



**HAL**  
open science

# Regulation of immune signalling in the malaria mosquito vector, *Anopheles*: the secreted mosquito leucine-rich repeat protein APL1C is a pathogen binding factor essential for immunity to *Plasmodium* ookinetes and sporozoites

Natalia Marta Zmarlak

## ► To cite this version:

Natalia Marta Zmarlak. Regulation of immune signalling in the malaria mosquito vector, *Anopheles*: the secreted mosquito leucine-rich repeat protein APL1C is a pathogen binding factor essential for immunity to *Plasmodium* ookinetes and sporozoites. Molecular biology. Sorbonne Université, 2021. English. NNT : 2021SORUS053 . tel-03634084

**HAL Id: tel-03634084**

**<https://theses.hal.science/tel-03634084>**

Submitted on 7 Apr 2022

**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.

# Sorbonne Université

École doctorale *Complexité du Vivant*

*Laboratoire de Génétique et Génomique des Insectes Vecteurs, Institut Pasteur,  
Paris*

## **Regulation of immune signalling in the malaria mosquito vector, *Anopheles***

*The secreted mosquito leucine-rich repeat protein APLIC is a  
pathogen binding factor essential for immunity to Plasmodium  
ookinetes and sporozoites*

Par Natalia Marta Zmarlak

Thèse de doctorat de Science du Vivant

Dirigée par Christian Mitri et Kenneth Vernick

Présentée et soutenue publiquement le 28.06.2021

Devant un jury composé de :

**Higuet Dominique.** Professeur de la Sorbonne Université, Président.

**Christophides George.** Chargé de recherche au Royaume-Uni, Rapporteur.

**Marois Eric.** Chargé de recherche INSERM, Rapporteur.

**Smith Ryan.** Chargée de recherche aux États-Unis, Examineur.

**Mitri Christian.** Thèse Co-Encadrant.

**Vernick Kenneth.** Directeur d'unité. Directeur de thèse.

1

2  
3

*This work is dedicated for my Parents. Thank you for raising me in love, faith  
and hope.*

4  
5

*Pracę te dedykuje moim Rodzicom. Dziękuję Wam za wychowanie mnie w  
miłości, wierze i nadziei.*

6

7

# SUMMARY

8		
9		
10		
11	<b>SUMMARY .....</b>	<b>3</b>
12	<b>ACKNOWLEDGMENTS.....</b>	<b>4</b>
13	<b>TABLE OF CONTENTS.....</b>	<b>5</b>
14	<b>LIST OF FIGURES.....</b>	<b>7</b>
15	<b>LIST OF TABLES.....</b>	<b>9</b>
16	<b>LIST OF ANNEX CONTENT.....</b>	<b>10</b>
17	<b>ABBREVIATIONS.....</b>	<b>11</b>
18	<b>I. GENERAL INTRODUCTION.....</b>	<b>17</b>
19	<b>II. CURRENT KNOWLEDGE, GAPS AND OBJECTIVES.....</b>	<b>41</b>
20	<b>III. RESULTS .....</b>	<b>50</b>
21	<b>IV. DISCUSSION .....</b>	<b>88</b>
22	<b>V. CONCLUDING REMARKS AND FUTURE PERSPECTIVES.....</b>	<b>96</b>
23	<b>VI. MATERIALS AND METHODS.....</b>	<b>99</b>
24	<b>REFERENCES.....</b>	<b>112</b>
25	<b>ANNEX .....</b>	<b>123</b>
26	<b>ABSTRACT .....</b>	<b>140</b>
27		

## ACKNOWLEDGMENTS

28

## 29 **ACKNOWLEDGMENTS**

30

31 I would like to thank my supervisors. Ken, Christian thank you for the opportunity  
32 to do this thesis. Thank you for all the time you both always had for me, our discussions  
33 and chance to observe your work. Finally, thank you that I could raise under your  
34 supervision as a scientist but also as a person.

35 I would also like to thank all the colleagues of the Genetics and Genomics of  
36 Insect Vectors Unit (Kenneth Vernick laboratory) who also attributed to this work,  
37 especially Emma who helped me with functional genomics studies, Corinne who  
38 always provided me the mosquitoes, Inge who taught me cell culture, Adrien who  
39 supervised my RNA sequencing analysis and Manu for advices concerning data  
40 analysis. Besides work, I would like to thank all present and previous members of  
41 GGIV for the everyday atmosphere that cheered me up during better and worse days  
42 of my PhD.

43 My deep gratitude to the members of Malaria Infection and Immunity (Rogerio  
44 Amino laboratory), especially Eduardo Aliprandini and Pauline Formaglio for  
45 discussions and advices concerning my immunofluorescence analyses and  
46 microscopy. Jean-Michel, thank you for your patience while teaching me salivary  
47 glands dissection and infection analysis.

48 I am grateful to Audrey Salles and Julien Fernandes for providing me skills for  
49 Laser Scanning Microscopy and to Jean-Yves Tinevez who helped me with pictures  
50 analysis.

51 My sincere thanks to all the members of CEPIA (Patricia Baldacci platform), for  
52 providing me mosquitoes. Special thanks to deceased Marek Szatanik (responsible for  
53 CEPIA platform 2015-2019) under whom supervision I made my first internship in  
54 Institut Pasteur and who will always stay in my memory as dear friend.

55 Thanks to all the members of Labo Prepa who took care of my working  
56 environment.

57 Thanks to all my friends in Poland and France. I would like give special thanks  
58 to my best friend Ilona who was always available to give me a call as well as to visit  
59 me in France. Aurelia and Kuba, thank you for providing scientific advices but also  
60 friend shoulder.

61 Finally, I would like to thank my family, my father Janusz Zmarlak and mother  
62 Marta Zmarlak, my brothers Marcin Zmarlak, Szymon Maj and Bartłomiej Zmarlak and  
63 their families. Last but not the least I want to thank my fiancé Maxence Feher who is  
64 my biggest supporter. Thank you for the faith you have in me and care that you took  
65 of our everyday life so I could focus on this thesis.

## TABLE OF CONTENTS

66	<b>TABLE OF CONTENTS</b>	
67		
68	<b>SUMMARY</b> .....	<b>3</b>
69	<b>ACKNOWLEDGMENTS</b> .....	<b>4</b>
70	<b>TABLE OF CONTENTS</b> .....	<b>5</b>
71	<b>LIST OF FIGURES</b> .....	<b>7</b>
72	<b>LIST OF TABLES</b> .....	<b>9</b>
73	<b>LIST OF ANNEX CONTENT</b> .....	<b>10</b>
74	<b>ABBREVIATIONS</b> .....	<b>11</b>
75	<b>I. GENERAL INTRODUCTION</b> .....	<b>17</b>
76	<b>1.1 Malaria – overview</b> .....	<b>18</b>
77	1.1.1 Female <i>Anopheles</i> mosquitoes – vector of malaria .....	18
78	1.1.2 History of malaria, mosquito-vector discovery and vector control .....	18
79	1.1.3 Why is Africa the most malaria endemic continent – what makes a competent mosquito vector? .	21
80	1.1.4 The complex lifecycle of <i>Plasmodium</i> parasite in vertebrate and mosquito.....	22
81	1.1.5 Two bottlenecks of <i>Plasmodium</i> development in mosquito vector .....	26
82	<b>1.2 Microbial sensing by the immune system</b> .....	<b>27</b>
83	1.2.1 Basis of microbial immune sensing and evolution.....	27
84	1.2.2 LRR domain proteins in immunity .....	29
85	1.2.2.1 LRR proteins in the innate immune system .....	30
86	1.2.2.1.1 <i>Drosophila</i> Toll receptors .....	30
87	1.2.2.1.2 Toll-like receptors (TLR) .....	31
88	1.2.2.1.3 Nucleotide-binding leucine-rich repeat receptors.....	33
89	1.2.2.1.4 Plant LRRs.....	34
90	1.2.2.2 LRR-coding receptors in adaptive-like immunity of jawless vertebrates.....	38
91	<b>II. CURRENT KNOWLEDGE, GAPS AND OBJECTIVES</b> .....	<b>41</b>
92	<b>2.1 Natural resistance to <i>Plasmodium</i> highlighted a family of LRR genes termed APL1 (<i>Anopheles</i></b>	
93	<b><i>Plasmodium</i>-responsive leucine-rich repeat)</b> .....	<b>42</b>
94	<b>2.2 APL1 functional activity and the mosquito complement system</b> .....	<b>46</b>
95	<b>2.3 Scope of the thesis</b> .....	<b>48</b>
96	<b>III. RESULTS</b> .....	<b>50</b>
97	<b>3.1 APL1C protein is a pathogen binding factor for the midgut ookinete stage of <i>Plasmodium</i>..</b>	<b>51</b>
98	3.1.1 Blood feeding induces an extracellular layer of APL1C surrounding the midgut.....	51
99	3.1.2 APL1C presence on the basal side of the midgut epithelium is microbiome independent.....	55
100	3.1.3 APL1C binds to <i>Plasmodium</i> ookinetes exiting the basal side of the midgut epithelium .....	56
101	3.1.4 LRR proteins bind autonomously to the ookinete surface (coll.: C. Lavazec, Institut Cochin, Paris) .	59
102	3.1.5 Phagocytic hemocytes are required for wildtype APL1C levels in the hemolymph.....	63
103	3.1.6 Nitration pathway activity is required for full APL1C abundance in hemolymph.....	64
104	<b>3.2 APL1C protein is a pathogen binding factor for the hemocoel sporozoite stage of <i>Plasmodium</i></b>	
105	.....	<b>67</b>
106	3.2.1 Sporozoites affect APL1C protein abundance .....	67
107	3.2.2 APL1C activity in the hemocoel limits sporozoite invasion of salivary glands.....	69
108	3.2.3 APL1C binds to free <i>Plasmodium</i> sporozoites in the mosquito hemolymph.....	71

## TABLE OF CONTENTS

109	3.2.4 APL1C protective activity against sporozoites does not require the complement proteins TEP1 or	
110	TEP3.....	72
111	<b>3.3 Implication of the APL1 LRR family in immune signal transduction.....</b>	<b>74</b>
112	3.3.1 APL1A and APL1C gene silencing in <i>A. coluzzii</i> .....	75
113	3.3.2 Kinetics of APL1A and APL1C silencing.....	77
114	3.3.3 RNAseq and bioinformatic analysis .....	81
115	3.3.4 qPCR validation of the RNAseq gene candidates.....	84
116	3.3.5 APL1C controls expression of immune-like genes in the mosquito .....	86
117	<b>IV. DISCUSSION .....</b>	<b>88</b>
118	4.1 Dual functionality for APL1C as guard and PRR-like factor .....	89
119	4.2 The HdMvs, a source of secreted APL1C, mediated by the nitration pathway?.....	91
120	4.3 APL1C presence in hemolymph at the basal side of the midgut epithelium is not PAMP	
121	dependent.....	91
122	4.4 The sophisticated sporogony control by APL1C.....	93
123	4.5 Whole transcriptome analysis reveals differential signalling function of APL1 genes .....	94
124	<b>V. CONCLUDING REMARKS AND FUTURE PERSPECTIVES.....</b>	<b>96</b>
125	<b>VI. MATERIALS AND METHODS.....</b>	<b>99</b>
126	<b>REFERENCES.....</b>	<b>112</b>
127	<b>ANNEX .....</b>	<b>123</b>
128	<b>ABSTRACT .....</b>	<b>140</b>
129		
130		
131		
132		

## LIST OF FIGURES

133	<b>LIST OF FIGURES</b>	
134		
135	<b>Figure 1. <i>Plasmodium</i> parasite asexual development in a human host.....</b>	<b>24</b>
136	<b>Figure 2. <i>Plasmodium</i> development in female <i>Anopheles</i>.....</b>	<b>25</b>
137	<b>Figure 3. LRR-containing immune receptors, modes of pathogen perception</b>	
138	<b>and activation in plants. ....</b>	<b>36</b>
139	<b>Figure 4. Schematic structure of a VLR protein and VLR gene assembly.....</b>	<b>39</b>
140	<b>Figure 5. Classification of mosquito LRIM proteins.....</b>	<b>45</b>
141	<b>Figure 6. Complement pathway activation and outcome in vertebrates.....</b>	<b>47</b>
142	<b>Figure 7. Blood feeding induces the presence of APL1C protein on the basal</b>	
143	<b>side of the mosquito midgut. ....</b>	<b>53</b>
144	<b>Figure 8. Anti-APL1C pAb labelling is specific to the APL1C protein.....</b>	<b>54</b>
145	<b>Figure 9. The mosquito microbiome does not influence APL1C protein</b>	
146	<b>abundance on the midgut 24 h after NBM.....</b>	<b>56</b>
147	<b>Figure 10. APL1C binds to ookinetes in <i>A. coluzzii</i> midguts. ....</b>	<b>57</b>
148	<b>Figure 11. APL1C binds to the ookinetes located on the basal side of the</b>	
149	<b>midgut. ....</b>	<b>59</b>
150	<b>Figure 12. APL1C and LRIM1 both bind to <i>P. berghei</i>.....</b>	<b>61</b>
151	<b>Figure 13. APL1C and LRIM1 both bind to <i>P. berghei</i> ookinetes independently</b>	
152	<b>of TEP1 presence. ....</b>	<b>62</b>
153	<b>Figure 14. Full APL1C gene expression and protein abundance in the</b>	
154	<b>hemolymph requires phagocytic hemocytes. ....</b>	<b>64</b>
155	<b>Figure 15. Nitration pathway activity mediates APL1C protein presence in</b>	
156	<b>mosquito hemolymph.....</b>	<b>66</b>
157	<b>Figure 16. APL1C protein level is increased by blood feeding and sporozoite</b>	
158	<b>release in <i>A. coluzzii</i>. ....</b>	<b>68</b>
159	<b>Figure 17. APL1C activity limits sporozoite invasion of salivary glands. ....</b>	<b>70</b>
160	<b>Figure 18. APL1C protein binds to free sporozoites in mosquito hemolymph.</b>	
161	<b>.....</b>	<b>72</b>
162	<b>Figure 19. APL1C protective activity against sporozoites does not require the</b>	
163	<b>complement proteins TEP1 or TEP3.....</b>	<b>73</b>
164	<b>Figure 20. Overview of the RNAseq experiment.....</b>	<b>75</b>
165	<b>Figure 21. 3' APL1A and APL1C fragments identified by 3' RACE overlap with</b>	
166	<b>dsAPL1A and dsAPL1C fragments.....</b>	<b>77</b>
167	<b>Figure 22. The potential time effect after APL1 depletion on the number of the</b>	
168	<b>affected genes. ....</b>	<b>78</b>
169	<b>Figure 23. APL1 gene silencing efficiency after dsRNA treatment.....</b>	<b>79</b>
170	<b>Figure 24. The kinetics of APL1C protein levels after dsAPL1C.....</b>	<b>80</b>



