

# Taking Insect Immunity to the Single-Cell Level

Sarah Hélène Merkling, Louis Lambrechts

## ▶ To cite this version:

Sarah Hélène Merkling, Louis Lambrechts. Taking Insect Immunity to the Single-Cell Level. Trends in Immunology, 2020, 41 (3), pp.190-199. 10.1016/j.it.2020.01.002 . pasteur-02889446

# HAL Id: pasteur-02889446 https://pasteur.hal.science/pasteur-02889446

Submitted on 3 Jul 2020  $\,$ 

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers. L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution - NonCommercial - ShareAlike 4.0 International License

1	Taking insect immunity to the single-cell level
2	
3	
4	Sarah Hélène Merkling <sup>1*</sup> and Louis Lambrechts <sup>1*</sup>
5	
6	<sup>1</sup> Insect-Virus Interactions Unit, Institut Pasteur, UMR2000, CNRS, Paris, France
7	
8	
9	*Correspondence to: Sarah Merkling (sarah.merkling@pasteur.fr)
10	
11	Keywords: Insect immunity, single-cell sequencing, drosophila, mosquito, infection
12	modeling

## 13 Abstract (120 words)

14

For decades, insect immunology has contributed groundbreaking discoveries on the 15 intricacies of innate immunity. It has profoundly impacted our understanding of innate 16 immunology in mammalian systems and improved human therapeutic interventions, 17 18 from the composition of vaccines to cell-based immune therapies. Our current 19 knowledge of insect immunity mainly encompasses detailed molecular mechanisms and 20 systemic responses to pathogen infection. However, the organs and specific cell 21 populations involved in immune responses remain elusive. Here, we highlight the need 22 for a better understanding of insect immune responses at the level of organs and cell 23 populations. Not only will it improve our current understanding of tissue or cell-specific 24 immune processes across species, but it will also pave the way for spatial modeling of 25 within-host infection dynamics.

### 26 (Re)-discovering insect immunity using single-cell technologies

27

28 Over the last few decades, intense research efforts exploiting the power of the model 29 organism Drosophila melanogaster have led to fruitful discoveries in the field of insect 30 immunology and impacted the field of innate immunity at large [1]. Innate immune 31 responses constitute the first line of defense against microbial agents and are defined 32 by their rapid activation, efficiency, and lack of memory. Pathogen recognition is 33 mediated by germline-encoded receptors recognizing pathogen-associated molecular 34 patterns (PAMPs). Innate immunity is often distinguished from adaptive immunity, 35 which emerged approximately 500 million years ago in vertebrates. Adaptive immune 36 responses are based on pathogen recognition by highly specific antibodies generated 37 by somatic gene rearrangements and long-term protective memory to counter 38 pathogen re-exposures [2].

39

40 Detailed insights into innate antibacterial, antifungal, and antiviral immunity have been 41 obtained in Drosophila, notably because of the vast genetic toolbox available in this 42 organism. Early studies focused on the mechanisms of pathogen recognition and 43 subsequent signal transduction. The discovery of an inducible antibacterial immune response in the fruit fly, first reported in 1972, led to the discovery of antimicrobial 44 45 peptides (AMPs) that are not only conserved from insects to mammals, but are also 46 controlled by highly conserved nuclear factor-kB (NF-kB) signaling cascades [3]. In 47 1997, the discovery of the Toll receptor as a key factor in *D. melanogaster* antifungal 48 immunity [4] was a turning point for the field, as this signaling pathway turned out to be 49 homologous to the myeloid differentiation primary response protein 88 (MYD88)dependent Toll-like receptor (TLR) pathway in mammals [1]. 50

51

52 Since those early days, critical molecular mechanisms for pathogen recognition have 53 been elucidated, and the coordination of immune responses at the organismal level 54 have been largely characterized [5]. Briefly, in insects (*Drosophila sp*)., humoral 55 immunity is controlled by a set of highly conserved signaling pathways. The Toll and 56 the Immune deficiency (IMD) pathways mainly control bacterial and fungal infections, 57 and the Janus kinase-signal transducer and activator of transcription (Jak-Stat) and 58 RNA interference (**RNAi**) pathways are primarily involved in antiviral defense [6]. In 59 reality, these pathways are **pleiotropic**, and their roles against a specific class of 60 pathogens are not always clearly distinguishable. Toll and IMD pathways also 61 participate in antiviral immune responses [6], and the Jak-Stat pathway mediates the 62 response to tissue damage that occurs upon bacterial, fungal, and parasitic infections 63 [7].

64

65 While humoral immunity is relatively well-characterized, cellular immunity in insects (Drosophila sp.) remains more enigmatic. The best characterized immune cells are 66 67 hemocytes, which circulate in the insect hemolymph and differentiate in four main 68 subtypes: plasmatocytes, granulocytes, lammelocytes, and oenocytoids [8]. 69 Our understanding of the immune pathways activated in those cells remains a black 70 box. Moreover, little is known about their movement and communication within immune 71 organs, somewhat due to the technical challenges linked to their study. In fact, 72 harvesting hemocytes from an insect's body cavity requires great technical skills [9] and 73 their number ranges from ≈500 to 3,000, depending on the age, genetic strain, 74 developmental stage, and infection status of the specific insect species [10, 11].

75

Beyond immune cells, immune organs are also poorly characterized. In adult insects, the fat body tissue has, in addition to its metabolic function, a central role in immunity [5]. It is a secretion site for antimicrobial peptides, but also a central signaling hub for the activation of cellular immunity [12]. In *Drosophila*, the fat body activates hemocytes, through a yet to be discovered mechanism that is Toll-dependent [13]. The exact cellular composition of the fat body remains unknown, as is the identity of the cells that can communicate with hemocytes, or coordinate immune processes.

