge* was detected in ~15% of fixed cells (Figure 5B), a much higher percentage than that of *SCRB9*$^{+}$ cells demarcating *An. gambiae* oenocytoids (Figure 3). When we more closely examined the

expression of *lozenge* and *SCRB9*, we see that co-localization of both markers only occurs in a subset of *lozenge*$^{+}$ cells (Figure 5C), suggesting that *lozenge* is expressed in other immune cell subtypes in addition to mosquito oenocytoid populations. This is supported by the expression of *lozenge* in other immune cell clusters (Figure 5A), the ability of a subset of lozenge+ cells to undergo phagocytosis, as well as the depletion of lozenge^{+} cells and lozenge expression following depletion of phagocytic cell populations (Figure S15). To evaluate the influence of *lozenge* on oenocytoid cell fate, we silenced *lozenge* expression by RNAi (Figure S16) and examined the co-localization of *LRIM15/SCRB9* by RNA-FISH. In *lozenge*-silenced mosquitoes, we see a significant decrease in *LRIM15*$^{-}$*/SCRB9*$^{+}$ cells (Figure 5D), suggesting that *lozenge* is integral to defining the oenocytoid lineage. This is further supported by the specific reduction of *PPO1/3/8* expression (Figure 5E), PPOs that are enriched in Clusters 7 and 8 corresponding to the oenocytoid cell fate (Figure 2F). Together, these data support that *lozenge* expression is an important driver of the oenocytoid lineage in *An. gambiae* (Figure 5F). Based on our cell trajectories proposed in Figure 4, it suggests lozenge^{+} prohemocytes promote the differentiation into an oenocytoid. However, the presence of *lozenge* in a subset of phagocytic granulocyte populations (Figure 5, Figure S15) may alternatively support a model of transdifferentiation in which oenocytoids can be derived from phagocytic granulocytes as previously proposed in *Drosophila* (Leitao and Sucena, 2015).

Discussion

Our understanding of mosquito immune cells has largely been shaped by studies in other insects (Banerjee et al., 2019; Lavine and Strand, 2002). From morphological observations of size and shape, three cell types have been described in mosquitoes: prohemocytes, oenocytoids, and granulocytes (Castillo et al., 2006). However, with the advent of additional molecular tools to study mosquito immune cell function, several studies have supported an increased complexity of hemocyte populations beyond these generalized cell subtype classifications (Bryant and Michel, 2016; Kwon and Smith, 2019; Pondeville et al., 2020; Raddi et al., 2020; Severo et al., 2018; Smith et al., 2016). Herein, we demonstrate through scRNA-seq experiments and additional molecular

characterization that there are at least seven conservatively defined immune cell populations in *An. gambiae*.

Similar to previous characterizations (Castillo et al., 2006), we identify prohemocyte, oenocytoid, and granulocyte populations in our RNA-seq analysis. However, in both the oenocytoid and granulocyte classifications, we identify multiple, distinct immune cell populations defined by developmental progression, activation state, or specialized immune function similar to those recently described in *Drosophila* (Cattenoz et al., 2020; Tattikota et al., 2020). Moreover, we provide an initial set of markers to accurately distinguish between mosquito immune cell populations using RNA-FISH and validate these identifications through co-localization experiments, phagocytosis experiments, and phagocyte depletion assays. Therefore, these experiments provide important and much-needed cell markers that reliably distinguish oenocytoid and granulocyte populations to advance the study of mosquito immune cells.

Through our scRNA-seq analysis, we identify at least four granulocyte subtypes in *An. gambiae* based on gene expression, previous proteomics studies (Smith et al., 2016), and phagocytic properties. This expansion of the general “granulocyte” classification is supported by previous morphological observations of granulocytes in mosquitoes (Kwon and Smith, 2019; Pondeville et al., 2020), and more recently by parallel scRNA-seq experiments in *Drosophila* (Cattenoz et al., 2020; Tattikota et al., 2020) and *An. gambiae* (Raddi et al., 2020). Of these four granulocyte subtypes, our data support that cells of Cluster 6 are intermediate or immature granulocyte forms that display distinct expression patterns from prohemocyte precursors, yet do not have the same properties of other granulocyte subtypes (Cluster 2-4). This is supported by our pseudotime lineage analysis, where these immature cells of Cluster 6 give rise to more specialized granulocyte populations, similar to that described for comparable immature plasmatocytes in *Drosophila* (Tattikota et al., 2020). This maturation in Clusters 2-4 includes the increased expression of the phagocytic cell markers *LRIM15* and *SCRASP1* (Smith et al., 2016) as well as Cyclin G2 as a marker of differentiated cells (Horne et al., 1997; Martínez-Gac et al., 2004) that result in more specialized granulocyte subtypes.

Of these “mature” granulocytes, Cluster 2 displays increased immune properties comparable to other recently described granulocyte or plasmatocyte populations in in *An. gambiae* (Raddi et al., 2020) and *Drosophila* (Cattenoz et al., 2020; Tattikota et al., 2020). Based on their increased expression of anti-microbial peptides (AMPs) and other immune components such as TEP1, these cell populations may have a primary role in the hemocyte-mediated immune responses that limit bacteria (Hillyer et al., 2003; Reynolds et al., 2020) or the recognition and killing of malaria parasites (Castillo et al., 2017; Kwon and Smith, 2019). Granulocytes of Cluster 3 can be delineated by the expression of *LRIM6*, as well as *cathepsin L* and *cathepsin F*. In other invertebrate systems, cathepsin L has been implicated in hemocyte lysosomes, serving important roles in the degradation of phagocytosed materials by phagocytic immune cells (Jiang et al., 2018; Tryselius and Hultmark, 1997). Both cathepsin L and cathepsin F have been associated with anti-microbial activity (Guo et al., 2018; Jiang et al., 2018), that together infer that these cells likely play an important role in immunity and immune homeostasis, although at present it is unclear why these cells are only detected under naïve conditions. Cells within Cluster 4 display high levels of *FBN10*, which resemble previously described PPO6^{high} phagocytic cell populations (Kwon and Smith, 2019; Severo et al., 2018). Our lineage analysis suggests that Cluster 4 may serve as an intermediate, potentially giving rise to cells of Clusters 2 and 3, or could alternatively represent different activation states among these granulocyte populations. Supporting this theory, Cluster 3 cells were only identified under naïve conditions, whereas both Clusters 2 and 4 displayed transcriptional changes in response to blood-feeding. However, additional studies are required to more closely resolve the impacts that feeding status (naïve, blood-fed, *Plasmodium* infection) may have on these granulocyte subtypes as transient cell states or as differentiated cell types.

In addition to identifying multiple granulocyte subtypes, we also define two populations of mosquito oenocytoids (clusters 7 and 8) that likely reflect immature and mature cell populations analogous to those recently described for *Drosophila* crystal cell populations (Cho et al., 2020; Koranteng et al., 2020; Tattikota et al., 2020). Mature oenocytoids (Cluster 8) are denoted in part by the increased expression of *lozenge*, *PPO1*, *pebbled*, *DNA J*, *MLF*, *Notch*, and *klumpfuss* as previously described (Cho et al., 2020; Koranteng et al., 2020; Tattikota et al., 2020), as well as the increase in *SCRB3* and *SCRB9* which

serve as a new marker of mosquito oenocytoids in our analysis. Despite these differences in cell maturation, the mosquito oenocytoid lineage is very distinct from that of granulocytes and relies on the expression of *lozenge* to promote oenocytoid differentiation as in *Drosophila* crystal cells (Fossett et al., 2003; Waltzer et al., 2003). In *Drosophila*, *lozenge* is a transcription factor that interacts with the GATA factor *serpent* to promote the crystal cell lineage from embryonic or larval lymph gland prohemocytes (Fossett et al., 2003; Waltzer et al., 2003), a process that is tightly regulated by *u-shaped* expression (Fossett et al., 2003). In our analysis, *u-shaped* is expressed in mature granulocytes (Cluster 2-4), implying that its expression can be a marker of differentiated granulocytes that are no longer able to adopt an oenocytoid cell fate. This is supported by our on lineage analysis, where differentiation likely occurs from circulating prohemocyte precursor populations (Cluster 5), yet we cannot rule out the potential that immature granulocyte populations are able to undergo transdifferentiation as previously described in *Drosophila* (Leitao and Sucena, 2015). Together with our lozenge gene-silencing data, this suggests that the regulation of oenocytoid differentiation may be highly conserved between *Drosophila* and *Anopheles*.

Our study also breaks down existing paradigms that insect oenocytoids/crystal cells are primarily associated with prophenoloxidase (PPO) production and melanization (Hillyer and Strand, 2014; Lavine and Strand, 2002; Lu et al., 2014). This largely stems from work in *Drosophila*, where two of the three PPOs (*PPO1* and *PPO2)* are expressed in crystal cells (Dudzic et al., 2015), and in *Bombyx mori* where PPOs are exclusively synthesized by oenocytoids (Iwama and Ashida, 1986). However, our scRNA-seq results suggest that both granulocytes and oenocytoids are involved in PPO production, and that distinct subsets of PPOs are differentially regulated in the granulocyte and oenocytoid lineages. Following *lozenge*-silencing, we see significant decreases in *PPO1*, *PPO3*, and *PPO8* expression, transcripts that were highly enriched in oenocytoids in our study, while the remaining PPO genes were unaffected. This is supported by similar RNAi experiments in *Aedes aegypti*, where lozenge and Toll signaling influence orthologous PPO gene expression (Zou et al., 2008). Recent studies of prostaglandin signaling in *An. gambiae* have also implicated the regulation of *PPO1*, *PPO3*, and *PPO8* in oenocytoid populations (Kwon et al., 2020), providing further support that a subset of PPOs are specifically

regulated in oenocytoids. In addition, several lines of evidence support that granulocyte populations also contribute to PPO expression in *An. gambiae*, including PPO6 transgene expression and staining in mosquito granulocytes (Bryant and Michel, 2016, 2014; Castillo et al., 2006; Kwon and Smith, 2019; Severo et al., 2018), the effects of phagocyte depletion on PPO expression (Kwon and Smith, 2019), and the identification of multiple PPOs in the mosquito phagocyte proteome (Smith et al., 2016). This is a departure from other insect systems and is most likely a reflection of the expansion of the PPO gene family in mosquito species, where a total of 9 *An. gambiae* PPOs have been annotated with yet undescribed function. Together, these results suggest new possible roles for PPOs in mosquito immune cells and their respective roles in the innate immune response.