## LIST OF FIGURES

171	<b>Figure 25. APL1 genes silencing controls prior to the RNAseq analysis. ....</b>	<b>81</b>
172	<b>Figure 26. Volcano plots visualize the differential expression analysis of</b>	
173	<b>RNAseq taken in dsAPL1A vs dsGFP (A) and dsAPL1C vs dsGFP conditions</b>	
174	<b>(B).....</b>	<b>82</b>
175	<b>Figure 27. MA-plots visualizing the log ratio of differential expression as a</b>	
176	<b>function of the mean intensity taken in dsAPL1A vs dsGFP (A) and dsAPL1C vs</b>	
177	<b>dsGFP conditions (B).....</b>	<b>83</b>
178	<b>Figure 28. qPCR gene expression analysis validates RNAseq APL1C gene</b>	
179	<b>targets. ....</b>	<b>85</b>
180	<b>Figure 29. Proposed model of APL1C function in anti-Plasmodium response.</b>	<b>98</b>
181		

182 **LIST OF TABLES**

183

184 **Table 1. Role of human TLRs in pathogen recognition..** .....32

185 **Table 2. Number of up-, down- and total number of differentially expressed**  
186 **genes in each compared condition.** .....84

187 **Table 3. List of the immune-like genes, which expression is downregulated**  
188 **upon APL1C depletion.**.....86

189 **Table 4. List of the immune-like genes, which expression is upregulated upon**  
190 **APL1C depletion.** .....87

191

192 **LIST OF ANNEX CONTENT**

193

194 **Figure S1. Workflow of IFA experiment..... 123**

195 **Figure S2. Mosquito hemolymph-injected dextran diffuses on mosquito**  
 196 **midguts in a way resembling APL1C protein deposition. .... 124**

197 **Figure S3. Tissue permeabilization enables antibody access to all the parasites**  
 198 **present in mosquito midguts..... 124**

199 **Figure S4. V5-tagged protein constructs are secreted from hemocyte-like cells..**  
 200 **..... 124**

201 **Figure S5. Workflow of salivary glands infection prevalence and intensity**  
 202 **analysis after dsRNA treatment..... 125**

203 **Figure S6. RNAseq bioinformatic analysis pipeline that identified differentially**  
 204 **expressed genes in APL1A and APL1C depleted mosquitoes. .... 126**

205 **Figure S7. Total sequence number reported in the RNAseq analysis for all the**  
 206 **studied samples. The results show that all samples displayed a close total**  
 207 **number of reads (30-40 M reads). .... 127**

208

209

210 **Table S1. Statistical analysis of APL1C signal on mosquito midguts..... 128**

211 **Table S2. Statistical analysis of P. berghei salivary glands infection prevalence**  
 212 **following gene silencing..... 129**

213 **Table S3. APL1C controls sporozoite intensity in mosquito hemolymph..... 130**

214 **Table S4. Full list of genes, which expression was downregulated upon APL1C**  
 215 **depletion in RNAseq experiment. .... 133**

216 **Table S5. Full list of genes, which expression is upregulated upon APL1C**  
 217 **depletion in RNAseq experiment. .... 138**

218 **Table S6. Primers sequences used for the plasmids construction. .... 138**

219 **Table S7. List of the primers used for dsRNA synthesis, qPCR and RACE..... 139**

220

221 **ABBREVIATIONS**

222

**A**

*A. gambiae* s.l. – *Anopheles gambiae* sensu lato

AMP – antimicrobial peptides

AP-1 – activator protein 1

APL1 – *Anopheles Plasmodium*-responsive leucine-rich repeat 1

ATB – antibiotic

ATP – adenosine triphosphate

Avr – avirulence

**B**

BCR – B-cell receptor

bp – base pair

Bti – *Bacillus thuringiensis* var. *israelensis*

**C**

C3 – complement component 3

CC – coiled coil

cDNA – complementary DNA

CLD – clodronate

CP – connecting peptide

CRISPRs – clustered regulatory interspaced palindromic repeats

**D**

DAMP – damage associated molecular pattern

DAPI – 4',6-diamidino-2-phenylindole

DC – dendritic cell

## ABBREVIATIONS

DNA – deoxyribonucleic acid

ds – double stranded

DTT – dichlorodiphenyltrichloroethane

### **E**

ETI – effector triggered immunity

### **F**

FC – fold change

### **G**

GAPDH – glyceraldehyde 3-phosphate dehydrogenase

GFP – green fluorescent protein

GMEC – Global Malaria Eradication Campaign

GNBP – Gram-negative binding-protein

### **H**

HdMv – hemocyte derived microvesicles

HELI – Health and Environment Linkages Initiative

HPX2 – heme peroxidase 2

HR – hypersensitive response

HSP – heat shock protein

### **I**

IBM – *Plasmodium* infected bloodmeal

IFA – immunofluorescence assay

Ig – immunoglobulin

Imd – immunodeficiency

IP – Interpro

## ABBREVIATIONS

IRF3 – interferon regulatory factor 3

IRS – indoor residual spraying

ITN – insecticide-treated nets

IVM – integrated vector management

I $\kappa$ B – inhibitor of nuclear factor kappa B

### **J**

JAK – Janus kinase

### **L**

LAM – lipoarabinomannan

LP – liposomes

LPS – lipopolysaccharide

LRIM – leucine-rich repeat immune

LRR – leucine-rich repeat

LRRCT – leucine-rich repeat C-terminus

LRRNT – leucine-rich repeat N-terminus

LRRV – leucine-rich repeat variable

LSM – laser scanning microscope

LTA – lipoteichoic acid

### **M**

mAb – monoclonal antibody

MAC – membrane attack complex

MAPK – mitogen activated protein kinase

MBL – mannose binding lectin

MHC – major histocompatibility complex

MyD88 – myeloid differentiation primary-response protein 88

## ABBREVIATIONS

### **N**

n.s. – not significant

NA – not applicable

NB – nucleotide binding

NBM – non-infected bloodmeal

NBS – nucleotide binding site

NFkB – nuclear factor kappa-B

NLR – nucleotide-binding leucine-rich repeat receptor

Nod – nucleotide-binding oligomerization domain-like

NOX5 – NADPH oxidase 5

### **P**

pAb – polyclonal antibody

PAMP – Pathogen Associated Molecular Pattern

PBS – phosphate-buffered saline

PCR – polymerase chain reaction

PGN – peptidoglycan

PGRP – peptidoglycan recognition protein

PRI – *Plasmodium* resistant island

PRR – Pathogen Recognition Receptor

qPCR – quantitative PCR

### **R**

RACE – rapid amplification of cDNA ends

RBC – red blood cell

rDNA – ribosomal DNA

RFP – red fluorescent protein

RLK – receptor-like kinase

## ABBREVIATIONS

RLP – receptor-like protein

RNA – ribonucleic acid

RNAi – RNA interference

ROS – reactive oxygen species

### **S**

s.l. – sensu lato

SEM – standard error of mean

SPE – Spätzle-processing enzyme

spp. – species pluralis

Spz – Spatzle

STAT – signal transducer and activator of transcription

sTLR – secreted Toll-like receptor

SUG – sugar meal

### **T**

TCR – T-cell receptor

TEP1 – thioester-containing protein 1

TIR – Toll/interleukine-1 receptor

TLR – toll-like receptor

TM – transmembrane

UP – UniPro

### **V**

VB – VectorBase

VC – vector competence

VLR – variable lymphocyte receptor

### **W**



## ABBREVIATIONS

WHO – World Health Organization

223

## GENERAL INTRODUCTION

224

225

226

227

228

229

230

231

232

233

234

## **I. GENERAL INTRODUCTION**

235

236 **1.1 Malaria – overview**

237 **1.1.1 Female *Anopheles* mosquitoes – vector of malaria**

238

239 Morbidity and mortality caused by mosquito-borne diseases makes a mosquito the  
240 most dangerous animal for a human being<sup>1</sup>. The hematophagous female mosquito  
241 acquires essential nutrients for egg development by taking blood from vertebrate  
242 animals including or humans. Through the bloodmeal, mosquitoes can ingest and  
243 eventually transmit pathogenic microbes with injected saliva to the vertebrate host  
244 during the next bloodmeal. Many mosquito species are vectors responsible for the  
245 transmission of multiple pathogens, including protozoans, viruses, bacteria and  
246 nematodes, causing diseases such as malaria, yellow fever, dengue, filariasis or  
247 encephalitis. This blood dependency has been used by the pathogens to co-evolve  
248 with their respective mosquito hosts to assure their transmission. Mosquito vectors  
249 thus bridge the physical barriers that would be impossible for the pathogen to cross  
250 through direct contact between humans or animals<sup>2</sup>.

251 Mosquitoes of the genus *Anopheles* are responsible for the transmission of the  
252 protozoan parasite *Plasmodium*, the causative agent of malaria, the world's deadliest  
253 parasitic disease. In 2019, there were 229 million diagnosed cases of malaria  
254 worldwide, with 409 000 human deaths reported. Approximately 94% of global malaria  
255 cases occur in sub-Saharan Africa. Children under 5 years old are the most vulnerable  
256 group to malaria and constitute two-thirds of all malaria deaths worldwide in 2019<sup>3</sup>.

257 **1.1.2 History of malaria, mosquito-vector discovery and vector**  
258 **control**

259

## GENERAL INTRODUCTION

260 Malaria is one of the oldest human diseases. The first recorded mention that  
261 corresponded to malaria symptoms appeared in notes from ~2700 BC. The word  
262 “malaria” comes from an Italian appellation of the disease during Middle Age, mal’aria,  
263 meaning “spoiled air”, reflecting the doctrine that lasted for over 2500 years, which  
264 considered malaria as an aerial infection arising from the marshes. Our understanding  
265 of this disease started at the end of the XIX century when a French physician, Alphonse  
266 Laveran, discovered *Plasmodium* parasites in a malaria infected patient<sup>4</sup>. For this  
267 finding, Laveran won the Nobel Prize in Physiology or Medicine in 1907.

268 In the 1890s, Patrick Manson proved an association between insects and pathogens  
269 causing human and animal diseases. His discovery of mosquito involvement in  
270 transmission of filarial worms responsible for lymphatic filariasis suggested that this  
271 insect could also be responsible for malaria transmission<sup>5</sup>. Scientists postulated the  
272 possibilities of human infection by drinking water contaminated by infected mosquitoes  
273 or inhaling dust from dried-up ponds in which infected mosquitoes died. Others  
274 suspected that the parasites could be transmitted mechanically from host to host on  
275 the proboscis of a mosquito<sup>4,5</sup>. But finally, in 1897 Ronald Ross, after many attempts,  
276 successfully infected female mosquitoes with *Plasmodium* from the blood of infected  
277 patients, thus resolving all previous speculations. His observations of circular parasites  
278 in dissected mosquito midguts (now called oocysts) were followed by research on  
279 avian malaria, where he demonstrated the rupture of oocysts leading to sporozoite  
280 release in the circulatory system of the mosquitoes (hemocoel) and their migration to  
281 the salivary glands, thus completing the parasite life cycle<sup>4,6</sup>.