83

In addition to *Drosophila*, non-model insect species are more challenging to study but also medically more relevant. This review mainly highlights recent studies in mosquitoes (*Anopheles* and *Aedes* genera) that contributed seminal discoveries to the field. The completion of genome sequencing and assembly in important vector species, as well as the advent of **RNAi-mediated knock-down** and **CRISPR-Cas9** technologies, have been instrumental in the study of mosquito immunity [14-16]. Comparative genomics 90 reveal a high degree of conservation of immune signaling pathways, combined with 91 specific adaptations to particular pathogens or lifestyles. Overall, the core signaling 92 pathways are well conserved between Drosophila and mosquitoes, whereas upstream 93 signaling factors and downstream effectors diverged during evolution [17, 18]. The 94 opinions expressed in this paper are mostly based on work in Drosophila and 95 mosquitoes but are widely applicable to the study of ticks (Ixodes genus), sandflies 96 (Phelobotominae subfamily) and further arthropods that are the focus of remarkable 97 research efforts (Box 1).

98

99 Single-cell technologies encompass a vast array of methods to study genomics, 100 epigenomics, proteomics and metabolomics at the single-cell level [19-21]. Here, we 101 focus on single-cell RNA sequencing (scRNA-seq) and single-cell transcriptomics, 102 which are on their own, revolutionary and paradigm-shifting. In 2009, the first study 103 using single-cell transcriptome analysis based on a high-throughput sequencing 104 platform led to the characterization of cells from early-stage murine embryos [22]. 105 Several protocols have been developed over the last decade, leading to an impressive increase in throughput and data quality [21]. Historically, for many technical reasons, 106 107 transcriptomic studies have typically been performed on large groups of cells originating 108 from multiple individuals, multiple tissues, or pools of individual or tissues, therefore 109 referred to as **bulk RNA-seq**. Hence, several levels of biological variation, including 110 organismal, tissular and cellular differences, could be "hidden" in one transcriptome 111 data set (Figure 1). Single-cell transcriptomics allow for the first time, to study cell-to-112 cell transcriptional heterogeneity in insects, including the gene expression stochasticity 113 that was first reported over 15 years ago [23].

114

How can single-cell technologies help to accomplish a leap forward for studies of insect immune systems? Could these technologies enable, in the near future, the acquisition of a more comprehensive picture of insect immunity? Can they integrate into the same model, our knowledge of molecular mechanisms and systemic responses by bridging those components with a better knowledge of cellular and tissular immunity? In this opinion article, we argue that the increasing availability and applicability of scRNA-seq

- 121 present a new opportunity to deepen our knowledge of insect immunity and impact our
- 122 global knowledge of innate immunology across species.
- 123

# How can single-cell transcriptomics be instrumental to the study of insect immunity?

126

## 127 Deciphering the cellular composition of organs

128

129 The cellular composition of organs, including immune sites, remains to be investigated 130 in insects. Early studies based on cellular morphology helped decipher the composition 131 of tissues and led to the classification of hemocytes [24] and digestive tract cells [25]. 132 However, a recent study revealed that Anopheles gambiae mosquito hemocytes display 133 a high level of complexity, given that two highly distinct cellular "populations" -- as 134 defined by their transcriptional profile -- actually share morphological and functional 135 features, such as their phagocytotic capacity [26]. Therefore, the classification of cell 136 types historically based on morphology or function might need to be revisited. In 137 addition, little is known about the cellular re-composition of organs after an infectious 138 episode. In Drosophila, infection of the intestinal tissue with the Gram-negative 139 bacterium Erwinia caratovora has resulted in dramatically increased stem cell 140 proliferation and epithelial renewal, as shown by the increase of mitotic recombination in 141 the intestinal tissue, accompanied by a higher number of stem cells (expressing the 142 escargot marker gene) upon infection [27]. However, it remains to be determined 143 whether the cellular content of the tissue pre and post pathogen exposure is 144 comparable, or if infection-induced changes of immune site composition could impact 145 subsequent immune responses. In fact, *Plasmodium berghei* (rodent malaria)-146 challenged mosquitoes constitutively secrete a hemocyte differentiation factor into their 147 hemolymph that induces **immune priming** when experimentally transferred into naive 148 mosquitoes [28]. The hemocyte differentiation factor was first discovered in mosquito 149 hemolymph by mass spectrometry, and is a complex consisting of a protein (called 150 Evokin) that is a lipid carrier of lipoxins [28]. Immune priming is characterized by an 151 increased number of granulocytes (a subpopulation of hemocytes) and can be 152 measured by extracting hemolymph from mosquitoes and counting the number of cells 153 present while assessing their morphology and binding properties [29]. P. berghei-154 challenged mosquitoes mount a strong immune response when the parasite breaks 155 through the gut and, by doing so, brings the gut bacteria in contact with injured 156 epithelial gut cell. This stimulation creates a state of systemic immune surveillance upon 157 which the immune system is primed to respond again to a new exposure. Upon re-158 infection, the immune system mounts a stronger antibacterial response that is also 159 detrimental to the survival of *Plasmodium sp.* parasites [29]. It will be interesting in the 160 future to analyze the transcriptome of these cell populations at the single-cell level and 161 understand the depth of immune priming processes. In fact, this could yield new 162 insights into the mechanisms of immune memory that are so far poorly understood in 163 insects.

164

### 165 Discovering new immune cell populations

166

167 In addition to discovering new cell populations that are induced by infection, single-cell 168 sequencing can also enable identification of cells that hold an immune function based 169 on their transcriptional similarity with existing immune cells, but have not yet been 170 characterized as such. In fact, it is likely that most tissues contain cells that mediate 171 their defense and viability upon infection but are rare, non-canonical, or difficult to 172 characterize. For instance, the immune capacity of glial cells in the Drosophila brain was 173 only discovered recently, with the observation that they were able to secrete AMPs [30]. 174 Additionally, infection of nervous tissues such as the brain by Zika virus has recently be 175 modeled in *Drosophila* and pointed to a role for the IMD pathway in glial cells [31]. In 176 fact, mutant flies that have a deficient IMD signaling failed to induce autophagy, an 177 antiviral process that helps reducing Zika virus infection [31]. Single-cell sequencing, 178 which is virtually applicable to any tissue or organs, is likely to contribute exciting 179 discoveries of new immune cell types in insects that were so far unreachable using 180 traditional methods.