Initial comparisons to recently published *Drosophila* immune cell scRNA-seq experiments (Cattenoz et al., 2020; Tattikota et al., 2020) reveal both similarities and differences in immune cell populations between dipteran species. Similar to *Drosophila*, our data in mosquitoes support the developmental progression and specialization of immune cells from precursor populations, the presence of multiple phagocytic cell populations, and multiple shared markers that delineate the oenocytoid/crystal cell lineage (Cattenoz et al., 2020; Tattikota et al., 2020). However, several differences are also noted in mosquitoes, including the absence of well-characterized *Drosophila* immune cell markers such as hemolectin (Goto et al., 2003; Pondeville et al., 2020) and hemese (Kurucz et al., 2003), and the lack of lamellocyte cell populations. We also identify several mosquito immune cell markers, such as LRIM15 and the expansion of PPO genes, which are unique to mosquito immune cell populations. Other respective differences in the isolation of cells from larvae or adults in *Drosophila* and mosquitoes, may further explain other disparities in cell types or steady states of activation. Ultimately, these questions require a more in-depth comparison of immune cells between these two species in the future.

When placed in the context of previously published scRNA-seq studies in mosquitoes (Raddi et al., 2020; Severo et al., 2018), our results provide additional resolution and perspective to the burgeoning study of mosquito immune cells. With regard to the study by Severo *et al.* (Severo et al., 2018), we expand upon the initial characterization of 25 cells and classification of PPO6^{high} and PPO6^{low} immune cell subtypes, providing an

increased number of cells examined (n=262) and the number of described hemocyte subtypes. Furthermore, our experiments support that the previously defined PPO6^{high} and PPO6^{low} populations (Severo et al., 2018) represent granulocyte subtypes based on their phagocytic ability similarities to gene expression profiles identified in our analysis. When compared to Raddi *et al.* (Raddi et al., 2020), our analyses are highly complementary, resulting in the description of comparable immune cell subtypes, despite having ~1/20th the number of single-cell transcriptomes in our study. Both studies are able to similarly define prohemocytes, multiple granulocyte populations (including an immune-enriched subtype), and oenocytoids. However, our studies are not without their differences. In our study we do not identify cell signatures of megacytes, rare immune cell populations denoted by *TM7318* and *LL3* expression (Raddi et al., 2020). At present, it is unclear if this is due to the smaller number of immune cells examined in our study, or if this is due to technical differences in the experimental approach between studies (Drop-seq vs. FACS isolation in our study). In addition, there are also differences in the description of oenocytoids between our studies, in which we suggest that oenocytoids are derived from prohemocytes and provide evidence that these cell populations undergo maturation similar to that in *Drosophila* (Koranteng et al., 2020; Tattikota et al., 2020). While there are additional details unique to each of these respective studies, we believe that they together represent a significant advancement in our understanding of mosquito immune cells, serving as a strong foundation for future studies of mosquito hemocyte biology.

A primary goal of our study has also been to integrate our dataset with previous descriptions of *Anopheles* hemocytes, thereby enhancing its role as a community resource by placing our analysis in the larger context of previously published work. As a result, we incorporate several markers detailed in previous immunofluorescence assays (Bryant and Michel, 2016, 2014; Castillo et al., 2006; Pinto et al., 2009), transcriptional studies (Pinto et al., 2009), and proteomic analysis (Smith et al., 2016) that have proven instrumental to the characterization of our immune cell clusters. Strengthened by homology to studies of *Drosophila* hemocytes (Cattenoz et al., 2020; Cho et al., 2020; Tattikota et al., 2020), we provide a reliable set of lineage-specific markers that accurately define the granulocyte and oenocytoid lineages. When paired with phagocytosis assays and methods of phagocyte depletion (Kwon and Smith, 2019), we provide an enhanced

set of tools to define mosquito immune cell populations. In contrast to previous reports of rare phagocytic events by oenocytoids when relying only on cellular morphology (Hillyer et al., 2003), our RNA-FISH data support that mosquito oenocytoid populations do not undergo phagocytosis. Therefore, our results improve upon existing knowledge to increase the consistency and resolution in the study of mosquito immune cells.

In addition, our study also provides several new insights into mosquito immune cell biology that warrant further study. This includes the expression of CLIP-domain serine proteases (CLIPs) and serine protease inhibitors (SRPNs) in distinct oenocytoid and granulocyte populations, the enrichment of chemosensory genes in oenocytoids, and the variable expression of tRNAs across immune cell clusters. Together, these are suggestive of distinct transcriptional repertoires within immune cell subtypes that may impart as of yet unknown biological functions specialized for each cell type. Our data also provide additional detail into the regulation of immune cell differentiation, demonstrating the integral role of lozenge in driving the oenocytoid cell lineage. Other transcription factors such as LL3 (Raddi et al., 2020; Smith et al., 2015) and STAT-A (Smith et al., 2015) have been implicated in the differentiation of the granulocyte and oenocytoid lineages in *Anopheles*, yet at present, we have very little understanding of the signals that promote mosquito hematopoiesis and hemocyte differentiation. It is also of interest to examine how immune cell populations differ between life stages (larvae and adult), to determine how immune cells are influenced by the microbiota, and how different pathogen signatures can influence immune cell development and differentiation. As a result, we believe that the candidate markers and cell lineage progressions proposed by our study provide an essential first step to approach this multitude of questions in *Anopheles* through future work.

In summary, our characterization of mosquito hemocytes by scRNA-seq and accompanying functional validation experiments provide an important advancement in our understanding of *An. gambiae* immune cell populations. Through the molecular characterization of at least seven immune cell subtypes and the development of dependable molecular markers to distinguish between cell lineages, we have developed new molecular tools where genetic resources were previously lacking. Through functional

data and efforts to incorporate existing knowledge of mosquito hemocytes, we believe that our data will serve as an important resource for the vector community, that together offer new insights into the complexity of mosquito immune cells and provide a strong foundation for comparative function of insect immune cells.

Materials and Methods

Mosquito rearing

Adult *An*. *gambiae* mosquitoes of the Keele strain (Hurd et al., 2005; Ranford-Cartwright et al., 2016) were reared at 27°C with 80% relative humidity and a 14/10 hour light/dark cycle. Larvae were reared on a diet of fish flakes (Tetramin, Tetra), while adult mosquitoes were maintained on 10% sucrose solution and commercial sheep blood for egg production.

Isolation and sorting of mosquito immune cells for single-cell RNA sequencing

Hemolymph was perfused from female mosquitoes (n=40) under naïve (3- to 5-day old) or blood-fed (~24h post-feeding) conditions using an anticoagulant solution as previously described (Kwon and Smith, 2019; Smith et al., 2016). Perfused hemolymph was diluted with 1X PBS to a total volume of 1 mL, then centrifuged for 5 min at 2000×g to pellet cells. After the supernatant was discarded, cells were washed two times in 1X PBS with an additional centrifugation step of 5 min at 2000×g between washing steps. Cells were incubated with WGA (1:5000, Sigma), DRAQ5 (1:1000, Thermo Fisher Scientific) and Live/Dead Fixable Dead Cell Stain (1:1000 Thermo Fisher Scientific) for 90 min at room temperature. Following incubation, cells were washed twice in 1X PBS to remove excess stain with a centrifugation step of 5 min at 2000 ×g and run on a BD FACSCanto cytometer (BD Biosciences). Based on the previous flow cytometry data for establishment of threshold values for gating (Kwon and Smith, 2019), cells smaller or larger than single cells were excluded. Cell viability was determined by the intensity of the blue fluorescent signal from the Live/Dead Fixable Dead Cell Stain, where dead cells display higher fluorescent signal. Following gating for cell viability, cell populations were distinguished by WGA and DRAQ5 signals and with individual cells sorted into each well of a 384-well plate (twintec™ PCR plates, Eppendorf, Germany) containing 2.3 µl lysis buffer (Picelli et al., 2013).

cDNA libraries were generated using a slightly modified version of Smart-seq2 as previously described (Picelli et al., 2013), where 23 cycles for cDNA amplification were used. Single-cell libraries were sequenced on the HiSeq2500 platform (Illumina) using 56 base pair single-end sequencing. Library preparation and sequencing was performed at ESCG and NGI, SciLifeLab, Sweden.