282 The ground-breaking discovery of R. Ross opened a new avenue in fighting malaria  
283 disease, through controlling the essential part of *Plasmodium* parasite transmission by  
284 targeting the *Anopheles* mosquito vector. The historic success of malaria eradication

## GENERAL INTRODUCTION

285 in various parts of the world was mainly achieved by vector control programs. The early  
286 malaria vector control programs attempted to eliminate mosquito through  
287 environmental management based on knowledge of insect development. The  
288 mosquito lifecycle includes four stages: egg, larva, pupae and adult, of which the first  
289 three are aquatic. Therefore, early strategies concentrated on limiting the aqueous  
290 environment, such as drainage of habitats, oil application to open water bodies, grid-  
291 ditching of marshes, modification of river boundaries and other engineering  
292 approaches, which decreased mosquito populations by denial of larval habitat<sup>7,8,9</sup>.  
293 However, the most efficient tool appeared with the use of insecticides. Since the first  
294 half of the 20th century, different classes of insecticides have been successively used:  
295 organochlorines, organophosphates, carbamates and synthetic pyrethroids. Use of  
296 insecticides greatly limited mosquito-borne diseases and in some areas even  
297 eradicated malaria. Indoor residual spraying (IRS) using  
298 dichlorodiphenyltrichloroethane (DDT) implemented in the Global Malaria Eradication  
299 Campaign (GMEC) in 1955, showed spectacular success in decimating disease  
300 vectors. However, despite the efficiency of DDT in reducing mosquito vector  
301 populations, safety concerns for humans and their environments led to its ban in many  
302 countries in the 1970s. Moreover, at least during the last 20 years, the widespread  
303 development of mosquito resistance to the most commonly used insecticides caused  
304 serious problems in many areas resulting in recurrent outbreaks of mosquito-borne  
305 diseases<sup>10,11,12,13,14,15,16</sup>. Alternative approaches that have already been implemented  
306 use natural insecticidal materials obtained from animals, plants, bacteria or minerals  
307 instead of chemical insecticides. Classified as biopesticides, entomopathogenic  
308 bacteria and fungi, such as *Bacillus thuringiensis* var. *israelensis* (Bti) or *Metarhizium*  
309 spp. displayed promising results in mosquito elimination. Nevertheless, their efficacy

## GENERAL INTRODUCTION

310 varies in different environments, affected by factors such as water pollution or  
311 instability to UV radiation<sup>17</sup>. Therefore, according to World Health Organization (WHO)  
312 recommendations, insecticide-based strategies, including insecticide-treated mosquito  
313 nets (ITNs) and IRS using pyrethroids, remain the mainstay interventions for mosquito  
314 vector control. At the same time, the recent WHO Health and Environment Linkages  
315 Initiative (HELI) stressed the importance of basic knowledge on mosquito vector  
316 behaviour and ecology, in order to develop tailor-made vector management actions in  
317 malaria endemic sites. This strategy enables implementation of integrated vector  
318 management (IVM), which besides conventional methods in fighting mosquito vector  
319 (ITNs and IRS) implements alternative strategies, such as environmental manipulation.  
320 IVM is a dynamic and still-evolving field necessary for sustainable and ecologically  
321 sound mosquito control strategies<sup>18</sup>.

322

### 323 **1.1.3 Why is Africa the most malaria endemic continent – what** 324 **makes a competent mosquito vector?**

325

326 *Plasmodium* parasite transmission between humans requires a competent mosquito  
327 vector. Vector competence defines the intrinsic ability of a vector to carry, tolerate the  
328 development and transmit a microbe and it reflects the susceptibility of a mosquito  
329 vector to such microbe<sup>19</sup>. Vector competence is usually measured experimentally as  
330 the prevalence of infected and infectious vectors at given times post-microbial  
331 exposure<sup>20</sup>. The reasons for high malaria endemicity cases in Africa is a combination  
332 of physiologically permissive vector species mainly belonging to the *Anopheles*  
333 *gambiae sensu lato (s.l.)* complex and, most importantly, their highly anthropophilic  
334 feeding behaviour. The anthropophily of *A. gambiae s.l.* probably originated from the  
335 integration of these mosquitoes into human ecosystems about 8000 years ago when

## GENERAL INTRODUCTION

336 the advent of agriculture led to the new ecological niche of fixed human settlements<sup>19</sup>.  
337 <sup>21,22</sup>.

338 Among approximately 100 *Plasmodium* species that are infective for different  
339 vertebrates, five of them cause malaria in humans: *Plasmodium falciparum*,  
340 *Plasmodium vivax*, *Plasmodium malariae*, *Plasmodium ovale* and *Plasmodium*  
341 *knowlesi*<sup>23,24</sup>. Today the majority of malaria cases in Africa are caused by *P.*  
342 *falciparum*<sup>23</sup>.

343 *A. gambiae* was initially considered as a single malaria vector species across the  
344 African continent. However, beginning in the 1960s hybridization experiments  
345 combined with genetic studies showed that *A. gambiae s.l.* is a complex of at least  
346 seven morphologically indistinguishable sibling species, and more taxa continue to be  
347 discovered<sup>25,26,27</sup>. Different species display diverse biting behaviour and ecological  
348 preferences, including indoor (endophagous) or outdoor (exophagous) biting  
349 behaviour. Therefore, indoor targeted vector control strategies are not completely  
350 efficient against outdoor-biting taxa or genotypes. This behavioural diversity of *A.*  
351 *gambiae s.l.* increases the challenge in vector control strategies. Mosquito genotyping  
352 combined with behavioural and infectious studies are helping to better understand the  
353 complexity of the efficient transmission system in order to develop more efficient  
354 control strategies.

355

### 356 **1.1.4 The complex lifecycle of *Plasmodium* parasite in vertebrate** 357 **and mosquito**

358

359 As pointed out, the success of parasite transmission to humans depends on the vector  
360 competence of the mosquito. *Plasmodium* has a complex lifecycle that requires two

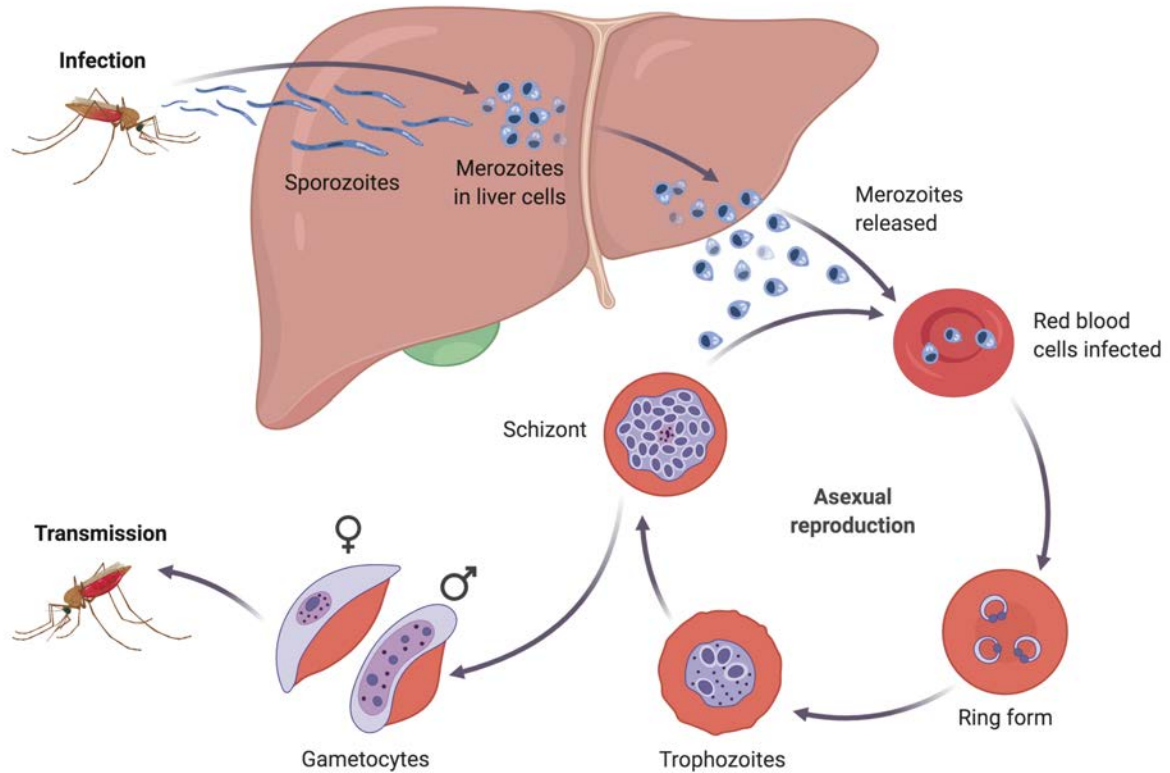
## GENERAL INTRODUCTION

361 hosts: the intermediate vertebrate host, and the definitive mosquito host. *Plasmodium*  
362 sporozoites are transmitted to the vertebrate host through the saliva of an infected  
363 mosquito (**Figure 1**). Sporozoites are carried by blood flow to the liver where they infect  
364 hepatocytes. There, the parasites grow and multiply to be released into the blood  
365 stream as merozoites, which infect red blood cells (RBCs). Repetitive intraerythrocytic  
366 cycles occur to form different stages: rings, trophozoites and schizonts, which release  
367 new merozoites that invade non-infected RBCs. The contents of the infected RBCs  
368 released upon their lysis stimulate the production of tumour necrosis factor and other  
369 cytokines, which are responsible for clinical manifestations of malaria disease. People  
370 infected with *Plasmodium* typically experience fever, shivering, cough, respiratory  
371 distress, joint pain, headache, diarrhoea, vomiting and convulsions. *P. falciparum* is  
372 the most dangerous human malaria species and can be sequestered in  
373 microvasculature of different organs, including the brain, which can result in cerebral  
374 malaria and potentially death<sup>23,28,29</sup>.

375



## GENERAL INTRODUCTION



376

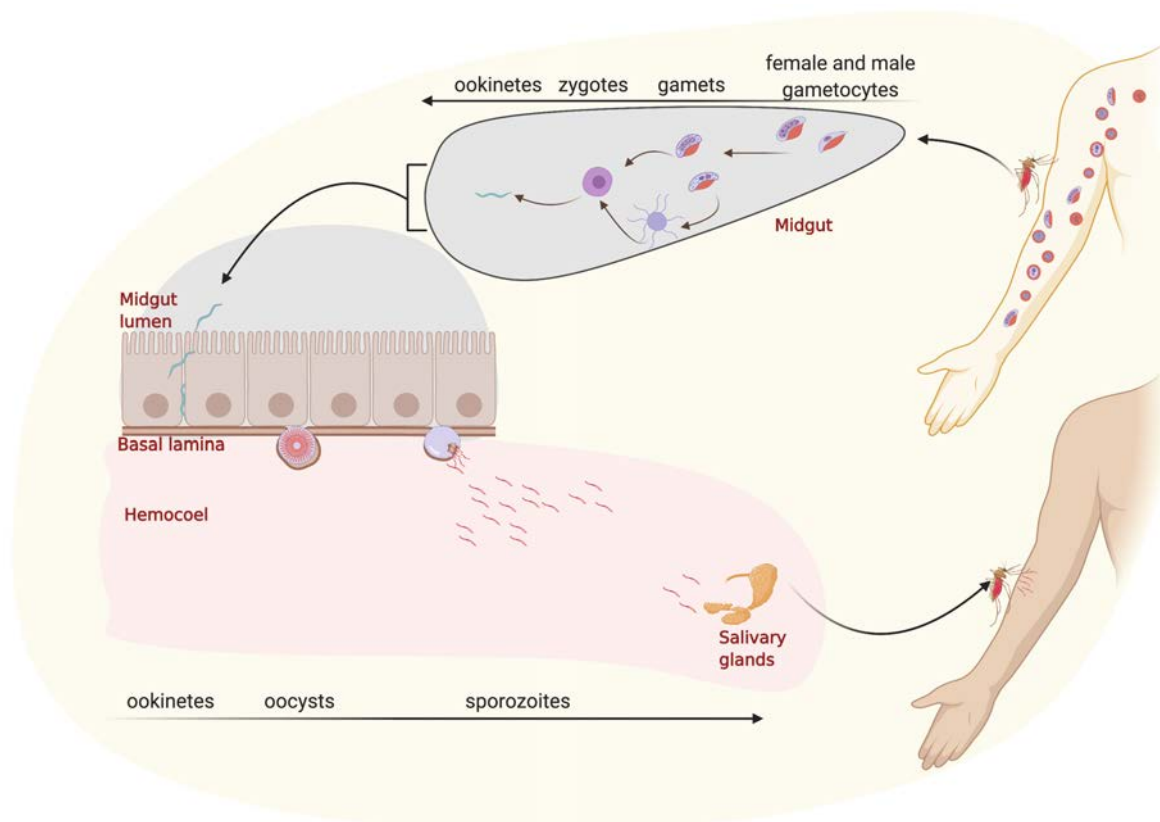
377 **Figure 1. Plasmodium parasite asexual development in a human host.** Detailed  
378 description in the text.

379

380 None of the parasite asexual blood stages are infectious to the mosquito. After several  
381 erythrocytic cycles, stress and other signals leads to parasite differentiation into female  
382 and male gametocytes, maturing through five different stages (I-V). Nevertheless,  
383 the mechanism of the parasite commitment to start sexual development still remains  
384 ambiguous, as approximately 10% of the merozoites undergo sexual development<sup>30</sup>.  
385 Only stage V mature female and male gametocytes are transmissible stages from  
386 human to mosquito. Upon a bloodmeal on an infected host, the female mosquito  
387 ingests both female and male gametocytes. In the midgut lumen, each male  
388 gametocyte produces eight flagellated motile microgametes in a process called  
389 exflagellation, induced by environmental and vector factors (**Figure 2**). Released  
390 motile male gametes attach and fertilize female gametes within ~30 min after ingestion,

## GENERAL INTRODUCTION

391 forming zygotes. The zygotes transform into motile ookinetes that traverse mosquito  
392 midgut epithelial cells, from the apical to the basal side where they round up to become  
393 oocysts. These first steps of the parasite development in the mosquito, from  
394 gametocyte fertilization to ookinete traversal, are relatively rapid, taking approximately  
395 24 hours. The subsequent steps are longer in time. Within the sessile oocysts, multiple  
396 rounds of mitosis result in thousands of sporozoites. Between 10-12 days after oocyst  
397 establishment, mature oocysts undergo membrane rupture, leading to the release of  
398 sporozoites in the mosquito hemocoel, which is an open circulatory system, from which  
399 sporozoites invade the salivary glands. There, the sporozoites remain metabolically  
400 quiescent for the life of the vector, and are transmitted in the saliva to a new host during  
401 later bloodmeals<sup>31,32,33,34,35,36</sup>.