181

182 Assessing heterogeneity of immune responses to infection

184 Recent studies on mammalian cells such as the human lung epithelial carcinoma line 185 A549, the human hepatoma line (Huh7), or the murine fibroblast L929 line, utilized 186 scRNA-seq to assess the diversity of the transcriptional responses to infections by 187 influenza, dengue, Zika and West Nile viruses [32-35]. Taken together, these studies 188 revealed astonishing heterogeneity at the level of host transcriptomes and viral 189 abundance from cell to cell. The measurement of both host and viral RNA transcripts 190 (also called visc-RNAseq) in the same cells can reveal correlations between gene 191 expression, and make possible the detection of strong pro- or antiviral candidate genes 192 [33]; indeed, this was previously done in a bulk RNA-seg study on dengue virus-193 infected midguts of Aedes aegypti mosquitoes, that identified the sterol regulatory 194 element-binding protein (SREBP) as a host factor with a proviral role in the mosquito 195 midgut [36]. Moreover, A scRNA-seg study performed on peripheral blood mononuclear 196 cells isolated from dengue virus-infected human patients identified a set of genes that 197 might be used as biomarkers for disease progression to severe dengue fever [37]. For 198 example, the expression of MX2 within naive B cells is strongly upregulated before the 199 onset of severe dengue fever [37]. The predictive power of such markers discovered by 200 visc-RNAseq remains to be tested with larger study cohorts. Finally, quantifying the 201 heterogeneity of immune responses to infection in relation to pathogen presence might 202 also help identify by-stander effects of cells near infection sites that are not infected 203 themselves, but which deploy immune responses linked to the detection of infection-204 linked signals in their environment. For example, in Drosophila, reactive oxygen species 205 are released from necrotic cells and act as an immediate danger signal which can 206 induce he recruitment of hemocytes to the injured tissue [38]. Indeed, it will be 207 interesting to compare immune profiles of infected cells versus uninfected cells that 208 have been alerted of an ongoing infection by danger signals in their intra-host 209 environment.

210

211 Assessing heterogeneity in pathogen genetic diversity upon infection

212

213 Pathogen-inclusive scRNA-seq opens possibilities to assessing pathogen genetic 214 diversity such as **viral quasispecies** at the intra-cellular level, as has already been 215 done for hepatitis C virus, human immunodeficiency virus and vesicular stomatitis virus 216 [39]. In insect vectors of viruses, it will be possible to identify cellular subtypes that are 217 responsible for increasing viral diversity, or on the contrary, act as population 218 bottlenecks. In fact, the organs and cell types that are sites of viral replication and 219 enable the virus to genetically diversify or homogenize, are not characterized despite 220 being important targets for future vector-control strategies. Intra-host viral diversity can 221 impact viral fitness, and perhaps the transmission efficacy of a given virus population, 222 as well as potentially, the epidemiology of an outbreak [40, 41]. For example, it was 223 shown that the fitness of West Nile virus decreased as it spread from one mosquito 224 organ to another: the virus replicated poorly in avian cells after a passage through Culex 225 quinquefasciatus mosquitoes [41]. Single-cell technologies might help identify cellular 226 sites and molecular mechanisms of viral diversity within the insect host and potentially 227 support genetic interventions aiming at reducing pathogen transmission.

228

229 Beyond viral infections, a recent scRNA-seq study that focused on *Plasmodium* sp. 230 parasites revealed discrete transcriptional signatures during the parasite's life cycle, as 231 well as unexpected sex-specific roles for contingency gene families (var in P. falciparum 232 and *pir* in *P. berghei*) that are known for enabling parasite adaptation to new 233 environments [42]. In addition, scRNA-seq enabled the Malaria Cell Atlas to characterize 234 thousands of single-parasite transcriptomes across the entire *Plasmodium* sp. life cycle 235 [43]. Those studies are relevant as they may open multiple new research avenues --236 from the analysis of nonculturable life stages of *Plasmodium* species (such as *P. vivax*), 237 to the analysis of cell-to-cell variability in response to drugs and vaccines, as well as the 238 development of blocking strategies against infection in humans. Overall, pathogen-239 inclusive scRNA-seq offers exciting research avenues that will allow to assess the 240 impact of the host cells on pathogen evolution and adaptation.

241

242 What are the main challenges of single-cell sequencing implementation in 243 insects?

244

245 Dissociation and fixation

247 The first step of any single-cell sequencing experiment consists of the preparation of the 248 cell suspension. With the exception of hemocytes that circulate in the insect's 249 hemolymph, cells are embedded in tissues and organs that have to be dissociated in 250 order to obtain a suspension of single cells. Dissociation methods are varied, ranging 251 from the use of a diverse set of enzymes to the use of mechanical force (Table 1). 252 Most often, several methods need to be tested, as their efficiency is often dependent on 253 the cell type, tissue, organism, and insect species. Dissociation is an aggressive 254 process that causes cellular stress and therefore transcriptional changes related to 255 stress responses. Accordingly, if feasible, it is preferable to fix cells and thus preserve 256 the transcriptome before dissociation. Several fixation agents are available 257 (paraformaldehyde, acetone, methanol, Lomant's reagent, RNAlater™, vivoPHIX™); 258 some have been tested in side-by-side studies with live cells and do not seem to bias 259 or alter the quality of the transcriptome [44, 45]. In all cases, a set of quality indicators 260 should be measured: (i) dissociation efficacy, (ii) cell viability and (iii) cell integrity (Box 261 2). Importantly, dissociation methods can potentially affect the transcriptome of cells or 262 a subset of cells [46]. It was recently shown that a subset of cells detected in several 263 single-cell studies, for example in mouse muscle cells or zebra fish osteoblasts, was 264 over expressing genes associated with tissue injury (i.e. heat shock proteins such as 265 Hspa1a and Hspa1b). However, those cells were not found as a functional subset in 266 vivo, pointing to a likely experimental artifact [46]. Noteworthy, these 'injured' cells might 267 also be present as potential contaminants in bulk RNA-seq studies. However, such 268 transcriptional signals would be harder or likely impossible to identify, as they would be 269 averaged at the level of an entire tissue or individual. By contrast, In a single-cell study, 270 'injured' cells could be identified and excluded from subsequent analyses.