Computational analysis

Sequencing reads were mapped to the *Anopheles gambiae* AgamP4 reference genome (Ensembl release 40) and ERCC sequences with the HISAT2 aligner version 2.1.0 (Kim et al., 2015). Only alignments with mapping quality 60 were retained. Quantification of gene expression was performed on the gene level. Overlapping exons of the same gene were merged based on their annotation coordinates. Counting of alignments on genome annotations was performed with the program subread version 1.6.2 (Liao et al., 2014) with the '-s' flag set to 0. Quality control of the data was performed by examining the fraction of sequencing output from ERCC templates versus the genome. Moreover, we applied a threshold of a minimum of 10,000 uniquely mapped reads per cell, only considering reads mapped in exons, i.e. intronic and intergenic reads are not counted towards the 10,000-minimum threshold; cells with fewer reads were not included in downstream analyses. Raw read counts were normalized to RPKM to adjust for gene length (Mortazavi et al., 2008). RPKM values were transformed with the function log2(**x**+1), where **x** is a vector of expression values. The complete quality-filtered gene expression data (as RPKM values) for all 262 cells is found in Table S1. Statistical analyses and data visualization were performed with the R software (www.r-project.org) version 3.5.3. Hierarchical clustering with Euclidean distance was used to define cell clusters; the ward.D2 agglomeration method was used for linkage. The final clusters were defined using a combination of manual examination of the tree structure and the cutreeDynamic function of the R package dynamicTreeCut (with method set to hybrid and deepSplit set to 4) (Langfelder et al., 2008).

Clustering and functional analysis of single-cell data

Raw reads count for single cells data were scaled and normalized using scale factor function (log10) for genes of interest. The percentage and the average expression of the

selected genes were calculated in a scaled normalized expression along a continuous color scale. Seurat Dotplot version (3.1.5) was used for visualization in R (version 3.5.3) software. The color intensity is proportional to a scaled average gene expression level for the selected genes across all clusters and the size of the circle is correspondence to the percentage of cells within each cluster expressing a gene (Stuart and Satija, 2019).

Heatmaps were produced using the pheatmap package (version 1.0.12), where the average expression of the selected genes was calculated from the normalized scaled RPKM values from clustering data. The working data frame matrix was prepared using tidy-verse package version (1.3.0) for heatmap construction on selected gene sets corresponding (hemocyte gene orthologs, immune genes, etc.) to visualize expression across clusters. The clustering distance applied to the heatmaps was based on spearman.complete.obs with the scale set to be the selected genes as a comparison for visualization. Color intensity corresponds to the average transcript abundance encoded with navy, white, red and firebrick 3, with the latter indicating a significant increase in the expression of a given gene.

Cell cycle gene analysis was performed using the Seurat function “AddModuleScore” to calculate the average expression levels of transcripts annotated to be involved in the cell cycle (GO:0007049) 191 transcripts similar to Raddi *et al.* (Raddi et al., 2020). After filtering out cells expressing less than 1% of the transcripts, 172 remaining transcripts were used in a combined expression score to calculate the enrichment of cell cycle genes among clusters. Positive scores indicate higher expression of genes involved in cell cycle regulation suggestive of cell proliferation.

To perform gene ontology (GO) analysis on each of the defined cell clusters, transcripts expressed in >80% of each respective cell cluster (Table S5) were examined to characterize the molecular composition of each cell type. Gene IDs (AGAP accession numbers) were classified based on gene ontology as previously (Mendes et al., 2011; Smith et al., 2016) to identify the functional categories of proteins within each cell cluster and to enable comparisons between cell clusters.

Differential expression analysis was performed using linear models as implemented in the alona software (https://github.com/oscar-franzen/alona/) (Franzén and Björkegren,

2020). Genes with false discovery rate (FDR; Benjamini-Hochberg's procedure) <5% and the absolute value of the log fold change>1.5 were considered significantly differentially expressed. All data can be visualized using the alona server (Franzén and Björkegren, 2020) at the following project link: https://alona.panglaodb.se/results.html?job=2c2r1NM5Zl2qcW44RSrjkHf3Oyv51y_5f09d74b770c9.

Cell trajectory and pseudotime analysis

Cells were assigned to cell groups using Monocle3 with UMAP clustering (Cao et al., 2019; Packer et al., 2019; Qiu et al., 2017; Trapnell et al., 2014). Data was normalized to remove batch effects using PCA clustering to 100 dimensions (Haghverdi et al., 2018). Pseudotime was calculated in Monocle3, with colors representing pseudotime changes among the cell clusters (Cao et al., 2019; Packer et al., 2019; Qiu et al., 2017; Trapnell et al., 2014). Bioinformatic methods for Monocle analyses can be found at: https://github.com/ISUgenomics/SingleCellRNAseq_RyanSmith.

RNA-FISH

In order to classify hemocyte populations by detecting specific RNA expression, we used RNAscope®Multiplex Fluorescent Reagent Kit v2 Assay (Advanced Cell Diagnostics), and *in situ* hybridization was performed using the manufacturer's instruction. Using anticoagulant solution, hemolymph was perfused from non-blood fed mosquitoes (3 to 5 days old) and placed on a superfrost microscopic slide (Fisher Scientific) to adhere at RT for 20 min. Cells were fixed with 4% paraformaldehyde for 15 min at 4°C, then washed three times with 1X PBS. Hydrogen peroxide was applied to the cells, and slides were incubated for 10 min at room temperature (RT). After washing three times in sterile distilled water, cells were treated Protease IV and incubated for 30 min at RT. To delineate hemocyte populations, a nimrod B2 (AGAP029054) RNA probe conjugated with C1 (Severo et al., 2018) was used as a universal maker and was mixed with a specific RNA probe conjugated with C2, either leucine rich-repeat immune protein 15 (LRIM15: AGAP007045; regions 2-874 of XM_308718.4), or lozenge (Lz: AGAP002506; regions 168-1372 of XM_312433.5). For the identification of oenocytoid populations, SCRB3 (AGAP005725; regions 337-1276 of XM_315741.5) or SCRB9 (AGAP004846; regions

402-1306 of XM_001688510.1) RNA probes conjugated with C1 were co-incubated with either LRIM15 or Lz probes. All RNAscope probes are commercially available through Advanced Cell Diagnostics. Fixed hemocyte slides were hybridized with the respective mixtures of RNA probes in a HybEZ™Oven for 2 h at 40°C. After washed two times with wash buffer for 2 min, hybridized probes were incubated with respective AMP reagents (AMP1 and AMP2) for 30 min at 40°C and with AMP3 for 15 min at 40°C. Cells were washed two times with wash buffer between AMP incubations. Cells were incubated with RNAscope®Multiplex FLv2 HRP-C1 for 15 min at 40°C, labeled with selected Opal Fluorophore reagent (Akoya Bioscience) at dilution facto (1:1000) for 30 min at 40°C and treated with RNAscope®Multiplex FLv2 HRP blocker for 15 min at 40°C. Cells were washed two times with wash buffer for 2 min between incubations. Following C1 labeling, cells were incubated with a specific RNAscope®Multiplex FLv2 HRP-C2 conjugated solution, desired Opal Fluorophore reagent (1:1000; Opal™520 or Opal™570) and HRP blocker. Slides were initially treated DAPI (Advanced Cell Diagnostics) for 30 sec at RT and mounted with ProLong®Diamond Antifade Mountant with DAPI (Life Technologies). Cells displaying a positive signal were quantified as the percentage of positive cells of the total number of cells examined. Counts were performed from >50 adherent cells per mosquito from randomly chosen fields using fluorescence microscopy (Nikon Eclipse 50i, Nikon).

Phagocytosis assays

Phagocytosis assays were performed by injecting 69 nl of 2% green fluorescent FluoSpheres (vol/vol) in 1X PBS to naïve female mosquitoes (3- to 5-day old) using a Nanoject II injector (Drummond Scientific). After injection, mosquitoes were kept at 27°C for 2 hours before hemolymph was perfused on a superfrost slide. To define phagocytic cell populations, phagocytosis assays were paired with RNA-FISH experiments as described above using *LRIM15*, *SCRB9*, and *Lz* RNA probes. The proportion of phagocytic cells was quantified as the number of cells that had phagocytosed one or more beads of the total number of cells examined that displayed a positive signal for each of the respective RNA-FISH probes. Counts were performed from >50 adherent cells per mosquito from randomly chosen fields using fluorescence microscopy (Nikon Eclipse 50i, Nikon).

Phagocyte depletion using clodronate liposomes

To confirm the identification of phagocytic cells from our defined immune cell clusters, validation experiments were performed using clodronate liposomes (CLD) to deplete phagocytic hemocytes (Kwon and Smith, 2019). Naïve female mosquitoes (3- to 5-day old) were injected with either 69 nl of control liposomes (LP) or CLD (Standard macrophage depletion kit, Encapsula NanoSciences LLC) at 1:5 dilution in 1X PBS. At 24 h post-injection, hemolymph was perfused, and RNA-FISH was performed to differentiate affected cell populations using RNA probes for NimB2 (C1), LRIM15 (C2), and Lz (C2) as described above. Hemocytes displaying a positive signal were quantified from 50 or more cells per mosquito.

qRT-PCR

Gene expression analysis using qRT-PCR was performed to validate the influence of phagocyte depletion on non-phagocytic and phagocytic cells. cDNA was prepared from the previous studies (Kwon and Smith, 2019) corresponding to naïve adult female *An. gambiae* treated with either control liposomes (LP) or clodronate liposomes (CLD) 24 h post-treatment. Hemocyte cDNA was prepared as previously described (Kwon and Smith, 2019) to analyze relative gene expression of *lozenge* between LP and CLD treatments. Specific transcripts representative of non-phagocytic and phagocytic cell populations were examined by qRT-PCR using primers listed in Table S6.