402

403 **Figure 2. Plasmodium development in female Anopheles.** Detailed description in the text.

404

### 405 **1.1.5 Two bottlenecks of *Plasmodium* development in mosquito** 406 **vector**

407

408 Even though *Anopheles* female mosquitoes are essential for *Plasmodium*  
409 transmission, not all of them are competent vectors for the parasites. In other words,  
410 the fact that a mosquito takes a bloodmeal from *Plasmodium*-infected host does not  
411 mean that the parasites will successfully develop within the insect vector. Indeed,  
412 during the parasite lifecycle, the largest losses in the number of parasites occur inside  
413 the mosquito host<sup>37</sup>. This indicates that the mosquito mounts defence mechanisms  
414 towards *Plasmodium* that can lead to complete elimination of the parasites from the  
415 mosquito host. During the early stages of *Plasmodium* development in the mosquito,  
416 from gamete fertilization to ookinete formation, the parasites are still exposed to the  
417 vertebrate host immune factors present in the ingested blood sera, which can impede  
418 parasite development in the mosquito<sup>38</sup>. The mosquito microbiome, which expands in  
419 abundance after the bloodmeal, could have deleterious, neutral, or beneficial effects  
420 on early parasite development, and is not yet well understood<sup>39-42</sup>. Finally, digestive  
421 enzymes in the mosquito midgut as well as the development of the chitin-based  
422 peritrophic matrix around the ingested blood bolus could serve as a mechanical barrier  
423 that some ookinetes fail to pass<sup>38</sup>. However, the most dramatic bottlenecks in the  
424 *Plasmodium* lifecycle appear during ookinete traversal through midgut epithelium, as  
425 out of ~1000 ookinetes invading the midgut epithelium, ~5 may succeed in forming  
426 oocysts<sup>43,44</sup>.

427 The next bottleneck is when sporozoites are released from mature oocysts in the  
428 hemolymph to invade the salivary glands. It is estimated that approximately 80% of  
429 released sporozoites will fail to infect the salivary glands<sup>45,46,47</sup>. During these two main  
430 bottlenecks of parasite development in the mosquito, *Plasmodium* is in direct contact

## GENERAL INTRODUCTION

431 with the insect hemolymph, thus with soluble mosquito immune factors. Therefore, the  
432 mosquito hemolymph appears to be a battlefield in which *Plasmodium* parasites are  
433 attacked by mosquito defence factors. While the role of mosquito immune factors in  
434 the first bottleneck has been clearly demonstrated, this is not the case for the second  
435 one, concerning the large sporozoite clearance seen in the hemolymph, and this still  
436 needs to be clarified.

437

## 438 **1.2 Microbial sensing by the immune system**

### 439 **1.2.1 Basis of microbial immune sensing and evolution**

440

441 One of the challenges in immunology is to understand how the host organism detects  
442 the presence of infectious agents and mounts efficient killing mechanisms against  
443 these intruders, without destroying self-tissues. Sensing intruder as “non-self” or as  
444 lacking “self-markers” constitutes the starting point of the host response as a defence  
445 reaction in all living organisms. For microbial sensing, host-derived factors need to  
446 come in contact with the pathogen and generate the signals that will be transduced  
447 into a downstream immune response leading to the clearance of the microbes. Even  
448 such simple organisms as prokaryotes have defence mechanisms based on restriction  
449 enzymes and clustered regularly interspaced palindromic repeats (CRISPRs) that  
450 target invading pathogenic phages for degradation<sup>48</sup>. The foundational defence  
451 strategies in eukaryotic organisms are generally classified as the innate immune  
452 system, a host defence system in which plants and animals share similar inherited  
453 molecular modules including for pathogen sensing. The principle of innate immune  
454 recognition is based on the detection of constitutive and conserved products from  
455 structural microbial surface proteins or from their metabolism. Different pathogens

## GENERAL INTRODUCTION

456 share unique structures, for example lipopolysaccharide (LPS) or peptidoglycan, which  
457 are bacterial surface molecules, characterized as Pathogen Associated Molecular  
458 Patterns (PAMPs) and which are not synthesized by eukaryotic cells. The structure of  
459 these components is conserved, as they are often essential components for microbial  
460 survival and fitness. PAMPs are sensed by host germline-encoded proteins serving as  
461 pathogen recognition receptors (PRRs). Activation of PRRs through PAMP recognition  
462 transduces an immune signal inside the cells that leads to diverse killing mechanisms  
463 of the invaded pathogen (opsonization, activation of complement, phagocytosis,  
464 activation of pro-inflammatory signalling pathways, induction of apoptosis etc.).

465 In the first vertebrates, around 500 million years ago, a more sophisticated mechanism  
466 called adaptive immunity, was built upon the innate immune layer. One of the unique  
467 features of adaptive immunity is the rearrangement of immune receptor gene  
468 segments, which result in development of clonally diverse immune cells called  
469 lymphocytes, where each possesses a unique antigen recognition receptor. Somatic  
470 recombination leads to the development of a large repertoire of lymphocytes that can  
471 recognize distinct microbes and mount highly specific immune responses tailored to  
472 the threat. Pathogen recognition occurs through mature effector lymphocytes with  
473 cytotoxic and pro-inflammatory functions or plasma cells that secrete antibodies (the  
474 soluble forms of the antigen-specific receptors). Moreover, the clonal expansion and  
475 long life of primed cytotoxic lymphocytes and plasma cells provide a protective immune  
476 memory to prevent reinfection<sup>49,50,51</sup>.

477 The recognition of a microbial intruder leads to the activation of defence mechanisms  
478 in the host-organism, leading to the immune responses. The study of *Drosophila*  
479 *melanogaster* was highly relevant to mammalian biology, given the evolutionary  
480 conservation of many key signal transduction pathways. In *Drosophila*, the innate

## GENERAL INTRODUCTION

481 immune system recognizes infectious microbes<sup>52</sup>. *Drosophila* Toll and  
482 Immunodeficiency (Imd) immune pathways represent two distinct major arms of the  
483 immune response, which are activated by different microbes. Gram-positive bacteria  
484 and fungi activate the Toll signalling pathway, whereas Gram-negative bacteria  
485 activate the Imd pathway. Both pathways induce humoral and cellular changes in the  
486 insect host in order to eliminate or control the invasion of the intruder. Besides Toll and  
487 Imd immune-signalling pathway, the Janus kinase (JAK)-signal transducer and  
488 activator of transcription (STAT) and RNA interference (RNAi) pathways are immune  
489 arms of the antiviral response. Thus, invertebrate innate immunity is able to mount  
490 different defence response against distinct pathogens<sup>53</sup>.

491

### 492 **1.2.2 LRR domain proteins in immunity**

493

494 LRR motifs have been identified in proteins ranging from viruses to eukaryotes.  
495 Despite distinct evolutionary origin, many plant and animal immune receptors or factors  
496 including cell-adhesion molecules, hormone receptors, tyrosine kinase receptors or  
497 extracellular matrix-binding glycoproteins contain LRR domain. LRR proteins serve as  
498 significant mediators in both innate and, replacing immunoglobulin (Ig) domain, in  
499 alternative-adaptive immune systems where they play a dual function. Firstly, serving  
500 in pathogen sensing, through direct or indirect recognition of the pathogens, and  
501 secondly, as signal mediator, through the elicitors of downstream immune signalling  
502 they activate. LRR domains vary in their lengths and pattern of conserved residues,  
503 mediating ligand binding interactions<sup>54</sup>. In the next parts, I give an overview of distinct  
504 conserved immune protein families from different organisms, all carrying the LRR  
505 domain required for achieving their protective function.

506

507 **1.2.2.1 LRR proteins in the innate immune system**

508

509 **1.2.2.1.1 Drosophila Toll receptors**

510

511 Toll and the members of the Toll-dependent signalling pathway were initially found in  
512 *Drosophila* as involved in early embryonic development in the dorso-ventral patterning  
513 of the embryo<sup>55,56</sup>. Further studies demonstrated the contribution of Toll signalling  
514 pathway in defence against Gram-positive bacteria and fungi<sup>57-59</sup>. Toll protein is a  
515 transmembrane receptor with extracellular LRR domain that binds a ligand and an  
516 intracellular Toll/interleukine-1 receptor (TIR) domain that takes part in the intracellular  
517 signal transduction. Toll is not activated by direct recognition of a PAMP, but it binds  
518 instead to the active form of the protein ligand Spätzle (Spz) with a high affinity  
519 ( $kd=0.4nM$ )<sup>60</sup>. Spz is a cytokine that is produced as a pro-form. Spz activation takes  
520 place upon proteolytic cleavage in response to infection through serine protease  
521 cascades that lead to the activation of Spätzle-processing enzyme (SPE), which is  
522 required to activate Spz<sup>57</sup>. Spz binding to Toll induces receptor dimerization through  
523 disulphide bridges of conserved cysteine residues<sup>61</sup>. This dimerization induces an  
524 intracellular signalling, leading to the phosphorylation of the inhibitor of nuclear factor  
525 kappa B (IκB) Cactus (a negative regulator of the Toll pathway). Once phosphorylated,  
526 Cactus releases the nuclear factor kappa-B (NF-κB) transcription factor(s) Dorsal or  
527 Dif, which migrate to the nucleus and activates the transcription of several genes, such  
528 as antimicrobial peptides (AMP)<sup>57</sup>. Apart from humoral response, Toll receptor  
529 activation also induces cellular responses in *Drosophila*, such as phagocytosis,  
530 melanization or encapsulation<sup>52,57</sup>. The *Drosophila* genome encodes nine Toll  
531 receptors that all share a similar molecular structure, with an ectodomain composed  
532 mainly of LRR and cysteine-rich flanking motifs. Besides Toll (also known as Toll-1),

## GENERAL INTRODUCTION

533 only Toll-5 and Toll-9 induce the expression of the anti-fungal peptide *drosomycin*<sup>62</sup>,  
534 whereas Toll-7 was shown to interact with viral pathogens, inducing antiviral autophagy  
535 in *Drosophila*<sup>63</sup>. Therefore, Toll proteins are responsible for indirect recognition of  
536 pathogens and activation of humoral and cellular immune responses in  
537 *Drosophila*.<sup>57,61,64,65</sup>

538

### 539 **1.2.2.1.2 Toll-like receptors (TLR)**

540

541 Identification of Toll receptors in *Drosophila* opened new perspectives to understand  
542 the innate immune system, which was understudied in higher organisms in favour to  
543 the adaptive immune response. In deuterostomes, Toll-like receptors (TLRs), which  
544 are transmembrane proteins, were identified as the major PRRs that underwent  
545 purifying and diversifying selection, ranging from 222 TLR genes in sea urchin  
546 *Strongylocentrotus purpuratus* to 12 TLR in mouse<sup>66</sup>, in order to adapt to a variety of  
547 coevolving pathogens. Among mammals, human encodes at least 10 TLRs. In contrast  
548 to *Drosophila*, where the Toll-dependent pathway is induced by Gram-positive and  
549 fungal pathogens, TLR in mammals are activated upon structurally variant PAMPs  
550 from different organisms and that constitute a large repertoire of microbial structures  
551 **(Table 1)**.



## GENERAL INTRODUCTION

Toll-like receptor	Ligand(s)
	Only signals as a dimer when combined with TLR2 for all its ligands; recognizes <i>Borrelia burgdorferi</i> OspA; required for adaptive immune response
TLR1	Tri-acyl lipopeptides (bacteria, eg, <i>M tuberculosis</i> ) Soluble factors ( <i>Neisseria meningitidis</i> )
	Associates with CD11/CD18, CD14, MD-2, TLR1, TLR6, dectin 1; Lipoprotein/lipopeptides (a variety of microbial pathogens)
	Peptidoglycan
	Lipoteichoic acid
	Lipoarabinomannan (mycobacteria)
	Phenol-soluble modulins ( <i>Staphylococcus epidermidis</i> )
	Glycoinositolphospholipids ( <i>Trypanosoma Cruzi</i> )
	Glycolipids ( <i>Treponema maltophilum</i> )
	Porin ( <i>Neisseria</i> )
	Zymosan (fungi)
	Atypical LPS ( <i>Leptospira interrogans</i> )
	Atypical LPS ( <i>Porphyromonas gingivalis</i> )
	HSP70 (host)
	CMV virions
	Hemagglutinin (H) protein of wild-type measles
TLR2	Bacterial fimbriae
TLR3	Double-stranded RNA in viruses
	Gram-negative enteric LPS (requires coreceptors MD-2 and CD14)
	Additional ligands
	Chlamydial heat shock protein 60
	RSV F protein
	Taxol (plant)
	<i>M tuberculosis</i> HSP65
	Envelope proteins (MMTV)
	HSP60 (host)
	HSP70 (host)
	Type III repeat extra domain A of fibronectin (host)
	Oligosaccharides of hyaluronic acid (host)
	Polysaccharide fragments of heparan sulfate (host)
	Fibrinogen (host)
TLR4	Beta-defensin2
TLR5	Flagellin (monomeric) from bacteria
	See TLR2 (as dimers with TLR2)
	Phenol-soluble modulins
TLR6	Di-acyl lipopeptides (mycoplasma)
	Responds to imidazoquinoline anti-viral agents (synthetic compounds)
	Loxoribine (synthetic compounds)
	Bropiramine (synthetic compounds)
	Endogenous and exogenous ligands unknown.
TLR7	Single-stranded RNA
	Imidazoquinoline (synthetic compounds)
TLR8	Single-stranded RNA
TLR9	Bacterial DNA as "CpG" motifs
TLR10	Unknown

552

553 **Table 1. Role of human TLRs in pathogen recognition.** Table modified from<sup>67</sup>.