271

Insect cells present additional challenges at the dissociation step due to their limited number and reduced size. Insect cells are often smaller in size than their mammalian counterparts [47, 48]. For instance, *Drosophila* neurons are ten times smaller than mammalian neurons [49]. Moreover, insect organs and tissues are small in size. The *Drosophila* brain contains approximately 100,000 cells [48], while the human brain contains 100 billion cells [50]. While it is possible to dissect high numbers of organs for single-cell experiments in insects, never will it be possible to reach the order of magnitudes of larger organisms. This dramatically complicates the execution of singlecell experiments, not only because of the sheer number of cells available but also because cell recovery after dissociation and cell viability have to be extraordinary. Once dissociation and potential fixation are optimized, RNA sequencing can take place. Apart from cellular sizes that might not be adapted to all microfluidic platforms, which are generally designed for mammalian cells, no specific considerations need to be taken for the sequencing of insect cells.

- 286
- 287
- 288
- 289 Data analysis and interpretation

290

291 Amongst the significant applications of scRNA-seq is the identification of novel cell 292 types and states, founded on the detection of co-expression patterns and clustering of 293 cells based on their transcriptomic similarity. Cell clusters can be analyzed further by 294 identifying gene sets characteristic of each cluster, also called markers. Whereas 295 markers for mammalian immune cells are generally well known, they remain elusive for 296 most insect cells. Cluster analysis and marker discovery is, therefore, more challenging 297 for insect cells, as preliminary knowledge about sub-groups to be detected is scarce. 298 Recently, an unbiased classification of An. gambiae hemocytes was performed using a 299 combination of scRNA-seq, high-content imaging flow cytometry, and single-molecule 300 RNA hybridization [26]. It revealed an unexpected degree of complexity in the hemocyte 301 population together with the existence of cellular cross-talk via exchange of mRNA 302 This combination of required tools highlights the need for between cells [26]. 303 orthogonal validation methods, especially in non-model insect systems for which there 304 is little preconceived knowledge about cell populations, and almost no cell markers 305 available. Overall, scRNA-seg data analysis and interpretation is challenging for insect 306 cells and, given the poor availability of markers for many cell populations, will require 307 cross-validation with other experimental methods.

308

309 How can single-cell sequencing data be exploited to model immune310 responses in insects?

312 Single-cell sequencing data sets are comparatively more complex and multi-313 dimensional then classical bulk RNA-seq data sets. Beyond the standard analysis 314 methods, consisting of reducing dimensional space (by **PCA** and **t-SNE**), functional 315 clustering or trajectory analysis [21], even more sophisticated data analysis methods 316 could help informing insect immunity.

317

318 Subjecting multiple tissue samples over an infection's time course to scRNA-seq could 319 help inform infection dynamics in a deeper manner than has ever been possible. Each 320 data set might describe degree of infection, correlating it with transcriptional responses 321 at the cellular level in different tissues over time. Given the high dimensionality of such 322 data, integration into a mathematical modeling framework is highly desirable. Such 323 information might help identify key immune cells and processes, as well as model the 324 progression of the pathogen through the host's body. Once built, such a modeling 325 framework could ideally be used to predict infection outcomes, severity, or further 326 pathogen transmission potential.

327

For instance, when an Aedes mosquito bites a human infected with an arbovirus such 328 329 as Zika virus, the bloodmeal containing the virus is first ingested in the mosquito midgut. 330 The virus initially replicates in midgut cells before disseminating throughout the 331 mosquito's body. Eventually, the virus reaches the salivary glands and can be 332 transmitted again through a new bite [51]. The exact progression of the virus 333 throughout the mosquito's body remains elusive. How does the virus circulate in the 334 hemolymph? Is it carried by hemocytes or does it circulate freely? Which organs does 335 the virus replicate in, and when? Are similar immune pathways triggered in the midgut, 336 fat body, and salivary gland? Can we predict the transmission potential of a certain virus 337 based on the host immune response and viral levels in certain organs? Such 338 fundamental questions might be answered, at least in part, by the use of single-cell 339 technologies and modeling of infection dynamics.

340

### 341 Concluding remarks

For decades, the study of innate immune processes in insects has greatly contributed to our overall knowledge of innate immunology. Here, we propose that the recent developments of single-cell transcriptomics represent an opportunity to make a significant leap forward in this exciting research field. Improving our knowledge of immunity at the cellular and tissular level has the potential to be paradigm shifting at the fundamental level (see outstanding questions). Indeed, it may also have a wider impact on novel strategies deployed to control vector-borne diseases.

350

Nowadays, performing single-cell transcriptomics remains challenging technically, but also financially. Although the protocols and technologies evolve rapidly, the costs of the technology remain high. Hopefully, the constant decrease of sequencing costs and commercialization of scRNA-seq by more industrial actors will make it more accessible.

355

356 Spatial transcriptomics are a recent and exciting technological development [52]. In 357 addition to obtaining single-cell transcriptomes, this technology allows to map the 358 transcriptional activity to its tissular location and place it in a morphological context. In 359 fact, it relies on the deposition of tissue slices on arrays covered with probes that 360 capture RNA. After sequencing of the array, the transcriptomes are reconstituted in 361 space and mapped to their location of origin on the tissue slice [53]. This will help 362 identification of immune cells, infected cells, by-stander cells, rare cell types on a 3D 363 map, and may contribute to a more holistic understanding of immune responses at a 364 cellular, tissular and organismal level.