Gene silencing by RNAi

RNAi experiments were performed as previously described (Kwon et al., 2017; Kwon and Smith, 2019; Reynolds et al., 2020; Smith et al., 2016, 2015). T7 primers for lozenge (Lz; AGAP002506) were designed using the E-RNAi web application (http://www.dkfz.de/signaling/e-rnai3/idseq.php) and listed in Table S7. T7 templates for dsRNA synthesis were prepared from amplified cDNA from 4-day old whole naïve mosquitoes. PCR amplicons were purified using the DNA Clean & Concentration kit (Zymo Research), and dsRNAs were synthesized using the MEGAscript RNAi kit (Life Technologies). Subsequent dsRNA targeting GFP (control) or Lz was resuspended in nuclease free water to 3 µg/µl after ethanol precipitation. Injections were performed in three- to four-day old cold anesthetized mosquitoes by intrathoracic injection with 69 nl

(~200 ng) of dsRNA per mosquito using a Nanoject III. The effects of gene silencing were measured at 3 days post-injection in whole mosquitoes (n=15) by qRT-PCR as previously described (Kwon and Smith, 2019).

Acknowledgements

We would like to thank Shawn Rigby of the Iowa State Flow Cytometry Facilty for his assistance with FACS as well as Anna-Maria Divne at the Microbial Single Cell Genomics (MSCG) facility, SciLifeLab, for consultation regarding the sorting of individual hemocytes, the Eukaryotic Single Cell Genomics (ESCG) facility, SciLifeLab, for assistance with single-cell library preparation and The National Genomics Infrastructure, SciLifeLab, Sweden for assistance with Illumina sequencing. We also want to thank Maiara Severo-Witte and Elena Levashina for initial assistance with the RNA-FISH methods, Rick Masonbrink and Andrew Severin of the Iowa State Genome Informatics Facility for assistance with the single-cell RNA-seq analysis. This work was supported by the Swedish Society for Medical Research (SSMF) and the Swedish Research Council (VR-NT) to J.A., the Agricultural Experiment Station at Iowa State University and the National Institutes of Health, National Institute of Allergy and Infectious Diseases (R21 AI44705) to R.C.S.

References

Banerjee U, Girard JR, Goins LM, Spratford CM. 2019. *Drosophila* as a genetic model for hematopoiesis. *Genetics* **211**:367–417.

Baton LA, Robertson A, Warr E, Strand MR, Dimopoulos G. 2009. Genome-wide transcriptomic profiling of *Anopheles gambiae* hemocytes reveals pathogen-specific signatures upon bacterial challenge and *Plasmodium berghei* infection. *BMC Genomics* **10**:257.

Benoit JB, Vigneron A, Broderick NA, Wu Y, Sun JS, Carlson JR, Aksoy S, Weiss BL. 2017. Symbiont-induced odorant binding proteins mediate insect host hematopoiesis. *Elife* **6**:709–719.

Bryant WB, Michel K. 2016. *Anopheles gambiae* hemocytes exhibit transient states of activation. *Dev Comp Immunol* **55**:119–129.

Bryant WB, Michel K. 2014. Blood feeding induces hemocyte proliferation and activation in the African malaria mosquito, *Anopheles gambiae* Giles. *J Exp Biol* **217**:1238–45.

Butler A, Hoffman P, Smibert P, Papalexi E, Satija R. 2018. Integrating single-cell transcriptomic data across different conditions, technologies, and species. *Nat Biotechnol* **36**:411–420.

Cao J, Spielmann M, Qiu X, Huang X, Ibrahim DM, Hill AJ, Zhang F, Mundlos S, Christiansen L, Steemers FJ, Trapnell C, Shendure J. 2019. The single-cell transcriptional landscape of mammalian organogenesis. *Nature* **566**:496–502.

Castillo J, Brown MR, Strand MR. 2011. Blood feeding and insulin-like peptide 3 stimulate proliferation of hemocytes in the mosquito *Aedes aegypti*. *PLoS Pathog* **7**:e1002274.

Castillo JC, Beatriz A, Ferreira B, Trisnadi N, Barillas-mury C. 2017. Activation of mosquito complement antiplasmodial response requires cellular immunity. *Sci Immunol* **2**:eaal1505.

Castillo JC, Robertson AE, Strand MR. 2006. Characterization of hemocytes from the mosquitoes *Anopheles gambiae* and *Aedes aegypti*. *Insect Biochem Mol Biol*

36:891–903.

Cattenoz PB, Sakr R, Pavlidaki A, Delaporte C, Riba A, Molina N, Hariharan N, Mukherjee T, Giangrande A. 2020. Temporal specificity and heterogeneity of *Drosophila* immune cells . *EMBO J* 39:e104486.

Chaplin DD. 2010. Overview of the immune response. *J Allergy Clin Immunol* **125**:S3–S23.

Cho B, Yoon SH, Lee DD, Koranteng F, Tattikota SG, Cha N, Shin M, Do H, Hu Y, Oh SY, Lee DD, Vipin Menon A, Moon SJ, Perrimon N, Nam JW, Shim J. 2020. Single-cell transcriptome maps of myeloid blood cell lineages in *Drosophila*. *Nat Commun* **11**, 4483.

Cirimotich CM, Dong Y, Garver LS, Sim S, Dimopoulos G. 2010. Mosquito immune defenses against *Plasmodium* infection. *Dev Comp Immunol* **34**:387–395.

Danielli A, Loukeris TG, Lagueux M, Müller HM, Richman A, Kafatos FC. 2000. A modular chitin-binding protease associated with hemocytes and hemolymph in the mosquito *Anopheles gambiae*. *Proc Natl Acad Sci U S A* **97**:7136–7141.

Danielli A, Kafatos FC, Loukeris TG. 2003. Cloning and characterization of four *Anopheles gambiae* serpin isoforms, differentially induced in the midgut by *Plasmodium berghei* invasion. *J Biol Chem* **278**:4184–4193.

Dudzic JP, Kondo S, Ueda R, Bergman CM, Lemaitre B. 2015. *Drosophila* innate immunity: Regional and functional specialization of prophenoloxidases. *BMC Biol* **13**:81.

Estévez-Lao TY, Hillyer JF. 2014. Involvement of the *Anopheles gambiae* Nimrod gene family in mosquito immune responses. *Insect Biochem Mol Biol* **44**:12–22.

Evans CJ, Liu T, Banerjee U. 2014. *Drosophila* hematopoiesis: Markers and methods for molecular genetic analysis. *Methods* **68**:242–251.

Fossett N, Hyman K, Gajewski K, Orkin SH, Schulz RA. 2003. Combinatorial interactions of serpent, lozenge, and U-shaped regulate crystal cell lineage commitment during

Drosophila hematopoiesis. *Proc Natl Acad Sci U S A* **100**:11451–11456.

Franc NC, Dimarcq JL, Lagueux M, Hoffmann J, Ezekowitz RAB. 1996. Croquemort, a novel *Drosophila* hemocyte/macrophage receptor that recognizes apoptotic cells. *Immunity* **4**:431–443.

Franzén O, Björkegren JLM. 2020. alona: a web server for single cell RNA-seq analysis. *Bioinformatics* **36**: 3910–3912

Goto A, Kadowaki T, Kitagawa Y. 2003. *Drosophila* hemolectin gene is expressed in embryonic and larval hemocytes and its knock down causes bleeding defects. *Dev Biol* **264**:582–591.

Gulley MM, Zhang X, Michel K. 2013. The roles of serpins in mosquito immunology and physiology. *J Insect Physiol* **59**:138–147.

Guo H, Li Y, Zhang M, Li R, Li W, Lou J, Bao Z, Wang Y. 2018. Expression of Cathepsin F in response to bacterial challenges in Yesso scallop *Patinopecten yessoensis*. *Fish Shellfish Immunol* **80**:141–147.

Haghverdi L, Lun ATL, Morgan MD, Marioni JC. 2018. Batch effects in single-cell RNA-sequencing data are corrected by matching mutual nearest neighbors. *Nat Biotechnol* **36**:421–427.

Hillyer Julián F, Schmidt SL, Christensen BM. 2003. Rapid phagocytosis and melanization of bacteria and Plasmodium sporozoites by hemocytes of the mosquito *Aedes aegypti*. *J Parasitol* **89**:62–69.

Hillyer Julián F., Schmidt SL, Christensen BM. 2003. Hemocyte-mediated phagocytosis and melanization in the mosquito *Armigeres subalbatus* following immune challenge by bacteria. *Cell Tissue Res* **313**:117–127.

Hillyer JF, Strand MR. 2014. Mosquito hemocyte-mediated immune responses. *Curr Opin Insect Sci* **3**:14–21.

Horne MC, Donaldson KL, Goolsby GL, Tran D, Mulheisen M, Hell JW, Wahl AF. 1997. Cyclin G2 is up-regulated during growth inhibition and B cell antigen receptor-

mediated cell cycle arrest. *J Biol Chem* **272**:12650–12661.

Hurd H, Taylor PJ, Adams D, Underhill A, Eggleston P. 2005. Evaluating the costs of mosquito resistance to malaria parasites. *Evolution* **59**:2560–2572.

Iwama R, Ashida M. 1986. Biosynthesis of prophenoloxidase in hemocytes of larval hemolymph of the silkworm, *Bombyx mori*. *Insect Biochem* **16**:547–555.