554

555 Depending on the sensed pathogen, TLRs activate intracellular signalling pathways  
556 that induce transcription of various defence factors, including inflammatory cytokines,  
557 chemokines, major histocompatibility complex (MHC) and co-stimulatory molecules<sup>68</sup>.  
558 TLR activity require dimerization of LRR ectodomain that is proposed to bring together  
559 intracellular TIR domains, thus enabling activation of downstream signal transduction  
560 process<sup>69</sup>. TLRs dimerization can be homo- and heterologous, and the variety of TLR  
561 dimer combinations increases their microbial sensing repertoire<sup>70</sup>. An example of this  
562 cooperation was demonstrated for TLR2 heterodimer with either TLR1 or TLR6, where

## GENERAL INTRODUCTION

563 the TLR2 partner determines the specificity of the ligand recognition<sup>71</sup>. In response to  
564 microbial invasion, extracellular PAMPs or DAMPs can activate TLRs on the cell  
565 surface. However, TLRs can also be found in the intracellular compartment, where  
566 they are located in the membranes of endosomes and lysosomes. In particular,  
567 intracellular TLRs sense viral and bacterial nucleic acids during viral particles  
568 endocytosis and degradation in late endosomes or lysosomes<sup>72</sup>. Moreover, it was  
569 suggested that the intracellular localization of these TLRs could enable discrimination  
570 between viral DNA and self-DNA. Lastly, secreted TLR (sTLR) were also described,  
571 e.g. sTLR2, which has been shown to arise from ectodomain shedding of the  
572 extracellular domain<sup>73</sup>. Soluble form of receptor was found in blood and breast milk to  
573 inhibit inflammatory reaction. Such reaction is proposed to work in the way that sTLR2  
574 could negatively regulate TLR2 activation by behaving as a decoy receptor<sup>74,75</sup>.

575 Taken together, the TLR family serves mainly as direct PRRs for microbial sensing  
576 that mount innate immune responses as well as playing key role in linking innate with  
577 adaptive immunity <sup>67,68,76,77,78,79</sup>.

578

### 579 **1.2.2.1.3 Nucleotide-binding leucine-rich repeat receptors**

580

581 Nucleotide-binding leucine-rich repeat receptors (NLRs) represent the major class of  
582 intracellular innate immune receptors found in both animals and plants (described  
583 below). In animals, the cytoplasmic detection system induced by NLRs was proposed  
584 to play a defence function in host tissues where TLRs are absent or expressed at low  
585 levels<sup>80</sup>. This is the case for the intestinal epithelial cells, which are in constant contact  
586 with the microbial flora. Therefore, expression of TLRs is down-regulated in these cells  
587 in order to avoid high stimulation and aberrant inflammations due to the microbial

## GENERAL INTRODUCTION

588 PAMPs from microbiome. However, in case of cell infections with invasive pathogens,  
589 PAMPs can be transferred to the intracellular compartment and be sensed by NLRs,  
590 which initiate the defence response. Proteins classified as NLR belong to diverse  
591 families. The best characterized members of the NLR family are the Nucleotide-  
592 Binding Oligomerization Domain-like (Nod) receptors Nod1 and Nod2. Both Nod1 and  
593 Nod2 recognize peptidoglycan (PGN), but each of them requires distinct PGN  
594 molecular motifs. In the centre of the proteins, NLRs are characterized by a nucleotide-  
595 binding (NB) domain followed by a variable number of highly polymorphic C-terminal  
596 LRRs and diverse N-termini. Activation of the Nod molecules occurs when a bacterial  
597 ligand is sensed by the LRR domain<sup>80</sup>. Besides recognition of PAMPs, similarly to  
598 TLRs, NLRs can also react to modified-self in the form of DAMPs<sup>81</sup>. NLRs are normally  
599 in an “off” state due to the inhibitory function of the LRR domain, thus preventing  
600 constitutive protein activation via the NB molecular switch<sup>82,83</sup>. Recognition of the  
601 pathogen effector induces the release of the inhibited LRR domain and this  
602 conformational change, through ATP binding, leads to NLR oligomerization and  
603 activation<sup>82</sup>. Activated NLRs lead to inflammatory responses mediated by NF- $\kappa$ B,  
604 Mitogen Activated Protein Kinases (MAPK), or Caspase-1 activation, accompanied by  
605 subsequent secretion of proinflammatory cytokines<sup>82</sup>.

606 Thus, NLRs are intracellular sensors of innate immunity that activate defence in cells  
607 that do not express TLRs but might also serve as a second layer of immune responses  
608 after intruders have bypassed extracellular defence and invaded a cell.<sup>80,84,82,85</sup>

### 609 **1.2.2.1.4 Plant LRRs**

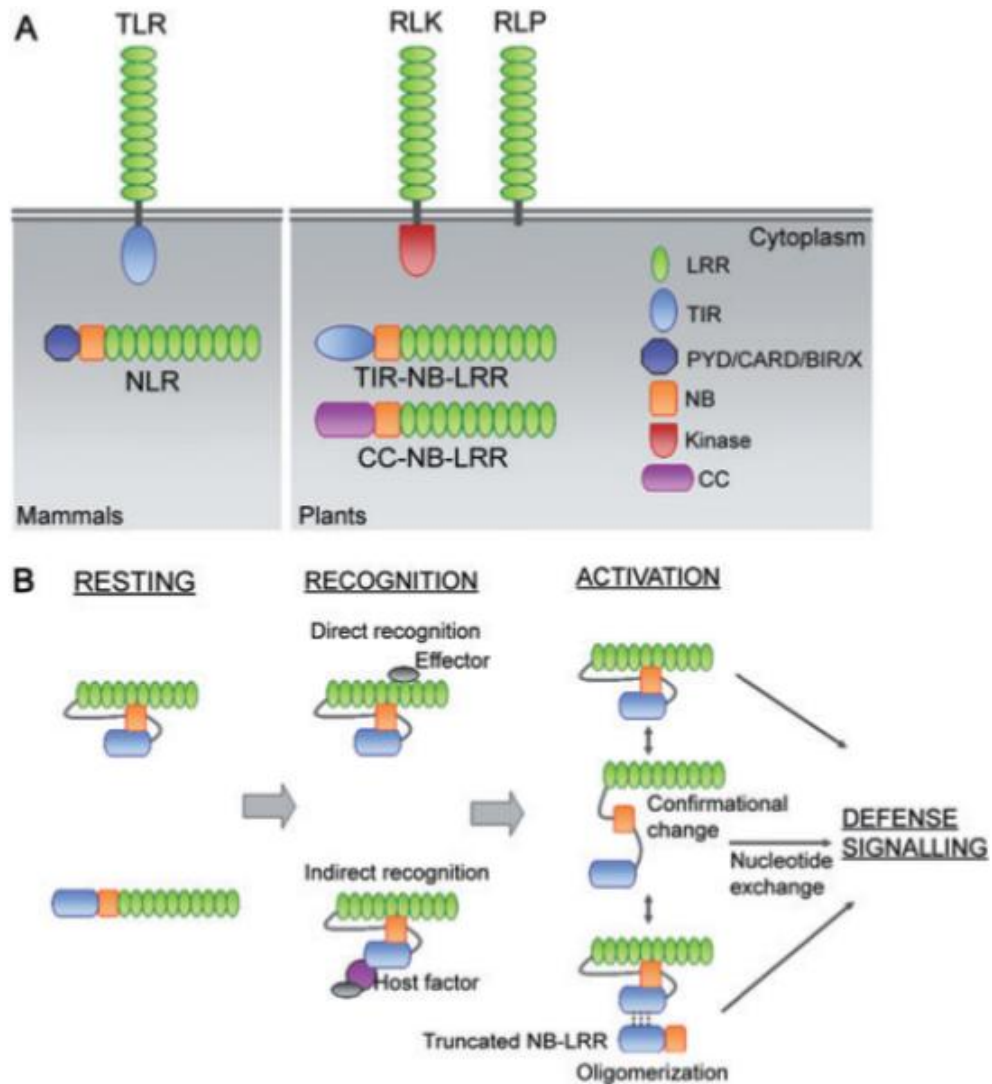
610

## GENERAL INTRODUCTION

611 Plants lack adaptive immune response. Nevertheless, during evolution plants acquired  
612 multilayer innate immune mechanisms. In this system, LRR proteins play a major role  
613 across both transmembrane and cytoplasmic PRRs that sense the pathogens and  
614 transduce the signals into activation of immune responses. Transmembrane LRR  
615 proteins compose the first layer of defence, sensing the extracellular signals from  
616 pathogens. This group of receptors is all composed of extracellular LRR domain and  
617 variable, intracellular C-terminus. Therefore, we can distinguish receptor-like kinase  
618 (RLK) proteins that have intracellular kinase domain and receptor-like protein (RLP)  
619 that lack intracellular kinase domain<sup>86</sup> (**Figure 3A**)

620

## GENERAL INTRODUCTION



621 **Figure 3. LRR-containing immune receptors, modes of pathogen perception and**  
 622 **activation in plants. A.** An overview of LRR-containing immune receptors in mammals and  
 623 plants. **B.** NB-LRR immune signalling activation pathway. Figure reproduced from<sup>86</sup>.

624

625 Based on N-terminal domain composition, the NB-LRR proteins can be categorized as  
 626 two types. The common core is composed of C-terminal LRR domain and a central  
 627 NBS domain, whereas the variable N-terminal domains carry coiled-coil (CC) domain  
 628 or a domain homologous to *Drosophila* Toll/human interleukin-1 receptor (TIR)<sup>86</sup>  
 629 (**Figure 3A**). These receptors mediate the activation of the mitogen-activated protein  
 630 kinase (MAPK) pathway that in turn leads to increased transcription of pathogen-  
 631 responsive genes, production of reactive oxygen species (ROS), all of which contribute  
 632 to prevention of pathogen spreading<sup>86,87</sup>. Nevertheless, if a microbe overcomes the

## GENERAL INTRODUCTION

633 first line of host-defence machinery and enters the cell, a second layer of defence takes  
634 over. Plant evolution has generated proteins that detect specific effector molecules, in  
635 a mechanism called effector-induced immunity (ETI)<sup>88</sup>. Plant effector-induced  
636 immunity is more akin to mammalian adaptive immunity in that pathogen effectors and  
637 virulence factors are specifically recognized. The genes encoding the specificity  
638 determinants of effector-induced immunity are known as resistance (R) genes. Most  
639 known R genes encode proteins belonging to the family of nucleotide binding (NB)  
640 LRR proteins (NB-LRR). NB-LRR proteins are involved in the recognition of specialized  
641 pathogen effectors (also called avirulence (Avr) proteins) that are thought to provide  
642 virulence function in the absence of the proper R gene<sup>88</sup>. Nevertheless, unlike  
643 mammalian adaptive immunity, plant defence relies on innate immunity. Therefore,  
644 their capability to recognize specifically a large repertoire of pathogen effectors  
645 suggests they encode a high number of NB-LRR genes. Indeed, in the genome of  
646 *Arabidopsis* there are approximately 150 predicted NB-LRRs<sup>87</sup>. However, this number  
647 is probably still below the wide range microbes that plants have to face. This questions  
648 of how an effector-induced immune response in plants coordinates resistance to a  
649 broad range of pathogens and their corresponding effectors, in sometimes very long-  
650 lived individuals, up to centuries, is still not completely clarified. It was proposed that  
651 NB-LRR bacterial effector recognition likely evolved as an indirect mechanism. Based  
652 on the fact that the majority of characterized bacterial effectors display enzyme activity  
653 that modify plant proteins, enzymatic functions of multiple distinct pathogen-effectors  
654 target the same host proteins. Therefore, rather than developing receptors for every  
655 possible effector, host plants have evolved mechanisms to monitor common plant  
656 proteins that are targets of pathogen virulence. Thus, by monitoring perturbations in  
657 key host proteins, R proteins such as the NB-LRR, indirectly detect the functional

## GENERAL INTRODUCTION

658 enzymatic activity of multiple or many pathogen-effectors<sup>87</sup> (**Figure 3B**). NB-LRRs are  
659 normally kept in a resting state by intramolecular interactions involving the LRR  
660 domain. Pathogen recognition occurs either by direct association between the LRR  
661 domain and a pathogen-effector protein or by indirect association mediated by a host  
662 protein (**Figure 3B**). Upon ligand binding, conformational changes and/or protein  
663 oligomerization leads to the activation of NB-LRR (**Figure 3B**). This is accompanied or  
664 followed by exchange of ADP for ATP at the NB domain that leads to defence  
665 signalling. The activity of NB-LRR proteins result in activation of defence factors  
666 transcription in nucleus or plant hypersensitive response (HR), which is a form of  
667 programmed cell death localized to infection sites<sup>86,89,90</sup>.