365

366 Here, we focused on the promise of single-cell transcriptomics to improve our 367 knowledge of insect immunology, without mentioning the development of other single-368 cell based technologies. Performing DNA sequencing at the level of single cells can also 369 help identify cell-to-cell heterogeneity in single-nucleotide polymorphisms and copy 370 number variations. Moreover, DNA mosaicism remains an unexplored territory in 371 insects although it has been extensively investigated in the context of human health 372 [54]. Single-cell proteomics and epigenomics have also been developed more recently, 373 offering multiple and diverse research possibilities in the near future [19, 20].

375 Finally, we posit that single-cell transcriptomics will need integration at the systems 376 level, i.e. using a holistic approach to capture emerging properties of cells and tissues 377 functioning as a system. This will be possible with the complementary single-cell based 378 technologies aforementioned, and more generally, with various approaches of 379 genomics, proteomics, epigenomics, metabolomics and lipidomics [55]. For now, it 380 remains challenging to perform experiments and analyze samples in such a multi-381 dimensional manner for technical and financial reasons. However, we predict that within 382 the next decade, the development of sophisticated computational methods will enable 383 seamless data integration and analysis at a systems level. Alexander von Humboldt was 384 the first scientist to suggest that, while each element in nature demanded specialized 385 analytical attention in order to establish its own specific rules of order and organization, 386 the most important challenge was that of integrating all this knowledge. With his 250<sup>th</sup> 387 birthday, his vision remains as relevant and applicable to systems biology in the 21<sup>st</sup> 388 century.

- 389
- 390
- 391
- 392

#### 393 Acknowledgements

394 We thank members of the Lambrechts lab and Silvain Pinaud, Virginia Howick, Mara 395 Lawniczak, Gianmarco Raddi, Andrew Goldsborough, Alex Perkins and John Huber for 396 insightful discussions. This work was supported by a Pasteur-Roux fellowship to S.H.M. 397 L.L. is supported by the French Agence Nationale de la Recherche (grants ANR-16-398 and ANR-18-CE35-0003-01), and the French Government's CE35-0004-01 399 Investissement d'Avenir program Laboratoire d'Excellence Integrative Biology of 400 Emerging Infectious Diseases (grant ANR-10-LABX-62-IBEID).

- 401
- 402
- 403
- 404
- 405
- 406

- Figure legend Figure 1. Schematic overview of the advantages of scRNA-seq over bulk RNA-seq. Over the last few decades, RNA sequencing technologies have evolved from using
- 425 pools of individual insects, to single individuals, to single organs. At every step, the 426 transcriptome obtained is an "averaged" picture of gene expression. It reflects several 427 individuals, several organs, or several cell types. Only using single-cell sequencing is it 428 possible to get a unique cellular transcriptome.
- **Table 1**. Summary of the published work applying scRNA-seq to insect cells.

Organism	Cell type	Dissociation method	Fixation method	Sequencing method	Reference
D. melanogaster	Total embryonic cells	Douncer homogenizer (mechanical)	Methanol	Drop-Seq	[56]
D. melanogaster	Olfactory projection neurons	Papain Liberase™	/	Smart-Seq2	[57]
D. melanogaster	Eye disc cells	Collagenase Trypsin	/	Drop-Seq	[58]
D. melanogaster	Total brain cells	Dispase Collagenase I	/	Chromium Smart-Seq2	[48]

		Trypsin		Cel-Seq 2 ATAC-seq	
D. melanogaster	Midbrain cells	Papain Collagenase I	/	Drop-seq	[49]
A. gambiae	Blood cells	/	/	Smart-Seq2	[26]

432 Box 1. scRNA-seq for studying immunity in arthropod vectors beyond mosquitoes.

433 The majority of published work describing the use of scRNA-seg in insects has been 434 performed in Drosophila sp. and mosquitoes (Table 1). However, scRNA-seq is a 435 technology that is applicable to arthropod vectors beyond mosquitoes, such as fleas 436 (Siphonaptera order), ticks (Ixodes genus), tsetse flies (Glossina genus) and sandflies 437 (Phlebotominae subfamily). Most current knowledge on insect immune signaling 438 pathways has been acquired in Drosophila, and served as a roadmap for studying immunity in other arthropods. However, recent studies have revealed that immune 439 440 signaling pathways in.... are mechanistically more distinct than anticipated, 441 emphasizing the need for in-depth studies in other arthropods [59, 60]. For example, ticks have a smaller repertoire of Toll receptors compared to Drosophila [61] and have 442 443 also undergone gene losses in the IMD pathway, despite both pathways remaining 444 functional [62, 63]. This suggests that unidentified immune pathway components might 445 be replacing predicted canonical factors to mediate the recognition of microbial 446 invaders in ticks, although this remains to be further investigated [64]. It might be 447 possible in the future to isolate individual cells in which those pathways are active (e.g. 448 using flow cytometry) and analyze their transcriptome by scRNAseq to find the 449 unknown components that keep the pathway functional. Another recent study [65] 450 analyzed the role of a mutualist bacterial symbiont on the adult immune response of the 451 tsetse fly, Wigglesworthia [65]. It revealed that the symbiont induced the upregulation 452 of an odorant binding protein (Obp6) in the gut of intrauterine tsetse larvae relative to...., 453 which was necessary and sufficient to induce the subsequent production of crystal 454 cells mediating the melanotic immune response [65]. In the future, it might be 455 possible to determine, using scRNA-seq, exactly which cell types secrete Obp6, and 456 which tissues are targeted by this molecule. Beyond these examples, we argue that 457 scRNA-seq will help elucidate fascinating biological processes, and enable in-depth