Jiang S, Qiu L, Wang L, Jia Z, Lv Z, Wang M, Liu C, Xu J, Song L. 2018. Transcriptomic and quantitative proteomic analyses provide insights into the phagocytic killing of hemocytes in the Oyster *Crassostrea gigas*. *Front Immunol* **9**:1280.

Kanost MR, Jiang H. 2015. Clip-domain serine proteases as immune factors in insect hemolymph. *Curr Opin Insect Sci* **11**:47–55.

Kim D, Langmead B, Salzberg SL. 2015. HISAT: A fast spliced aligner with low memory requirements. *Nat Methods* **12**:357–360.

King JG, Hillyer JF. 2013. Spatial and temporal in vivo analysis of circulating and sessile immune cells in mosquitoes: hemocyte mitosis following infection. *BMC Biol* **11**:55.

Kocks C, Cho JH, Nehme N, Ulvila J, Pearson AM, Meister M, Strom C, Conto SL, Hetru C, Stuart LM, Stehle T, Hoffmann JA, Reichhart J-M, Ferrandon D, Rämet M, Ezekowitz RAB. 2005. Eater, a transmembrane protein mediating phagocytosis of bacterial pathogens in *Drosophila*. *Cell* **123**:335–346.

Koranteng F, Cha N, Shin M, Shim J. 2020. The role of lozenge in *Drosophila* hematopoiesis. *Mol Cells* **43**:114–120.

Krishna S, Yim DG, Lakshmanan V, Tirumalai V, Koh JL, Park JE, Cheong JK, Low JL, Lim MJ, Sze SK, Shivaprasad P, Gulyani A, Raghavan S, Palakodeti D, DasGupta R. 2019. Dynamic expression of tRNA-derived small RNAs define cellular states. *EMBO Rep* e47789.

Kurucz E, Zettervall C-J, Sinka R, Vilmos P, Pivarcsi A, Ekengren S, Hegedüs Z, Ando I, Hultmark D. 2003. Hemese, a hemocyte-specific transmembrane protein, affects the cellular immune response in *Drosophila*. *Proc Natl Acad Sci U S A* **100**:2622–2627.

Kwon H, Arends BR, Smith RC. 2017. Late-phase immune responses limiting oocyst survival are independent of TEP1 function yet display strain specific differences in *Anopheles gambiae*. *Parasit Vectors* **10**:369.

Kwon H, Hall DR, Smith RC. 2020. Identification of a prostaglandin E2 receptor that regulates mosquito oenocytoid immune cell function in limiting bacteria and parasite infection. *bioRxiv* 2020.08.03.235432

Kwon H, Smith RC. 2019. Chemical depletion of phagocytic immune cells in *Anopheles gambiae* reveals dual roles of mosquito hemocytes in anti-*Plasmodium* immunity. *Proc Natl Acad Sci* **116**:14119–14128.

Langfelder P, Zhang B, Horvath S. 2008. Defining clusters from a hierarchical cluster tree: The Dynamic Tree Cut package for R. *Bioinformatics* **24**:719–720. doi:10.1093/bioinformatics/btm563

Lavine M, Strand MR. 2002. Insect hemocytes and their role in immunity. *Insect Biochem Mol Biol* **32**:1295–1309.

Leitao AB, Sucena E. 2015. *Drosophila* sessile hemocyte clusters are true hematopoietic tissues that regulate larval blood cell differentiation. *Elife* **2015**:1–38.

Liao Y, Smyth GK, Shi W. 2014. FeatureCounts: An efficient general purpose program for assigning sequence reads to genomic features. *Bioinformatics* **30**:923–930.

Lu A, Zhang Q, Zhang J, Yang B, Wu K, Xie W, Luan YX, Ling E. 2014. Insect prophenoloxidase: The view beyond immunity. *Front Physiol* **5**:252.

Manaka J, Kuraishi T, Shiratsuchi A, Nakai Y, Higashida H, Henson P, Nakanishi Y. 2004. Draper-mediated and phosphatidylserine-independent phagocytosis of apoptotic cells by *Drosophila* hemocytes/macrophages. *J Biol Chem* **279**:48466–48476.

Martinek N, Shahab J, Saathoff M, Ringuette M. 2011. Haemocyte-derived SPARC is required for collagen-IV-dependent stability of basal laminae in *Drosophila* embryos. *J Cell Sci* **124**:670–670.

Martínez-Gac L, Marqués M, García Z, Campanero MR, Carrera AC. 2004. Control of

cyclin G2 mRNA expression by forkhead transcription factors: novel mechanism for cell cycle control by phosphoinositide 3-kinase and forkhead. *Mol Cell Biol* **24**:2181–2189.

Mendes AM, Awono-Ambene PH, Nsango SE, Cohuet A, Fontenille D, Kafatos FC, Christophides GK, Morlais I, Vlachou D. 2011. Infection intensity-dependent responses of *Anopheles gambiae* to the African malaria parasite *Plasmodium falciparum*. *Infect Immun* **79**:4708–4715.

Midega J, Blight J, Lombardo F, Povelones M, Kafatos F, Christophides GK. 2013. Discovery and characterization of two Nimrod superfamily members in *Anopheles gambiae*. *Pathog Glob Health* **107**:463–74.

Miller M, Chen A, Gobert V, Augé B, Beau M, Burlet-Schiltz O, Haenlin M, Waltzer L. 2017. Control of RUNX-induced repression of Notch signaling by MLF and its partner DnaJ-1 during *Drosophila* hematopoiesis. *PLoS Genet* **13**:1–28.

Mortazavi A, Williams BA, McCue K, Schaeffer L, Wold B. 2008. Mapping and quantifying mammalian transcriptomes by RNA-Seq. *Nat Methods* **5**:621–628.

Packer JS, Zhu Q, Huynh C, Sivaramakrishnan P, Preston E, Dueck H, Stefanik D, Tan K, Trapnell C, Kim J, Waterston RH, Murray JI. 2019. A lineage-resolved molecular atlas of *C. elegans* embryogenesis at single-cell resolution. *Science* **365:**eaax1971.

Papalexi E, Satija R. 2017. Single-cell RNA sequencing to explore immune cell heterogeneity. *Nat Rev Immunol*. **18**: 35–45.

Picelli S, Björklund ÅK, Faridani OR, Sagasser S, Winberg G, Sandberg R. 2013. Smart-seq2 for sensitive full-length transcriptome profiling in single cells. *Nat Methods* **10**:1096–1100.

Picelli S, Faridani OR, Björklund ÅK, Winberg G, Sagasser S, Sandberg R. 2014. Full-length RNA-seq from single cells using Smart-seq2. *Nat Protoc* **9**:171–181.

Pinto SB, Lombardo F, Koutsos AC, Waterhouse RM, McKay K, An C, Ramakrishnan C, Kafatos FC, Michel K. 2009. Discovery of *Plasmodium* modulators by genome-wide analysis of circulating hemocytes in *Anopheles gambiae*. *Proc Natl Acad Sci U S A*

106:21270–21275.

Pondeville E, Puchot N, Parvy JP, Carissimo G, Poidevin M, Waterhouse RM, Marois E, Bourgouin C. 2020. Hemocyte-targeted gene expression in the female malaria mosquito using the hemolectin promoter from *Drosophila*. *Insect Biochem Mol Biol* **120**:103339.

Proserpio V, Mahata B. 2016. Single-cell technologies to study the immune system. *Immunology* **147**:133–140.

Qiu X, Mao Q, Tang Y, Wang L, Chawla R, Pliner HA, Trapnell C. 2017. Reversed graph embedding resolves complex single-cell trajectories. *Nat Methods* **14**:979–982.

Raddi G, Barletta AB, Efremova M, Ramirez JL, Cantera R, Teichmann S, Barillas-Mury C, Billker O. 2020. Mosquito cellular immunity at single-cell resolution. *Science* **369**:1128–1132.

Rak R, Polonsky M, Eizenberg I, Dahan O, Friedman N, Pilpel YT. 2020. Dynamic changes in tRNA modifications and abundance during T-cell activation. *bioRxiv* 2020.03.14.991901.

Ramirez JL, de Almeida Oliveira G, Calvo E, Dalli J, Colas R a., Serhan CN, Ribeiro JM, Barillas-Mury C. 2015. A mosquito lipoxin/lipocalin complex mediates innate immune priming in *Anopheles gambiae*. *Nat Commun* **6**:7403.

Ramirez JL, Garver LS, Brayner FA, Alves LC, Rodrigues J, Molina-Cruz A, Barillas-Mury C. 2014. The role of hemocytes in *Anopheles gambiae* antiplasmodial immunity. *J Innate Immun* **6**:119–128.

Ranford-Cartwright LC, McGeechan S, Inch D, Smart G, Richterová L, Mwangi JM. 2016. Characterisation of species and diversity of *Anopheles gambiae* Keele colony. *PLoS One* **11**:e0168999.

Reynolds RA, Kwon H, Smith RC. 2020. 20-hydroxyecdysone primes innate immune responses that limit bacterial and malarial parasite survival in *Anopheles gambiae*. *mSphere* **5**:e00983-19.

Ribeiro C, Brehélin M. 2006. Insect haemocytes: What type of cell is that? *J Insect Physiol* **52**:417–429.

Rodrigues J, Brayner FA, Alves LC, Dixit R, Barillas-Mury C. 2010. Hemocyte differentiation mediates innate immune memory in *Anopheles gambiae* mosquitoes. *Science* **329**:1353–1355.