668

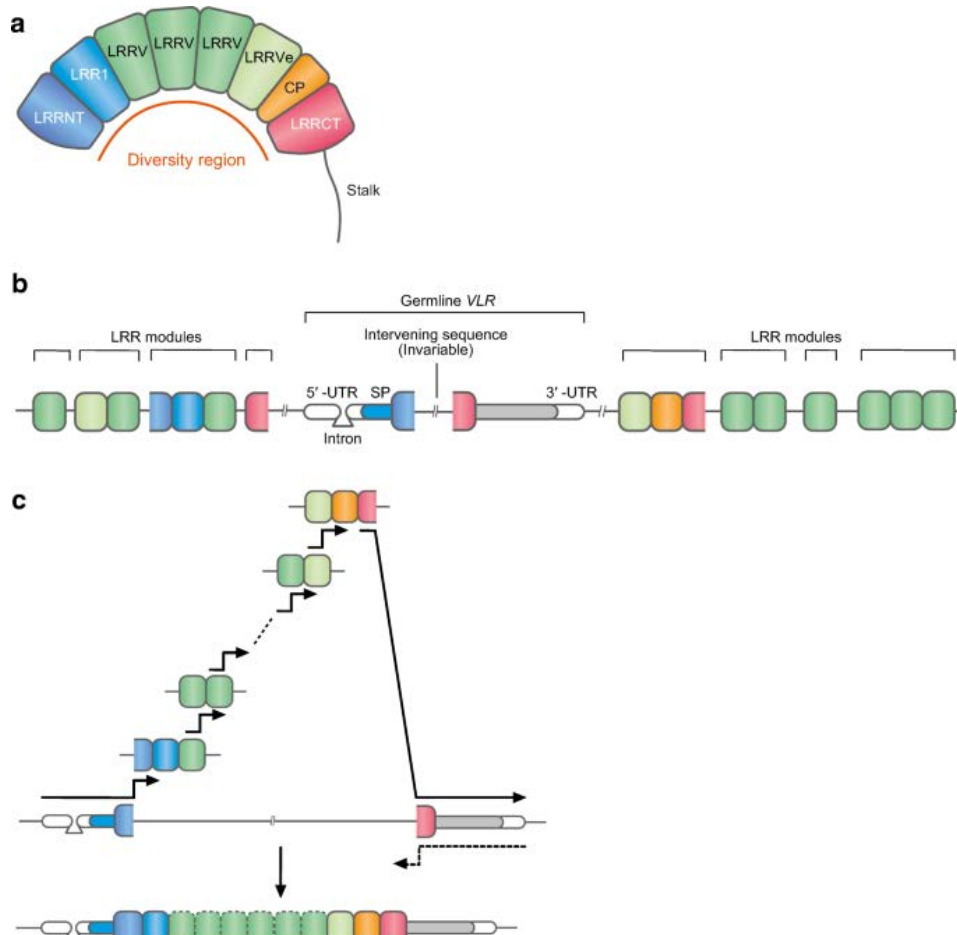
### 669 **1.2.2.2 LRR-coding receptors in adaptive-like immunity of jawless vertebrates**

670

671 Jawless vertebrates (agnathans) and jawed vertebrates (gnathostomes) diverged  
672 approximately 550 million years ago. While jawed vertebrates are predominant, the  
673 jawless clade has only two extant representatives: hagfishes and lampreys. As  
674 mentioned above, the arsenal of adaptive immune recognition in gnathostomes is  
675 based on Ig domain B and T lymphocyte receptors (BCRs and TCRs) and major  
676 histocompatibility complex (MHC) molecules. While none of these proteins are present  
677 in agnathans, hagfishes and lampreys were shown to mount adaptive-like immune  
678 responses to repetitive antigenic challenges through the expression of unique antigen  
679 receptors named variable lymphocyte receptors (VLRs). These VLRs generate  
680 antigen-recognition diversity through a gene conversion-like mechanism involving a  
681 variable LRR segment domain<sup>91</sup>. Three VLR genes have been identified, VLRA, VLRA  
682 and VLRC and they share a similar protein organization with a N-ending LRR-capping  
683 motif (LRRNT) followed by 18-residue N-terminal LRR module (LRR1), diverse region

## GENERAL INTRODUCTION

684 composed of multiple, variable LRR modules (LRRV), 24-residue called end LRRV  
685 (LRRVe), a connecting peptide (CP) and a C-terminal LRR-capping motif (LRRCT,  
686 **Figure 4**)<sup>92</sup>.



687  
688 **Figure 4. Schematic structure of a VLR protein and VLR gene assembly. A.** VLR  
689 monomers adopt a horseshoe-like solenoid fold characteristic of the LRR protein family. **B.**  
690 The germline VLR gene has a non-functional intervening region instead of the diversity region.  
691 Sequences coding for LRR modules (color-coded as in panel **A**) are located in multiple copies  
692 adjacent to the germline VLR gene. **C.** In developing lymphocytes, VLR is assembled by a  
693 gene conversion-like mechanism. Gene conversion can begin from either end of the gene.  
694 Figure reproduced from<sup>92</sup>.

695

696 This diverse region (LRRV) composed of multiple LRR modules determines receptor  
697 specificity. VLRA, VLB and VLB are a membrane-bound protein receptor. However, in  
698 addition to be a transmembrane protein, VLRB can also be secreted as a soluble  
699 protein. Secreted VLRB molecules occur as pentamers or tetramers of dimers and



## GENERAL INTRODUCTION

700 have eight to ten antigen binding sites<sup>93</sup>, leading to the neutralization of the pathogen  
701 and subsequent elimination by phagocytosis or complement pathway activation<sup>92</sup>. In  
702 summary, suggesting a convergent evolution, jawed and jawless vertebrates use  
703 remarkably different modules (Ig and LRR, respectively) in common strategies, forming  
704 transmembrane or soluble antibody receptors as the arm of adaptive immune response  
705 <sup>91,92,93,94,95</sup>.

706

707

708

## **II. CURRENT KNOWLEDGE, GAPS AND OBJECTIVES**

709 **2.1 Natural resistance to *Plasmodium* highlighted a family**  
710 **of LRR genes termed APL1 (*Anopheles Plasmodium-***  
711 **responsive leucine-rich repeat)**

712 Like all invertebrates, mosquitoes lack adaptive immune response, thus using their  
713 innate immune system to face microbial invasion. The availability of complete genome  
714 sequences as well as comparative genome studies between mosquitoes and the  
715 arthropod model organism, *Drosophila*, contributed to identification of mosquito  
716 immune factors<sup>96,97</sup>. However, *Drosophila* and *Anopheles* diverged from a common  
717 ancestor around 250 million years ago and have extremely different lifestyles,  
718 especially with the obligate hematophagy of mosquito females. Therefore, while the  
719 main immune pathways were found to be conserved, flies and mosquitoes are exposed  
720 to different microbial repertoires and thus underwent different evolutionary pressures  
721 that have shaped their respective immune systems. While it is difficult to measure  
722 fitness trade-offs in nature, identifying the molecular components of natural resistance  
723 could help understand how they are genetically structured in wild mosquito vector  
724 populations<sup>98</sup>. A field-based genetic mapping survey of wild mosquito populations  
725 revealed that about 50% of the *A. gambiae* s.s. tested carried a significant locus  
726 controlling susceptibility to natural *P. falciparum*. This genomic locus located on  
727 chromosome 2L was mapped in both East and West African wild *A. gambiae*<sup>99,100,101</sup>,  
728 suggestive of a widespread and shared genetic mechanism of malaria resistance.  
729 Candidate gene filtering of the locus by bioinformatics and transcriptional regulation  
730 highlighted a family of three genes encoding LRR proteins and named *Anopheles*  
731 *Plasmodium*-responsive leucine-rich repeat 1 (APL1). The genes, APL1A, APL1B and  
732 APL1C, share  $\geq 50\%$  peptide identity and have distinct functional activities<sup>102,103</sup>.

## CURRENT KNOWLEDGE, GAPS AND OBJECTIVES

733 Functional analysis performed in the laboratory by measuring the presence of midgut  
734 oocysts after gene silencing showed that a functional APL1A gene protects against  
735 infection with *P. falciparum*, while functional APL1C protects against the rodent malaria  
736 parasites, *P. berghei* and *P. yoelii*<sup>102,103</sup>. Furthermore, protection against *P. falciparum*  
737 correlates with the transcriptional regulation of APL1A by the isoform Imd/Rel2-S but  
738 not Imd/Rel-2F<sup>103</sup>, whereas protection against *P. berghei* correlates with the  
739 transcriptional regulation of APL1C by the Toll/Rel1 pathway<sup>102</sup>. Gene APL1B protects  
740 against both human and rodent malaria parasites in collaboration with APL1A and  
741 APL1C, respectively<sup>104</sup>. The APL1 genes display high polymorphism with a rate of non-  
742 synonymous variation 10-fold higher than any another *A. gambiae* genes. This pattern  
743 is consistent with adaptive maintenance of polymorphism, which can be seen in so-  
744 called “trench warfare” models of fluctuating recurrent pathogen epidemics<sup>105</sup>. Such  
745 polymorphism results in different encoded protein variants that likely lead to protective  
746 phenotypes against distinct pathogens. Indeed, in a study of the *A. gambiae* Ngousso  
747 colony, which segregates three different major structural alleles of APL1A, only one  
748 allele was associated with protection against *P. falciparum*<sup>106</sup>. This observation  
749 highlights that APL1 allelic variation could contribute to natural variation in mosquito  
750 susceptibility to *Plasmodium*, and also that studies in different mosquito colonies  
751 should take into consideration presence of different genetic forms that could affect the  
752 phenotype.

753

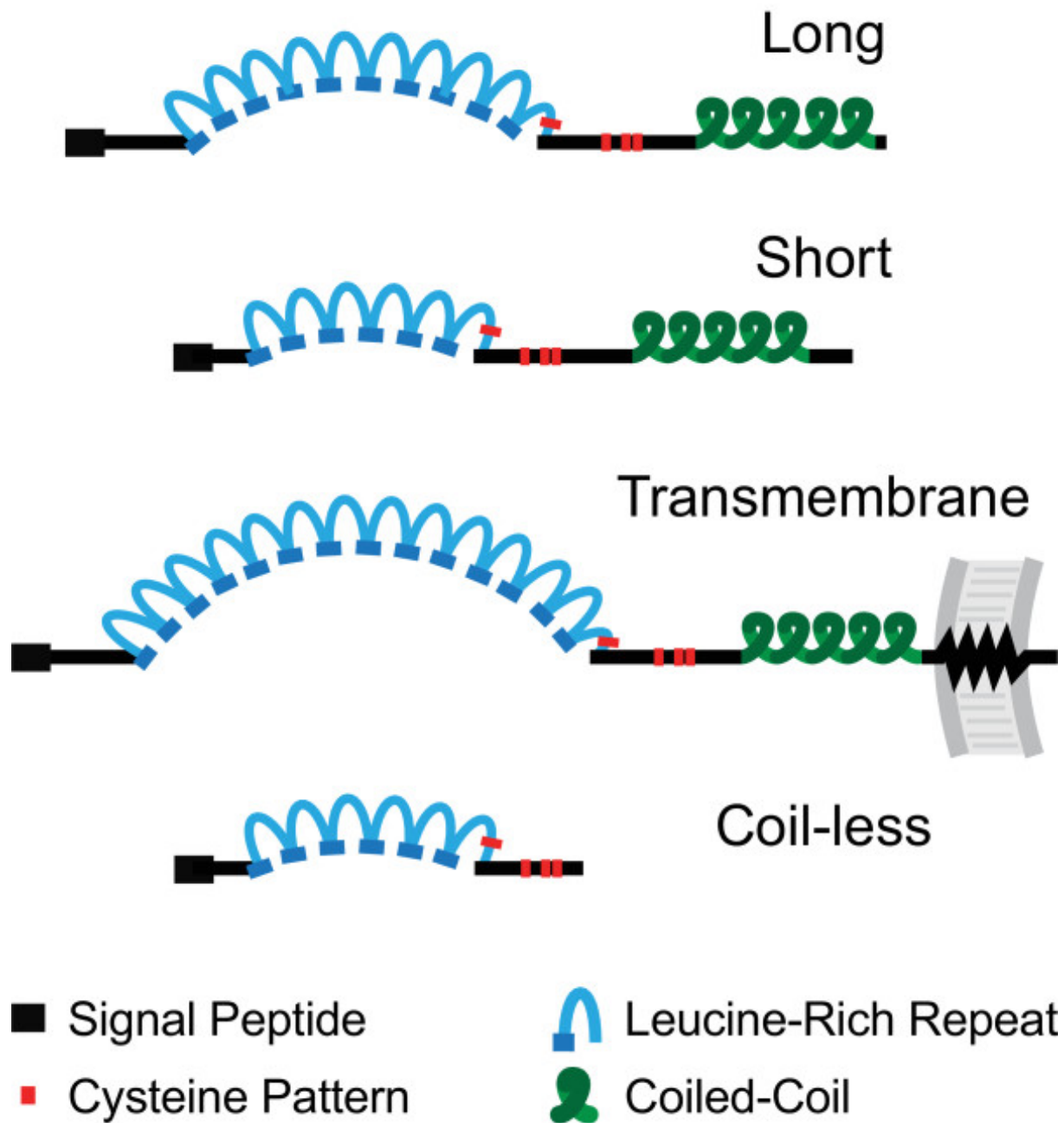
754

755 All three APL1 gene genes are comprised of a small 5' exon followed by a second  
756 longer exon separated by a short intron. The encoded proteins are characterized by  
757 an N-terminal signal peptide, a series of LRR motifs spanning ~300 amino acids and

## CURRENT KNOWLEDGE, GAPS AND OBJECTIVES

758 a coiled-coil domain at the C-terminus. In addition to common core sequence among  
759 the three genes, APL1C (and some allelic variants of APL1A) also carry an N-terminal  
760 repeated motif of the amino acids P-A-N-G-G-L (called PANGGL region), with an  
761 unknown function<sup>107 106</sup>.