458	studies in a broad range of arthropods that have so far been understudied because of				
459	the limited toolbox available.				
460					
461	Box 2. Essential questions to ask during sample preparation for scRNA-seq.				
462	How many cells are recovered after dissociation? Is that number consistent?				
463	Is the tissue fully dissociated?				
464	What is the percentage of cell death?				
465	Are the live cells intact and healthy? Are there apoptotic or necrotic cell bodies in				
466	the cellular suspension?				
467	• Are the cells metabolically active during dissociation potentially leading to				
468	changes in the transcriptome?				
469	<ul> <li>Does the cell suspension contain cellular debris to be filtered?</li> </ul>				
470	Are cells recovered after fixation?				
471	Are the cells potentially infectious?				
472	<ul> <li>Do the cells need to be used immediately or can they be stored before use?</li> </ul>				
473					
474	Glossary (450 words)				
475					
476	Antimicrobial peptides: In insects, AMPs are a class of small peptides with potent				
477	antimicrobial activity.				
478					
479	Humoral immunity: Immune responses mediated by factors present in extracellular				
480	fluids, mainly the hemolymph in insects.				
481					
482	Toll pathway: Immune pathway of insects mainly triggered by Gram-positive bacteria				
483	leading to antimicrobial peptide production.				
484					
485	IMD pathway: Immune pathway of insects mainly triggered by Gram-negative bacteria				
486	leading to antimicrobial peptide production.				
487					
488	Jak-Stat pathway: Immune pathway of insects triggered by the Unpaired cytokines				
489	mainly upon viral infection.				

490	
491	RNAi: RNA interference refers to a set of gene-silencing mechanisms induced by
492	double-stranded RNA among which the small interfering RNA (siRNA) pathway is
493	considered a major antiviral defense of insects.
494	
495	Pleiotropic: A gene or pathway is called pleiotropic when it affects two or more
496	seemingly different phenotypic traits.
497	
498	Cellular immunity: Immune responses mediated mainly by hemocytes.
499	
500	CRISPR-Cas9: Prokaryotic immune system based on Clustered Regularly
501	Interspaced Short Palindromic Repeats and the endonuclease Cas9, which was
502	converted into a powerful a gene-editing technique targeting specific sequences
503	recognized by a complementary guide RNA.
504	
505	Bulk RNA-seq: Conventional RNA sequencing technique that probes the
506	transcriptome of a large number of pooled cells.
507	
508	Immune priming: analogous to trained immunity, it describes the phenomenon by
509	which a first pathogen exposure enhances host resistance to a subsequent challenge.
510	
511	viscRNAseq: virus-inclusive RNA sequencing probes the host transcriptome together
512	with intra-cellular viral RNA at the single-cell level.
513	
514	Viral quasispecies: Cloud of diverse virus variants that are genetically linked through
515	mutation and collectively contribute to the characteristics of the viral population.
516	
517	Population bottleneck: In evolutionary biology, it is defined as an event that
518	drastically reduces the size of a population.
519	
520	Viral fitness: Usually equated to replicative fitness, it is defined as the capacity of a
521	virus to produce infectious progeny in a given cellular environment.

- 522
- 523 PCA: Principal component analysis is a linear dimensionality reduction method
- 524 commonly used to identify cell subpopulations by clustering their features in 2D or 3D.
- 525
- 526 t-SNE: t-distributed stochastic neighbor embedding is a dimensionality reduction
- 527 method that helps identifying cell subpopulations and sometimes provides clearer
- 528 visualization of clusters then PCA.
- 529
- 530 RNAi-mediated knock-down: Experimental gene-silencing technique using double531 stranded RNA or siRNAs to neutralize target RNA molecules by RNAi.
- 532
- 533 Plasmatocytes: Macrophage-like cells involved in phagocytosis and encapsulation
- that are the major constituents of the insect hemocyte population.
- 535
- 536 Granulocytes: Adhesive cells mediating phagocytosis, encapsulation and537 degranulation in insects.
- 538
- 539 Lammelocytes: Large cells involved in encapsulation of invading pathogens in540 insects.
- 541
- 542 **Oenocytoids:** Large cells responsible for the production of prophenoloxidase in 543 insects.

- 545 **Crystal cells:** Cells responsible for humoral melanization that accompanies some 546 immune reaction in insects.
- 547
- 548 Pathogen-Associated Molecular Patterns (PAMPs): Diverse set of molecules 549 that share a set of general "patterns," or structures, that alert immune cells to destroy 550 invading pathogens. One example is lipopolysaccharides present in the membrane of 551 Gram-negative bacteria.

552

553 **Toll-like receptor (TLR):** A class of pattern recognition receptors that initiate the 554 innate immune response by sensing pathogen-associated molecular patterns.

556 **Hemocytes:** A group of several cell types that are found in the hemolymph of insects.

557

558 **Copy number variation:** Phenomenon in which the number of copies of a particular

559 gene varies from one individual to the next.

- 560
- 561 **DNA mosaicism**: Phenomenon that involves the presence of two or more populations
- of cells with different genotypes in one individual that is called a mosaic.