Severo MS, Landry JJMM, Lindquist RL, Goosmann C, Brinkmann V, Collier P, Hauser AE, Benes V, Henriksson J, Teichmann SA, Levashina EA. 2018. Unbiased classification of mosquito blood cells by single-cell genomics and high-content imaging. *Proc Natl Acad Sci U S A* **115**:E7568-7577.

Smith RC, Barillas-Mury C, Jacobs-Lorena M. 2015. Hemocyte differentiation mediates the mosquito late-phase immune response against *Plasmodium* in *Anopheles gambiae*. *Proc Natl Acad Sci* **112**:E3412-20.

Smith RC, King JG, Tao D, Tomescu O, Brando C, Thallinger GG, Dinglasan RR, Zeleznik OA, Brando C, Thallinger GG, Dinglasan RR. 2016. Molecular profiling of phagocytic immune cells in *Anopheles gambiae* reveals integral roles for hemocytes in mosquito innate immunity. *Mol Cell proteomics* **15**:3373–3387.

Stofanko M, So YK, Badenhorst P. 2008. A misexpression screen to identify regulators of *Drosophila* larval hemocyte development. *Genetics* **180**:253–267.

Strand MR. 2008. The insect cellular immune response. *Insect Sci* **15**:1–14.

Stuart T, Satija R. 2019. Integrative single-cell analysis. *Nat Rev Genet* **20**:257–272.

Szabo PA, Levitin HM, Miron M, Snyder ME, Senda T, Yuan J, Cheng YL, Bush EC, Dogra P, Thapa P, Farber DL, Sims PA. 2019. Single-cell transcriptomics of human T cells reveals tissue and activation signatures in health and disease. *Nat Commun* **10**: 4706.

Tattikota SG, Hu Y, Liu Y, Cho B, Barrera V, Steinbaugh M, Yoon S-H, Comjean A, Li F, Dervis F, Hung R-J, Nam J-W, Sui SH, Shim J, Perrimon N. 2020. A single-cell survey of *Drosophila* blood. *Elife* **9**:e54818.

Terriente-Felix A, Li J, Collins S, Mulligan A, Reekie I, Bernard F, Krejci A, Bray S. 2013. Notch cooperates with Lozenge/Runx to lock haemocytes into a differentiation programme. *Dev* **140**:926–937.

Torrent M, Chalancon G, De Groot NS, Wuster A, Madan Babu M. 2018. Cells alter their tRNA abundance to selectively regulate protein synthesis during stress conditions. *Sci Signal* **11**:1–10.

Trapnell C, Cacchiarelli D, Grimsby J, Pokharel P, Li S, Morse M, Lennon NJ, Livak KJ, Mikkelsen TS, Rinn JL. 2014. The dynamics and regulators of cell fate decisions are revealed by pseudotemporal ordering of single cells. *Nat Biotechnol* **32**:381–386.

Tryselius Y, Hultmark D. 1997. Cysteine proteinase 1 (CP1), a cathepsin L-like enzyme expressed in the *Drosophila melanogaster* haemocyte cell line mbn-2. *Insect Mol Biol* **6**:173–181.

Villani A-C, Satija R, Reynolds G, Sarkizova S, Shekhar K, Fletcher J, Griesbeck M, Butler A, Zheng S, Lazo S, Jardine L, Dixon D, Stephenson E, Nilsson E, Grundberg I, McDonald D, Filby A, Li W, De Jager PL, Rozenblatt-Rosen O, Lane AA, Haniffa M, Regev A, Hacohen N. 2017. Single-cell RNA-seq reveals new types of human blood dendritic cells, monocytes, and progenitors. *Science* **356**:eaah4573.

Waltzer L, Ferjoux G, Bataillé L, Haenlin M. 2003. Cooperation between the GATA and RUNX factors Serpent and Lozenge during *Drosophila* hematopoiesis. *EMBO J* **22**:6516–6525.

Zou Z, Shin SW, Alvarez KS, Bian G, Kokoza V, Raikhel AS. 2008. Mosquito RUNX4 in the immune regulation of PPO gene expression and its effect on avian malaria parasite infection. *Proc Natl Acad Sci U S A* **105**:18454–18459.

Figures

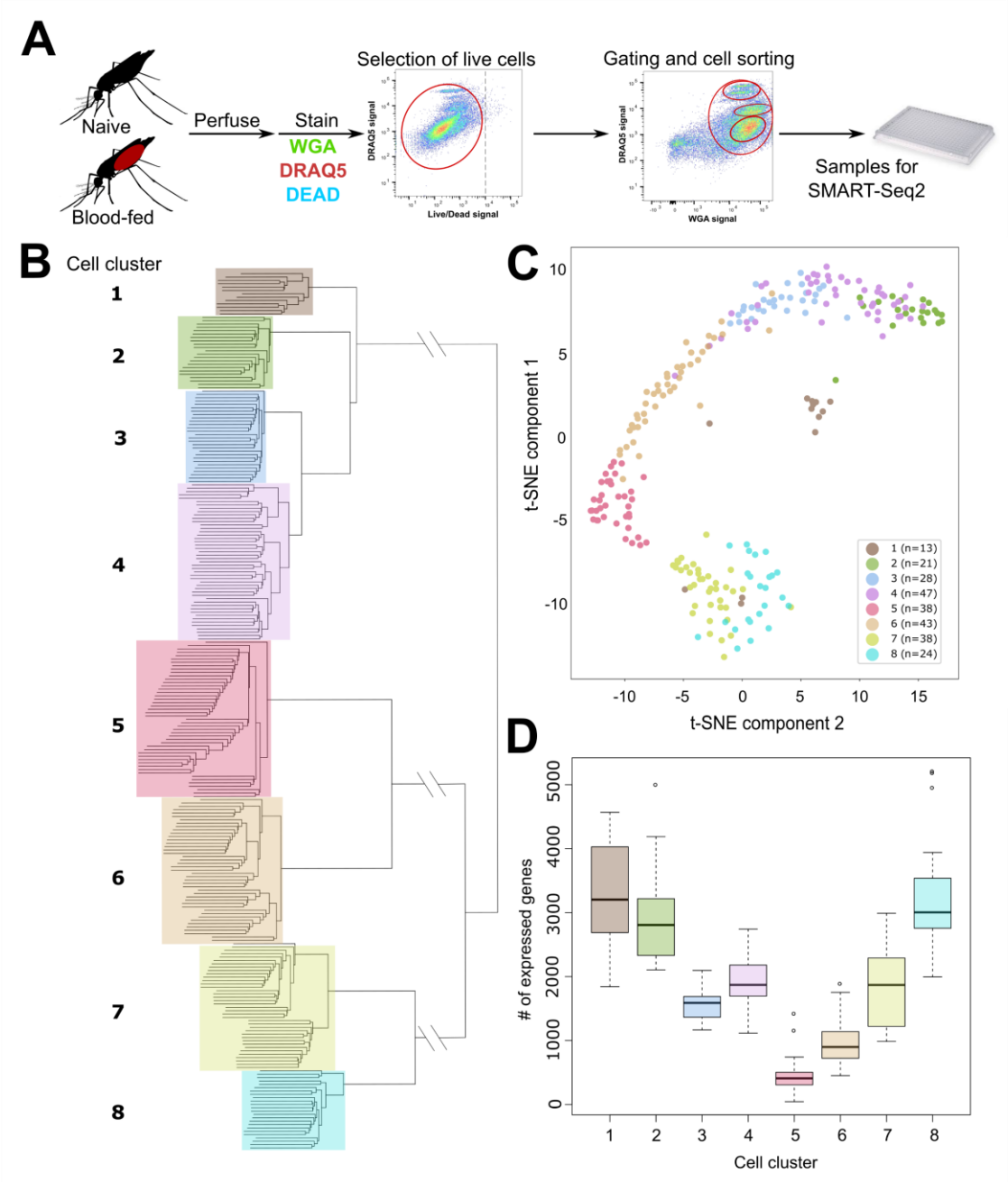

Figure 1. **scRNA-seq of *An. gambiae* immune cells. (A)** Graphical overview of the isolation of mosquito immune cells from naïve and blood-fed mosquitoes. Following perfusion, cells were stained to enable processing by fluorescent activated cell sorting (FACS) and isolation for scRNA-seq. Resulting immune cells data were separated into 8 cell clusters based on hierarchical clustering analysis **(B)** and visualized using a t-

Distributed Stochastic Neighbor Embedding (t-SNE) plot **(C)**. The number of expressed genes per cluster are displayed as a boxplot **(D)**.