762 Besides the common LRR domain, some of the other LRR genes in *A. gambiae* share  
763 a similar protein structure as the APL1 gene members and together constitute a  
764 subfamily of LRR immune genes, called leucine-rich repeat immune molecules  
765 (LRIM)<sup>107</sup>. Apart from LRR, these proteins include a coiled-coil domain and patterns of  
766 cysteine residues that give possibilities for disulphide bridges between two protein  
767 partners. The coiled-coil domain was proposed to drive or consolidate interactions  
768 such as combinatorial oligomerizations to promote the orientation of the cysteine  
769 residues for stabilizing the disulphide bridges. LRIMs also have a signal peptide with,  
770 for some members, a transmembrane domain, suggesting they are secreted into the  
771 mosquito hemolymph or exposed on cell membranes. The PANGGL domain present  
772 in APL1C and some APL1A allelic forms is exceptional for these proteins and was not  
773 found in any other mosquito protein<sup>107</sup>. The LRIM family can be classified in four  
774 groups: “Long” subgroup that includes APL1 genes, the “Short” subgroup, the  
775 Transmembrane (TM) LRIM and Coil-less subgroups. The “Long” and “Short”  
776 subgroups differ each other from the length of their LRR domain<sup>107</sup>. The variable LRRs  
777 domains among LRIMs form a stable arc or horseshoe-like fold that could facilitate  
778 protein interactions (**Figure 5**).



779

780 **Figure 5. Classification of mosquito LRIM proteins.** Figure reproduced from<sup>107</sup>.

781

782 The structure of LRIMs promotes genes interaction and was best studied in the case  
 783 of APL1C, which forms a heterodimer with LRIM1 via interaction of their C-terminal  
 784 coiled-coil domains and a single disulphide bond<sup>108</sup>. This dimerization is absolutely  
 785 necessary for the antimalarial activity of the *Anopheles* complement-like protein, TEP1  
 786 (thioester-containing protein)<sup>109,110</sup>.

## CURRENT KNOWLEDGE, GAPS AND OBJECTIVES

787 Nevertheless, not only LRIMs, but other LRR proteins can biochemically interact in  
788 binding assays with APL1C or LRIM1, and possess the same protective phenotypes  
789 as LRIM family genes, which could generate combinatorial complexity for immune  
790 recognition and response<sup>104,107,111,112</sup>.

791

792

### 793 **2.2 APL1 functional activity and the mosquito complement** 794 **system**

795

796 APL1C and LRIM1 form a heterodimer necessary for the antimalarial activity of the  
797 *Anopheles* complement-like protein, Thioester-containing Protein 1 (TEP1)<sup>108,111</sup>. The  
798 two other APL1 genes (APL1A and APL1B) also have the capacity to heterodimerize  
799 with the LRIM1 protein, suggesting that LRIM1 may act as a common core protein for  
800 different heterocomplex formation with APL1 as well as other LRR proteins, including  
801 the LRIM members<sup>104</sup>. This is consistent with the role of LRIM1 as an antagonist of  
802 both *P. berghei* and *P. falciparum* midgut infection, and also against Gram-negative  
803 bacteria<sup>104,113,114</sup>.

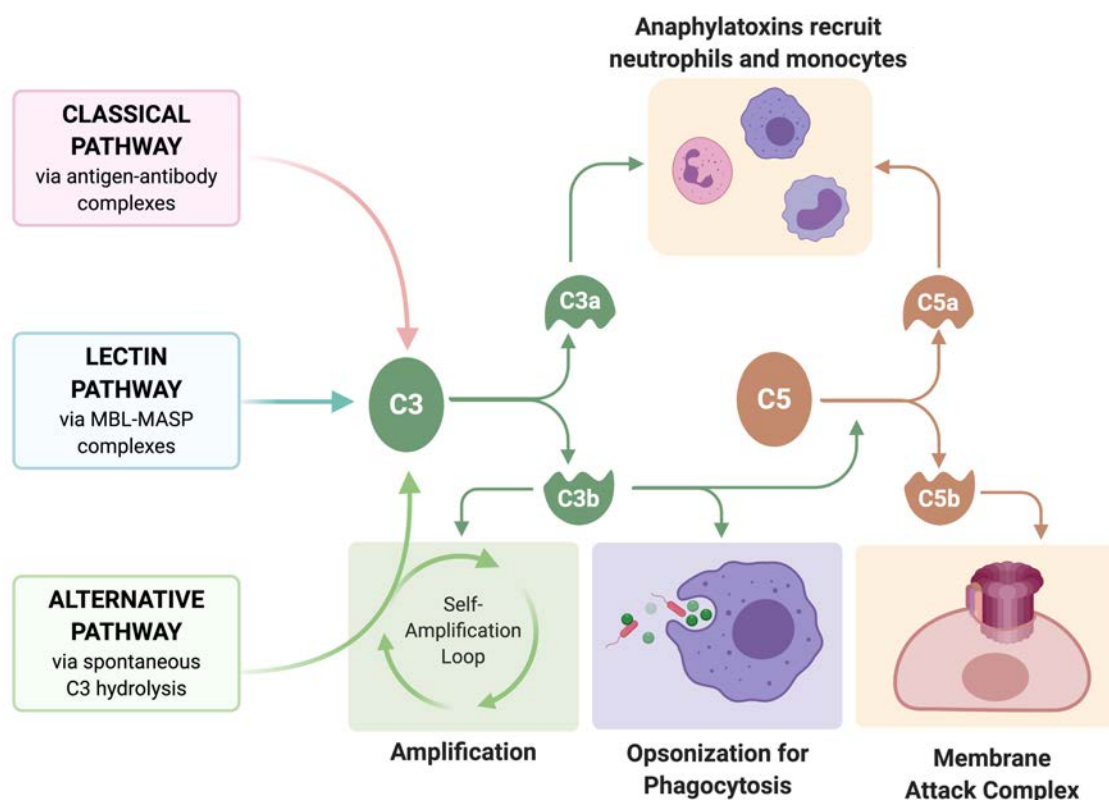
804 TEP1 is a structural and functional homolog of vertebrate complement factor C3 and  
805 of members of the  $\alpha$ 2-macroglobulin family (A2Ms)<sup>115,116</sup>.

806 In vertebrates, the complement system is composed of soluble and cell-membrane  
807 proteins, acting together in the cascade leading to pathogen clearance through their  
808 phagocytosis or membrane attack complex (MAC) formation. There are three  
809 complement pathways (**Figure 6**): i) the classical pathway, which is initiated by an  
810 antibody-antigen complex ii) the lectin pathway where the mannose or different sugar

## CURRENT KNOWLEDGE, GAPS AND OBJECTIVES

811 moieties present on the microbial surface are bound by mannan-binding lectin (MBL)  
812 iii) the alternative pathway which is initiated by the spontaneous hydrolysis of the C3  
813 factor. This pathway is either self-amplified on the surfaces of microbes or inhibited in  
814 the case of host self-surfaces. Despite distinct arms of pathogen recognition, all three  
815 pathways induce subsequent proteolysis events and converge to the central C3 factor.  
816 Upon proteolytic cleavage, C3 produces anaphylatoxin C3a that recruits neutrophils  
817 and monocytes at the site of infection. This promotes inflammation and C3b opsonin  
818 inducing phagocytosis and/or downstream reactions with other complement factors, all  
819 ending with MAC formation (Figure 6)<sup>117,118</sup>.

### Roles of the Complement Cascade in Innate Immunity



820

821 **Figure 6. Complement pathway activation and outcome in vertebrates.**

822

823 In *A. gambiae*, the TEP1-mediated complement-like pathway is the best described



## CURRENT KNOWLEDGE, GAPS AND OBJECTIVES

824 immune effector mechanism. TEP1 is secreted into the hemolymph as a full-length  
825 protein before being cleaved to the active form. The cleaved TEP1 form called TEP1<sub>cut</sub>  
826 is stabilized in the hemolymph by the heterodimer LRIM1/APL1C to form a complex.  
827 This interaction with the LRR heterodimer is necessary to maintain a functional  
828 thioester bond that can covalently attach to the parasite surface<sup>119</sup>. In the absence of  
829 APL1C or LRIM1, TEP1 is unable to specifically bind to the parasite surface, and  
830 deposits apparently non-specifically on self-tissues, the thioester bond being  
831 hydrolysed over time, resulting in non-functional precipitation of TEP1<sup>109,110</sup>. TEP1  
832 was demonstrated to bind to different pathogens, including Gram-positive and Gram-  
833 negative bacteria and *P. berghei* ookinetes in mosquito midguts<sup>115</sup>, mediating their  
834 killing<sup>120</sup>. Like the APL1 gene family, TEP1 genes are highly polymorphic in *A. gambiae*  
835 genome with distinct alleles correlating with susceptibility or resistance to  
836 *Plasmodium*<sup>115,121</sup>. Variation of resistance and susceptible alleles are associated with  
837 variable loops within the thioester domain<sup>119</sup>.

838 Besides the interaction with TEP1, the APL1C-LRIM1 heterocomplex can also interact  
839 with other TEP family members, such as TEP3 and TEP4, indicating that LRIM  
840 members can form multiple distinct immune complexes<sup>111</sup>. Functional analysis *A.*  
841 *coluzzii* detected a protective phenotype of TEP4 for *P. falciparum* but not rodent  
842 malaria species, whereas TEP3 displayed the inverse protective phenotype<sup>104</sup>. This  
843 protective specificity mirrors the distinct antagonistic activity observed for APL1A and  
844 APL1C proteins against *P. falciparum* and *P. berghei*, respectively.

845

### 846 **2.3 Scope of the thesis**

## CURRENT KNOWLEDGE, GAPS AND OBJECTIVES

847 The phenotype of APL1 function in the anti-*Plasmodium* response has been described,  
848 but the molecular mechanism of activity is not known. The only structural mechanistic  
849 information is related to their ability to form, in the case of APL1C, a heterodimer with  
850 another LRR, LRIM1, which stabilizes the active TEP1<sub>cut</sub> protein. In addition, unlike  
851 LRIM1, which dimerizes with all three APL1 genes, APL1A and APL1C display a  
852 discriminating protective phenotype against human and rodent malaria parasites  
853 species, respectively. Consistently, a recent study revealed specificity in complement-  
854 like pathway activation towards Gram-negative bacteria through protease cleavage of  
855 the APL1C C-terminus in the LRIM1/APL1C complex<sup>122</sup>. Thus, LRIM1 which has the  
856 capacity to bind to the three APL1 genes, might constitute a common structural scaffold  
857 where activities directed towards distinct pathogens may be determined by the APL1  
858 partner. Therefore, I postulate the following hypotheses:

- 859 1. APL1C might act as a pathogen specificity component in its heterocomplex with  
860 the common scaffold component LRIM1, which would explain how the  
861 complement-like effector activity of the otherwise nonspecifically-binding TEP1  
862 protein is brought into proximity with the target pathogen surface.
- 863 2. Because APL1C is secreted into the hemolymph, APL1C could also display  
864 functional activity against free sporozoites in the hemocoel to limit *Plasmodium*  
865 invasion of the salivary glands and thus transmission.
- 866 3. APL1 proteins, similar to other LRR proteins, may act in signal transduction to  
867 initiate immune signalling cascades, and if so, the architecture of the pathways  
868 may be reflected at least in part by transcriptional regulation of target genes.

869 To examine these assumptions, I used as a study model the rodent malaria *P. berghei*  
870 and *A. coluzzii* mosquitoes, and thus focused on APL1C as a model system for the  
871 APL1 immune gene family.

## RESULTS

872

873

### **III. RESULTS**

874

### 875 **3.1 APL1C protein is a pathogen binding factor for the** 876 **midgut ookinete stage of *Plasmodium***

877

#### 878 **3.1.1 Blood feeding induces an extracellular layer of APL1C** 879 **surrounding the midgut**

880

881 The active, cleaved TEP1 form is bound in the hemolymph by the heterodimer  
882 LRIM1/APL1C to form a complex to allow specific deposition on the surface of the  
883 ookinetes exiting the mosquito midgut<sup>115,109</sup>. However, while TEP1 was detected on  
884 midgut ookinetes, the localisation of the LRR partners during this process is not known.  
885 Therefore, I detected APL1C on mosquito midguts after a *P. berghei*-infected  
886 bloodmeal (IBM) 24 h post-feeding by immunofluorescence assay (IFA) on  
887 permeabilized midguts, using a polyclonal anti-APL1C antibody (pAb)<sup>102</sup>. Non-infected  
888 bloodmeal (NBM) and unfed (UF) midguts served as the controls (for the experimental  
889 design see Annex **Figure S1**). IFA followed by confocal microscopy analysis showed  
890 that while no APL1C signal was detected on UF midguts, a layer of APL1C protein  
891 surrounding the midgut epithelial cells was detected on both NBM and IBM midguts  
892 (**Figure 7A**). Quantitative and comparative analysis of the APL1C protein signal  
893 intensity between IBM, NBM and UF mosquito midguts confirmed significant statistical  
894 difference either between UF and NBM ( $p < 0.0001$ ), or between UF and IBM  
895 ( $p < 0.0001$ ) (**Figure 7B**). The comparison between IBM and NBM midguts was not  
896 statistically different (Annex, **Table S1**). I queried the exact APL1C positioning through  
897 reconstructed optical sections of confocal stack midgut images (**Figure 7C**, projected

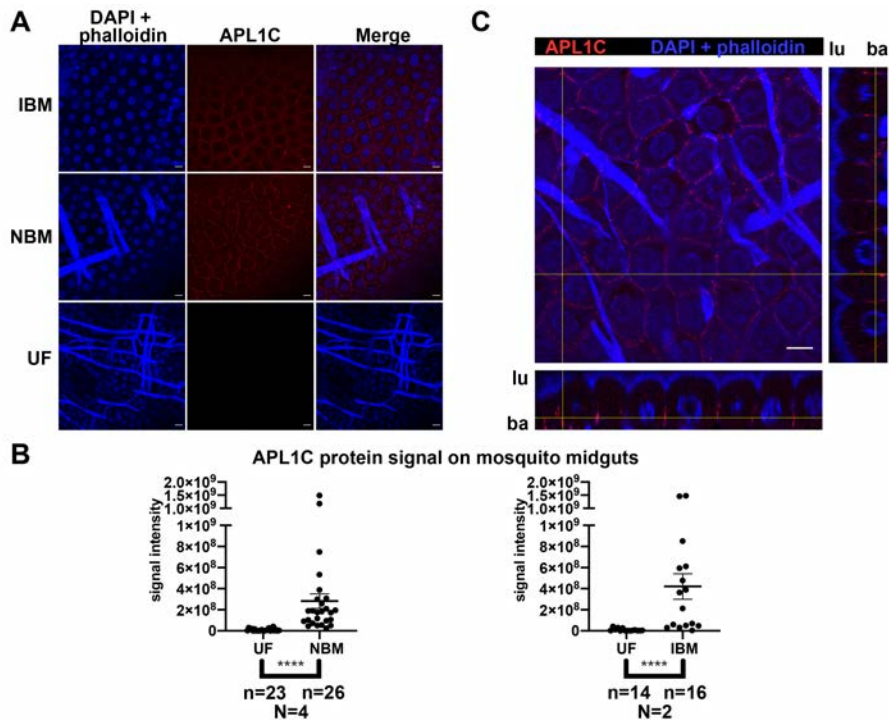
## RESULTS

898 as orthogonal views on the right side and below the stack picture). This analysis  
899 revealed that APL1C protein was localised on the basal side of the midguts tissue.