- 564 **Melanotic immune response:** A branch of the innate immune response that 565 involves the synthesis of melanin, a dark pigment that can encapsulate invading
- 566 pathogens or unhealthy cells.
- 567
- 568 References
- 569
- 1. Lemaitre, B. and Hoffmann, J. (2007) The host defense of Drosophila
- 571 melanogaster. Annu Rev Immunol 25, 697-743.
- 572 2. Hoffmann, J.A. et al. (1999) Phylogenetic perspectives in innate immunity.
- 573 Science 284 (5418), 1313-8.
- 3. Boman, H.G. et al. (1972) Inducible antibacterial defence system in Drosophila.
- 575 Nature 237 (5352), 232-5.
- 576 4. Lemaitre, B. et al. (1996) The dorsoventral regulatory gene cassette
- 577 sp./§tzle/Toll/cactus controls the potent antifungal response in Drosophila adults.
- 578 Cell 86 (6), 973-83.
- 579 5. Buchon, N. et al. (2014) Immunity in Drosophila melanogaster from microbial
- recognition to whole-organism physiology. Nature Reviews Immunology 14, 796.
- 581 6. Merkling, S.H. and van Rij, R.P. (2013) Beyond RNAi: antiviral defense strategies
- in Drosophila and mosquito. J Insect Physiol 59 (2), 159-70.
- 583 7. Myllymäki, H. and Rämet, M. (2014) JAK/STAT Pathway in Drosophila Immunity.
- 584 Scand J Immunol 79 (6), 377-385.
- 585 8. Fauvarque, M.O. and Williams, M.J. (2011) Drosophila cellular immunity: a story of 586 migration and adhesion. J Cell Sci 124 (Pt 9), 1373-82.
- 587 9. Hiroyasu, A. et al. (2018) Extraction of Hemocytes from Drosophila melanogaster 588 Larvae for Microbial Infection and Analysis. J Vis Exp (135).
- 589 10. Horn, L. et al. (2014) Phagocytic ability declines with age in adult Drosophila 590 hemocytes. Aging Cell 13 (4), 719-28.
- 591 11. Bryant, W.B. and Michel, K. (2014) Blood feeding induces hemocyte proliferation
- and activation in the African malaria mosquito, Anopheles gambiae Giles. J Exp Biol217 (Pt 8), 1238-45.
- 12. Yang, H. and Hultmark, D. (2016) Tissue communication in a systemic immune
- response of Drosophila. Fly (Austin) 10 (3), 115-22.

- 596 13. Schmid, M.R. et al. (2014) Control of Drosophila blood cell activation via Toll 597 signaling in the fat body. PLoS ONE 9 (8), e102568.
- 598 14. Matthews, B.J. et al. (2018) Improved reference genome of Aedes aegypti 599 informs arbovirus vector control. Nature 563 (7732), 501-507.
- 600 15. Neafsey, D.E. et al. (2015) Mosquito genomics. Highly evolvable malaria vectors:
- the genomes of 16 Anopheles mosquitoes. Science 347 (6217), 1258522.
- 16. Kistler, K.E. et al. (2015) Genome engineering with CRISPR-Cas9 in the
- 603 mosquito Aedes aegypti. Cell Rep 11 (1), 51-60.
- 17. Waterhouse, R.M. et al. (2007) Evolutionary dynamics of immune-related genes
- and pathways in disease-vector mosquitoes. Science 316 (5832), 1738-43.
- 606 18. Christophides, G.K. et al. (2002) Immunity-related genes and gene families in607 Anopheles gambiae. Science 298 (5591), 159-65.
- 19. Clark, S.J. et al. (2016) Single-cell epigenomics: powerful new methods for
- understanding gene regulation and cell identity. Genome Biol 17 (1), 72.
- 20. Specht, H. et al. (2019) High-throughput single-cell proteomics quantifies the
- 611 emergence of macrophage heterogeneity. bioRxiv, 665307.
- 612 21. Hwang, B. et al. (2018) Single-cell RNA sequencing technologies and
- 613 bioinformatics pipelines. Exp Mol Med 50 (8), 96.
- 614 22. Tang, F. et al. (2009) mRNA-Seq whole-transcriptome analysis of a single cell.615 Nat Methods 6, 377.
- 616 23. Elowitz, M.B. et al. (2002) Stochastic gene expression in a single cell. Science 617 297 (5584), 1183-6.
- 618 24. Lawrence, P.O. (2008) Hemocytes of Insects: Their Morphology and Function. In
- 619 Encyclopedia of Entomology (Capinera, J.L. ed), pp. 1787-1790, Springer
- 620 Netherlands.
- 621 25. Demerec, M. (1994) Biology of Drosophila, Cold Spring Harbor Laboratory Press.
- 622 26. Severo, M.S. et al. (2018) Unbiased classification of mosquito blood cells by
- single-cell genomics and high-content imaging. Proc Natl Acad Sci U S A 115 (32),
  E7568-E7577.
- 625 27. Buchon, N. et al. (2009) Drosophila Intestinal Response to Bacterial Infection:
- Activation of Host Defense and Stem Cell Proliferation. Cell Host Microbe 5 (2), 200-211.
- 628 28. Ramirez, J.L. et al. (2015) A mosquito lipoxin/lipocalin complex mediates innate 629 immune priming in Anopheles gambiae. Nat Commun 6, 7403.
- 630 29. Rodrigues, J. et al. (2010) Hemocyte differentiation mediates innate immune
- 631 memory in Anopheles gambiae mosquitoes. Science 329 (5997), 1353-5.
- 632 30. Lye, S.H. and Chtarbanova, S. (2018) Drosophila as a Model to Study Brain
- 633 Innate Immunity in Health and Disease. Int J Mol Sci 19 (12).
- 634 31. Liu, Y. et al. (2018) Inflammation-Induced, STING-Dependent Autophagy
- Restricts Zika Virus Infection in the Drosophila Brain. Cell Host Microbe 24 (1), 57-68e3.
- 637 32. Russell, A.B. et al. (2018) Extreme heterogeneity of influenza virus infection in 638 single cells. Elife 7.
- 639 33. Zanini, F. et al. (2018) Single-cell transcriptional dynamics of flavivirus infection.
  640 Elife 7.
- 641 34. Steuerman, Y. et al. (2018) Dissection of Influenza Infection In Vivo by Single-Cell
- 642 RNA Sequencing. Cell Syst 6 (6), 679-691 e4.