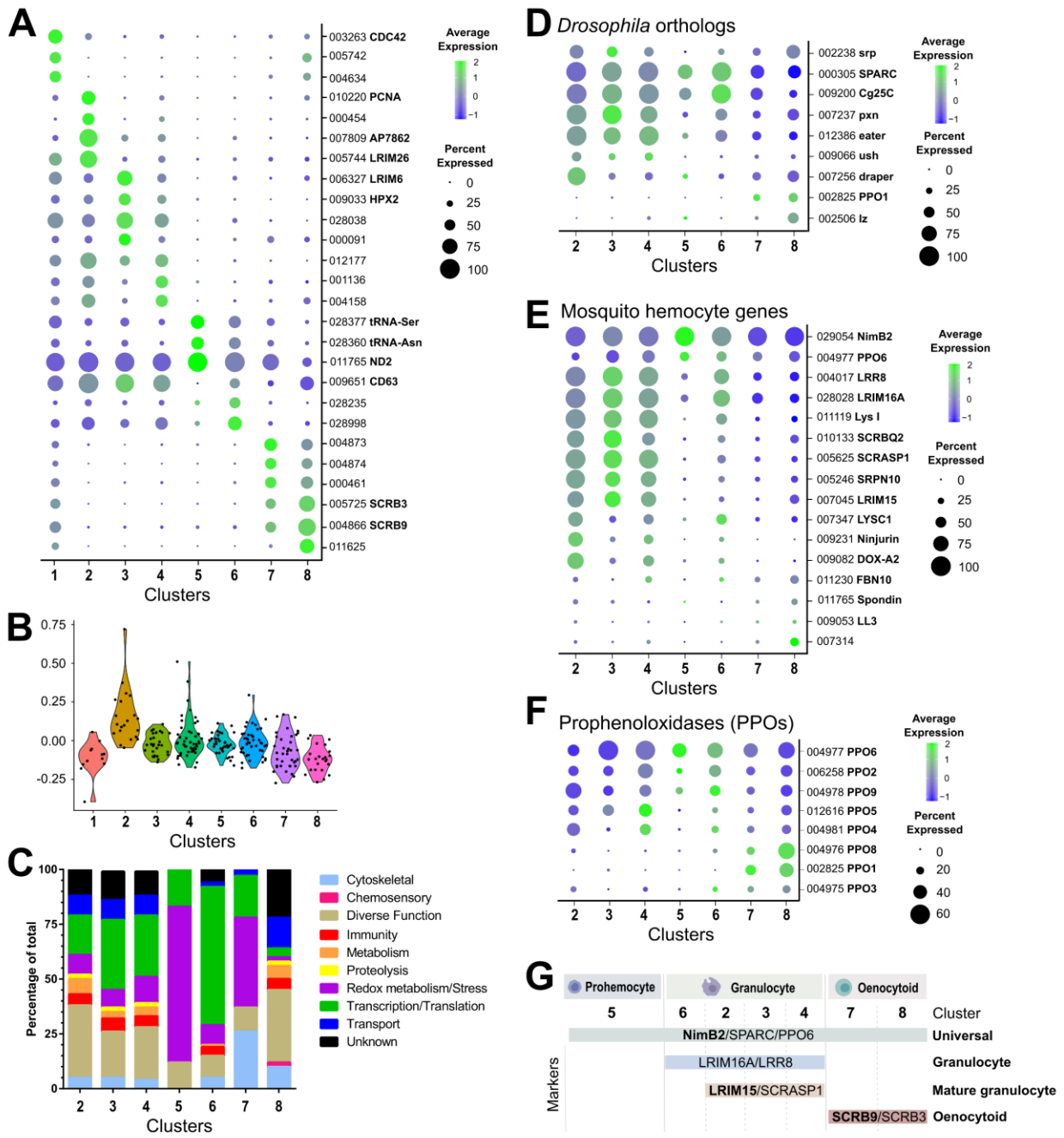

Figure 2. **Comparative analysis of mosquito immune cells. (A)** Marker gene expression displayed by dot plot across cell clusters. Dot color shows levels of average expression, while dot size represents the percentage of cells expressing the corresponding genes in each cell cluster. **(B)** Violin plot of cell cycle genes (GO::0007049) by cluster, displayed as the difference in average gene expression levels between the cell cycle gene set and random control genes. Positive numbers indicate higher levsl of

cell cycle gene expression compared to the random set for that cluster. **(C)** Gene ontology (GO) analysis of genes expressed in >80% cells within each respective cluster. Heat maps of candidate genes to *Drosophila* hemocyte orthologs **(D)**, described mosquito hemocyte genes **(E)**, or *An. gambiae* prophenoloxidases (PPOs) **(F)** to enable the characterization of immune cells form each cell cluster. **(G)** From these analysis, immune cells cluster were assigned to tentative cell types (prohemocytes, granulocytes, oenocytoids) based on the expression subtype specific marker expression. Genes in bold are featured prominently in our downstream analysis.

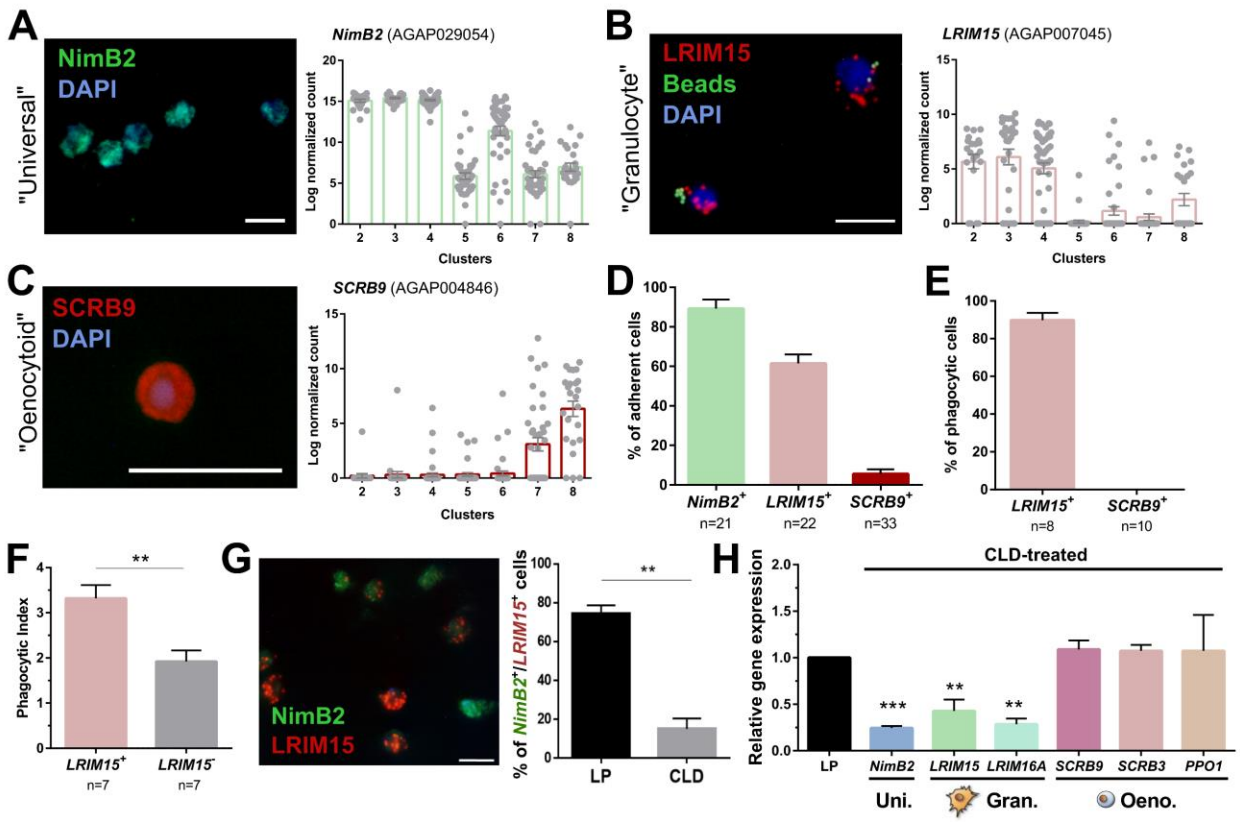

Figure 3. **Definition of mosquito immune cell subtypes.** RNA-FISH and gene expression profiles across cell clusters for the "universal" marker, *NimB2* **(A)**, the "granulocyte" marker, *LRIM15* **(B)**, and "oenocytoid" marker, *SCRB9* **(C)**. The percentage of adherent cells following fixation was evaluated for each of the respective *NimB2*, *LRIM15*, and *SCRB9* markers in four or more independent replicates **(D)**. To determine the phagocytic ability of granulocytes and oenocytoids, the uptake of fluorescent beads was evaluated in either LRIM15$^+$ or SCRB9$^+$ cells in two independent replicates **(E)**. LRIM15$^+$ cells display a higher phagocytic index (# beads engulfed per cell) than LRIM15$^-$ cell populations **(F)**. Data were analyzed using a Mann–Whitney test, with bars representing mean ± SE of two independent replicates. The phagocytic ability of LRIM15$^+$ cells was further validated by examining the abundance of NimB2$^+$/LRIM15$^+$ cells by RNA-FISH following perfusion after treatment with control (LP)- or clodronate (CLD) liposomes that deplete phagocytic cells **(G)**. Data were analyzed using a Mann–Whitney test. Bars represent mean ± SE of two independent replicates. Additional validation of clodronate (CLD) depletion of phagocytic cells was performed by qRT-PCR using primers

for universal (uni.), granulocyte (gran.), and oenocytoid (oeno.) cell markers. Data were analyzed using an unpaired t test to determine differences in relative gene expression between LP and CLD treatments. Bars represent mean ± SE of three independent replications (**H**). Asterisks denote significance (**$P < 0.01$, ***$P < 0.001$). Scale bar, 10 µm.

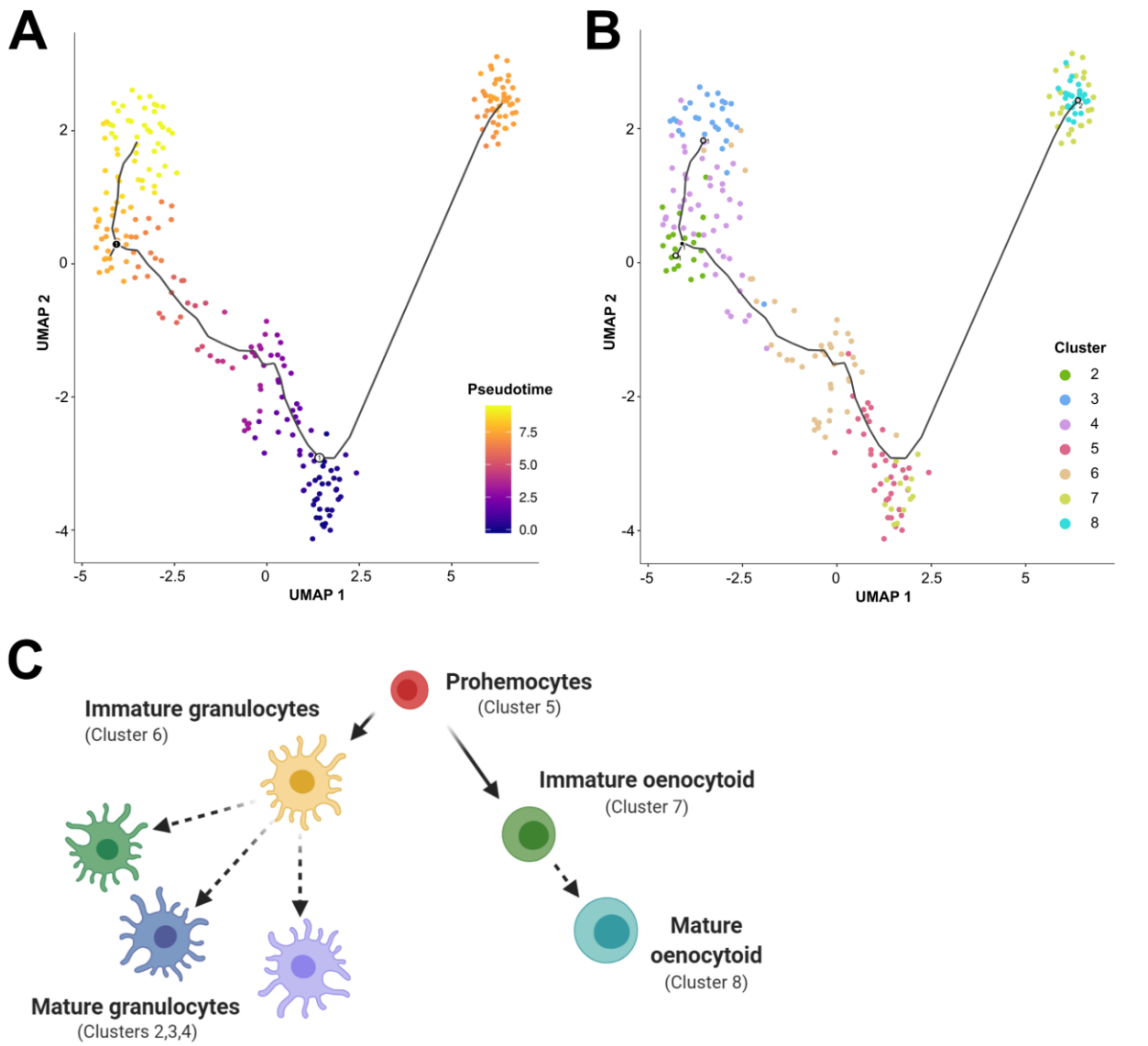

Figure 4. **Lineage analysis of mosquito immune cells.** Using Monocle3, mosquito immune cells were visualized by UMAP to reveal two distinct lineages in pseudotime **(A)** or to display cells of each respective immune cell cluster **(B)**. Based on the lineage analysis, gene expression, and other functional assays, our data support the following model of immune cell development and differentiation where prohemocytes serve as precursors for the granulocyte and oenocytoid lineages **(C)**. Each cell type is labeled with the corresponding cell cluster described in our analysis. Figure was created with BioRender.com.

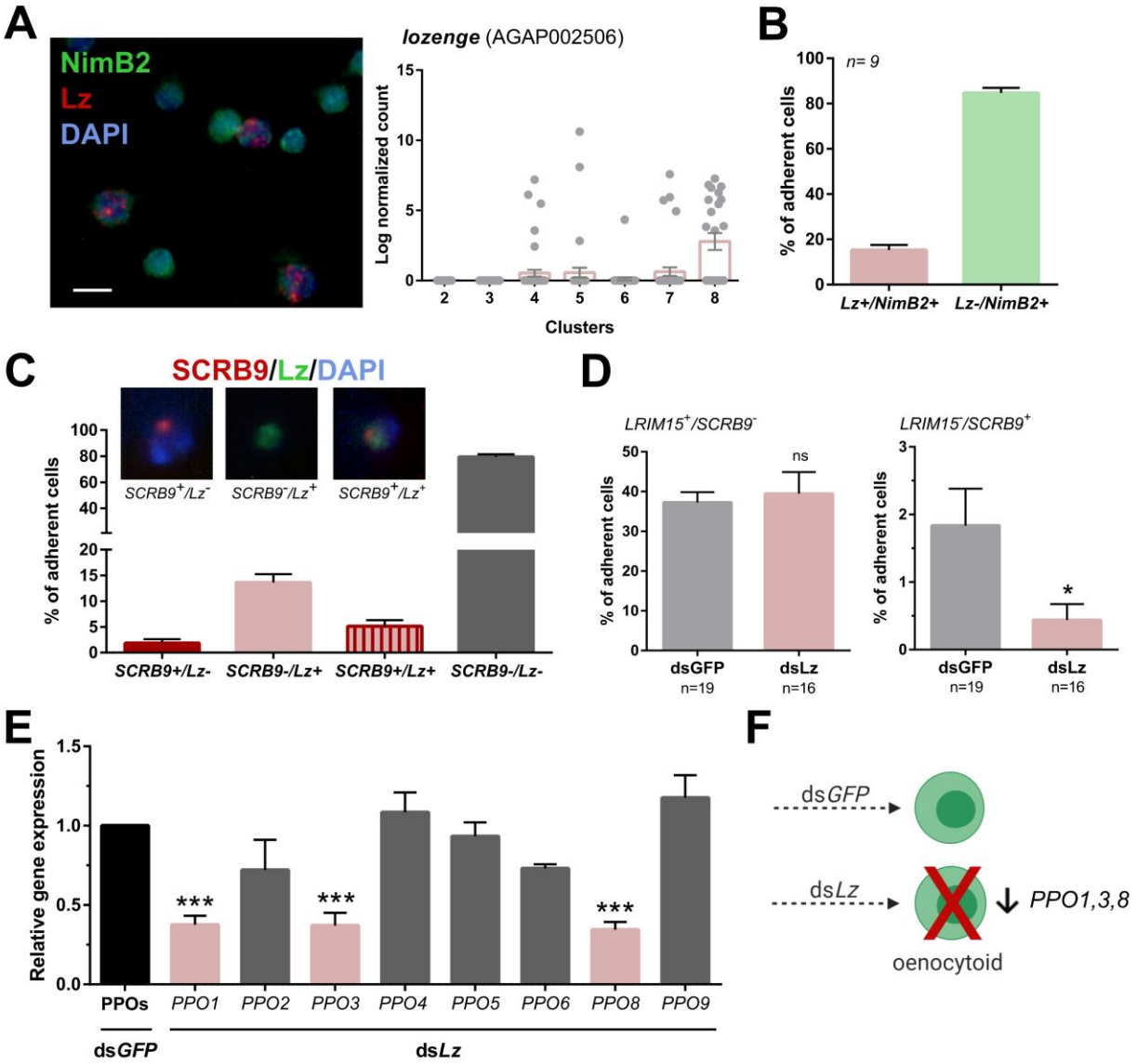

Figure 5. Lozenge promotes oenocytoid differentiation. RNA-FISH and gene expression profiles across cell clusters for *lozenge (Lz)* **(A)**. Scale bar, 10 µm. The percentage of adherent *Lz*$^{+}$*/NimB2*$^{+}$ or *Lz*$^{-}$*/NimB2*$^{+}$ cells were examined in naïve adult female mosquitoes to estimate cell abundance **(B)**. Data were collected from two independent experiments. To more closely examine the population of *Lz*+ cells, RNA-FISH experiments were performed double staining for *Lz* and the oenocytoid marker, *SCRB9* **(C)**. The percentage of fixed cells positive for one, both, or neither marker is displayed with representative images. Data are summarized from two independent experiments. To determine the effects of Lz on immune cell populations, the abundance

of *LRIM15*$^{+}$*/SCRB9*$^{-}$ (granulocyte) and *LRIM15*$^{-}$*/SCRB9*$^{+}$ (oenocytoid) cells were evaluated by RNA-FISH after *GFP* (control)- or *Lz*-silencing **(D)**. Data represent the mean ± SE of three independent replicates. Significance was determined using Mann-Whitney analysis and is denoted by an asterisk (*$P < 0.05$); ns, not significant. Since Lz expression has previously been associated with prophenoloxidase (PPO) expression, the expression of all 8 genes identified in our scRNA-seq analysis were examined by qRT-PCR in GFP (control) - *Lz*-silenced mosquitoes **(E)**. Data represent the mean ± SE of three or more independent replicates and were analyzed by a One-way ANOVA and Holm-Sidak's multiple comparison test using GraphPad Prism 6.0. Asterisks denote significance (***$P < 0.001$). **(F)** Summary of *Lz*-silencing experiments which display a reduction in oenocytoid numbers and a specific sub-set of PPO gene expression which support that Lz is integral to the differentiation of the mosquito oenocytoid lineage.