- 643 35. O'Neal, J.T. et al. (2019) West Nile Virus-Inclusive Single-Cell RNA Sequencing
  644 Reveals Heterogeneity in the Type I Interferon Response within Single Cells. J Virol
  645 93 (6).
- 646 36. Raquin, V. et al. (2017) Individual co-variation between viral RNA load and gene
- 647 expression reveals novel host factors during early dengue virus infection of the
- Aedes aegypti midgut. PLoS Negl Trop Dis 11 (12), e0006152.
- 649 37. Zanini, F. et al. (2018) Virus-inclusive single-cell RNA sequencing reveals the
- molecular signature of progression to severe dengue. Proc Natl Acad Sci U S A 115(52), E12363-E12369.
- 652 38. Razzell, W. et al. (2013) Calcium flashes orchestrate the wound inflammatory
- response through DUOX activation and hydrogen peroxide release. Curr Biol 23 (5),424-9.
- 39. Cristinelli, S. and Ciuffi, A. (2018) The use of single-cell RNA-Seq to understand
   virus-host interactions. Current Opinion in Virology 29, 39-50.
- 40. Lequime, S. et al. (2016) Genetic Drift, Purifying Selection and Vector Genotype
- 658 Shape Dengue Virus Intra-host Genetic Diversity in Mosquitoes. PLoS Genet 12 (6), 659 e1006111.
- 41. Grubaugh, N.D. et al. (2016) Genetic Drift during Systemic Arbovirus Infection of
- 661 Mosquito Vectors Leads to Decreased Relative Fitness during Host Switching. Cell 662 Host Microbe 19 (4), 481-92.
- 42. Reid, A.J. et al. (2018) Single-cell RNA-seq reveals hidden transcriptionalvariation in malaria parasites. Elife 7.
- 43. Howick, V.M. et al. (2019) The Malaria Cell Atlas: Single parasite transcriptomes
- across the complete <em>Plasmodium</em> life cycle. Science 365 (6455),
- 667 eaaw2619.
- 44. Attar, M. et al. (2018) A practical solution for preserving single cells for RNA
  sequencing. Scientific Reports 8 (1), 2151.
- 45. Alles, J. et al. (2017) Cell fixation and preservation for droplet-based single-cell
  transcriptomics. BMC Biol 15 (1), 44.
- 46. van den Brink, S.C. et al. (2017) Single-cell sequencing reveals dissociation-
- 673 induced gene expression in tissue subpopulations. Nat Methods 14, 935.
- 47. Zhai, Z. et al. (2017) A genetic framework controlling the differentiation of
- 675 intestinal stem cells during regeneration in Drosophila. PLoS Genet 13 (6),676 e1006854.
- 48. Davie, K. et al. (2018) A Single-Cell Transcriptome Atlas of the Aging Drosophila
  Brain. Cell 174 (4), 982-998 e20.
- 679 49. Croset, V. et al. (2018) Cellular diversity in the Drosophila midbrain revealed by 680 single-cell transcriptomics. Elife 7.
- 681 50. Herculano-Houzel, S. (2009) The human brain in numbers: a linearly scaled-up 682 primate brain. Front Hum Neurosci 3, 31.
- <sup>683</sup> 51. Ruckert, C. and Ebel, G.D. (2018) How Do Virus-Mosquito Interactions Lead to
  <sup>684</sup> Viral Emergence? Trends Parasitol 34 (4), 310-321.
- 52. Burgess, D.J. (2019) Spatial transcriptomics coming of age. Nature Reviews
- 686 Genetics 20 (6), 317-317.
- 53. Ståhl, P.L. et al. (2016) Visualization and analysis of gene expression in tissue
- 688 sections by spatial transcriptomics. Science 353 (6294), 78-82.
- 689 54. Campbell, I.M. et al. (2015) Somatic mosaicism: implications for disease and
- transmission genetics. Trends Genet 31 (7), 382-92.

- 55. Petit, M.J. and Shah, P.S. (2018) Mapping Arbovirus-Vector Interactions Using
  Systems Biology Techniques. Front Cell Infect Microbiol 8, 440.
- 693 56. Karaiskos, N. et al. (2017) The Drosophila embryo at single-cell transcriptome 694 resolution. Science 358 (6360), 194-199.
- 695 57. Li, H. et al. (2017) Classifying Drosophila Olfactory Projection Neuron Subtypes 696 by Single-Cell RNA Sequencing, Cell 171 (5), 1206-1220 e22.
- 697 58. Ariss, M.M. et al. (2018) Single cell RNA-sequencing identifies a metabolic 698 aspect of apoptosis in Rbf mutant. Nat Commun 9 (1), 5024.
- 59. Zakovic, S. and Levashina, E.A. (2018) Insects Go on a STING Operation to Tackle Intracellular Invaders. Immunity 49 (2), 195-197.
- 60. Zakovic, S. and Levashina, E.A. (2017) NF-kappaB-Like Signaling Pathway REL2
- in Immune Defenses of the Malaria Vector Anopheles gambiae. Front Cell InfectMicrobiol 7, 258.
- 61. Palmer, W.J. and Jiggins, F.M. (2015) Comparative Genomics Reveals the
- 705 Origins and Diversity of Arthropod Immune Systems. Mol Biol Evol 32 (8), 2111-29.
- 62. Rosa, R.D. et al. (2016) Exploring the immune signalling pathway-related genes
- 707 of the cattle tick Rhipicephalus microplus: From molecular characterization to
- transcriptional profile upon microbial challenge. Dev Comp Immunol 59, 1-14.
- 63. Severo, M.S. et al. (2013) The E3 ubiquitin ligase XIAP restricts Anaplasma
- phagocytophilum colonization of Ixodes scapularis ticks. J Infect Dis 208 (11), 1830-40.
- 712 64. Oliva Chavez, A.S. et al. (2017) Tick Humoral Responses: Marching to the Beat
- 713 of a Different Drummer. Front Microbiol 8, 223.
- 714 65. Benoit, J.B. et al. (2017) Symbiont-induced odorant binding proteins mediate
- 715 insect host hematopoiesis. Elife 6.
- 716

# Figure 1