900 The simplest interpretation is that the observed IFA midgut presence of APL1C is a  
901 proxy for the concentration of systemic APL1C in the mosquito hemolymph, which  
902 diffuses through the porous basal lamina<sup>123</sup>. This conclusion is strengthened by the  
903 similar localisation pattern detected using a diffusible dextran the size of APL1C  
904 (**Figure S2**). I cannot eliminate the possibility that the midgut extracellular APL1C  
905 detected after a bloodmeal could be specifically targeted to the midgut, although there  
906 is no known mechanism that would explain this. Thus, I conclude that mosquito blood  
907 feeding causes elevated levels of APL1C protein within the hemolymph, which can be  
908 conveniently measured using dissected midguts as a proxy assay that is simpler than  
909 perfusion of hemolymph from individual mosquitoes. The APL1C in the extracellular  
910 space surrounding the midgut epithelium could serve as part of an initial immune  
911 sentinel barrier.

912

## RESULTS



913

914 **Figure 7. Blood feeding induces the presence of APL1C protein on the basal side of the**  
 915 **mosquito midgut. A.** Immunostaining analysis of the midguts from mosquitoes subjected to  
 916 *P. berghei*-infected bloodmeal (IBM), normal bloodmeal (NBM) and unfed control (UF).  
 917 Depicted pictures are representative for at least two replicates in each group. For the in vivo  
 918 system, all immunostaining analysis were performed with an anti-APL1C pAb followed by  
 919 fluorophore-conjugated IgG secondary antibody (red). Nuclei and actin were stained with DAPI  
 920 and Phalloidin (blue). The scale bar is 10  $\mu$ m. **B.** Quantitative analysis of the APL1C signal  
 921 between IBM, NBM and UF midguts reveals increase of APL1C protein on the midguts after  
 922 IBM and NBM. Graph labels and statistical tests for this figure: tests of APL1C protein signal  
 923 intensity on the midgut indicate each midgut subjected for the analysis as a dot, bars represent  
 924 mean with  $\pm$ SEM. Sample sizes (N) show the number of independent replicate experiments,  
 925 (n) the total number of midguts dissected across replicates. Data were analysed by measuring  
 926 APL1C signal intensity (RawIntDen) using ImageJ v1.52p and the intensity from each midgut  
 927 was compared between the two conditions by Mann-Whitney test. All statistical differences  
 928 were first tested independently within replicates (individual p-values in Annex, **Table S1**), and  
 929 if individual replicates showed a common trend of change, individual p-values were combined  
 930 using the meta-analytical approach of Fisher (significance level of Fisher-combined p-value:  
 931 \*\*\*\* p-value <0.0001). **C.** The XZ and YZ orthogonal views of the NBM stack picture show  
 932 APL1C protein localising extracellularly, on the basal side (ba) of the midguts. No APL1C  
 933 signal was found on midgut lumen (lu). The scale bar is 10  $\mu$ m.

934

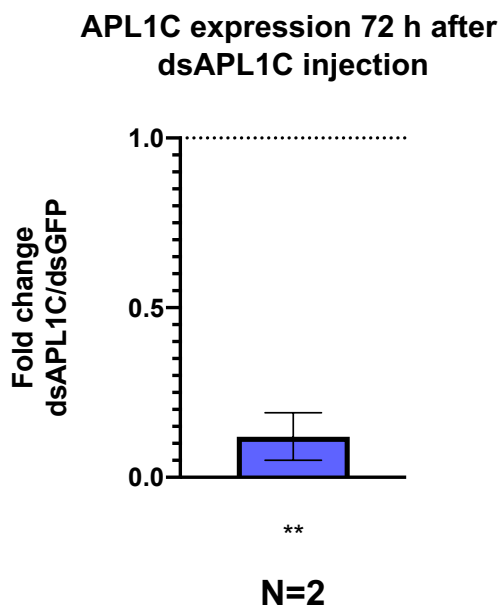
935 The specificity of the anti-APL1C pAb was confirmed on midguts collected from  
 936 mosquitoes depleted for APL1C gene expression by RNAi-mediated gene silencing.

937 *A. coluzzii* mosquitoes were injected with double-stranded RNA of APL1C gene

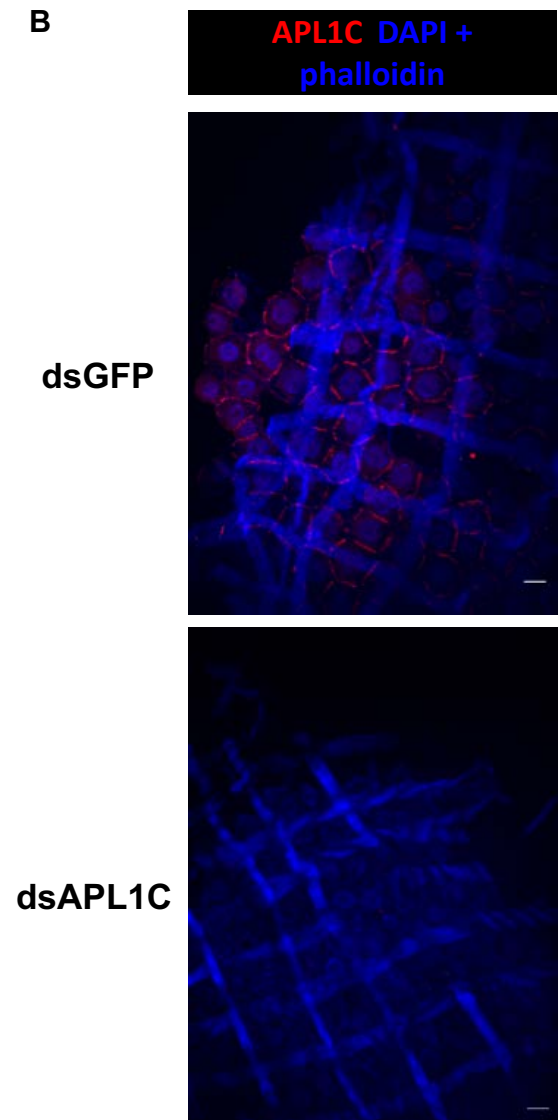
## RESULTS

938 (dsAPL1C) or dsGFP (for control group) and blood fed 72 h post-treatment. APL1C  
939 gene expression was efficiently silenced at 72 h post dsAPL1C treatment as compared  
940 to dsGFP-injected controls, measured by quantitative polymerase chain reaction  
941 (qPCR, **Figure 8A**). IFA analysis of the midguts 24 h after NBM revealed APL1C  
942 protein signal in dsGFP controls and no signal in APL1C-silenced midguts (**Figure 8B**).  
943 Hence, the IFA signal observed on the mosquito midgut samples is specific for  
944 presence of the APL1C protein.

A



B



945

946 **Figure 8. Anti-APL1C pAb labelling is specific to the APL1C protein.** A. APL1C silencing  
947 was verified by qPCR measurement between dsAPL1C and dsGFP (dotted line) 72 h post-  
948 treatment (corresponding to expression status at the time of the NBM). The ratio of the

## RESULTS

949 normalized APL1C cDNA detection in “dsAPL1C” versus “dsGFP” treatments was computed  
950 using triplicates from the same cDNA dilution. Graph represents mean with  $\pm$ SEM of the  
951 expression fold change between “dsAPL1C” and “dsGFP” control from two biological replicates  
952 (N=2). Data for qPCR analysis was analysed by unpaired t-test (significance level of t-test p-  
953 value: \*\* p-value <0.01). **B.** Immunostaining analysis of NBM mosquito midgut after APL1C  
954 gene silencing reveals no APL1C signal on APL1C-silenced midguts. APL1C protein images  
955 are from midguts of dsGFP-injected control mosquitoes. The scale bar is 10  $\mu$ m.

956

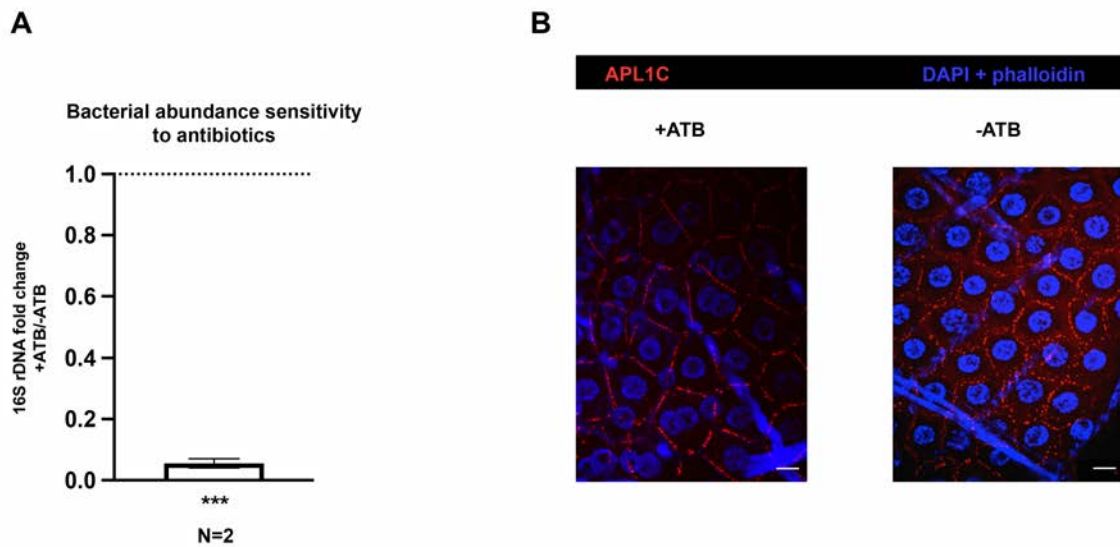
### 957 **3.1.2 APL1C presence on the basal side of the midgut epithelium is** 958 **microbiome independent**

959

960 I showed that NBM is sufficient to cause presence of the APL1C protein on *A. coluzzii*  
961 midguts. Since mosquito blood feeding induces expansion of the bacterial flora<sup>39</sup>, I  
962 queried the role of the microbiome as a potential signal promoting elevated APL1C  
963 abundance on the midgut epithelium. Prior to a bloodmeal, two groups of mosquitoes  
964 were treated or not with a wide-spectrum antibiotic (ATB) cocktail. The ATB efficiency  
965 on the microbiome was confirmed by qPCR measuring bacterial ribosomal DNA (rDNA  
966 16S) (**Figure 9A**). After the bloodmeal, IFA analysis showed that APL1C presence was  
967 still elevated on the midguts in the +ATB midguts group, and with an intensity similar  
968 to that of -ATB control group (**Figure 9B**), indicating the bacteria microbiome  
969 expansion post-NBM does not influence APL1C abundance detected on the midgut.



## RESULTS



970

971 **Figure 9. The mosquito microbiome does not influence APL1C protein abundance on**  
972 **the midgut 24 h after NBM. A.** Bacterial abundance was significantly reduced after antibiotic  
973 treatment. Antibiotic efficiency on bacterial abundance was confirmed by qPCR detection of  
974 16S rDNA in controls (-ATB, dotted line indicates  $y=1.0$ ) and ATB-treated (+ATB), 24h post  
975 NBM. The ratio of the normalized 16S rDNA detection in "+ATB" versus "-ATB" treatments was  
976 computed using triplicates from the same cDNA dilution. Graph represents mean with  $\pm$ SEM  
977 of the expression fold change between "+ATB" and "-ATB" control from two biological  
978 replicates (N=2). Bacterial qPCR data were analysed by unpaired t-test using (\*\*\*) p-value  
979  $<0.001$ ). **B.** Immunostaining analysis of the +ATB midguts group shows that bacterial depletion  
980 did not diminish APL1C abundance on NBM midguts. Images are representative of two  
981 independent biological replicates of 3–7 midguts each. The scale bar is 10  $\mu$ m.

982

### 983 3.1.3 APL1C binds to *Plasmodium* ookinetes exiting the basal side 984 of the midgut epithelium

985

986 Elevated hemolymph APL1C levels are induced by NBM, and are also present in the  
987 hemolymph within the midgut extracellular space delimited by the basal lamina.

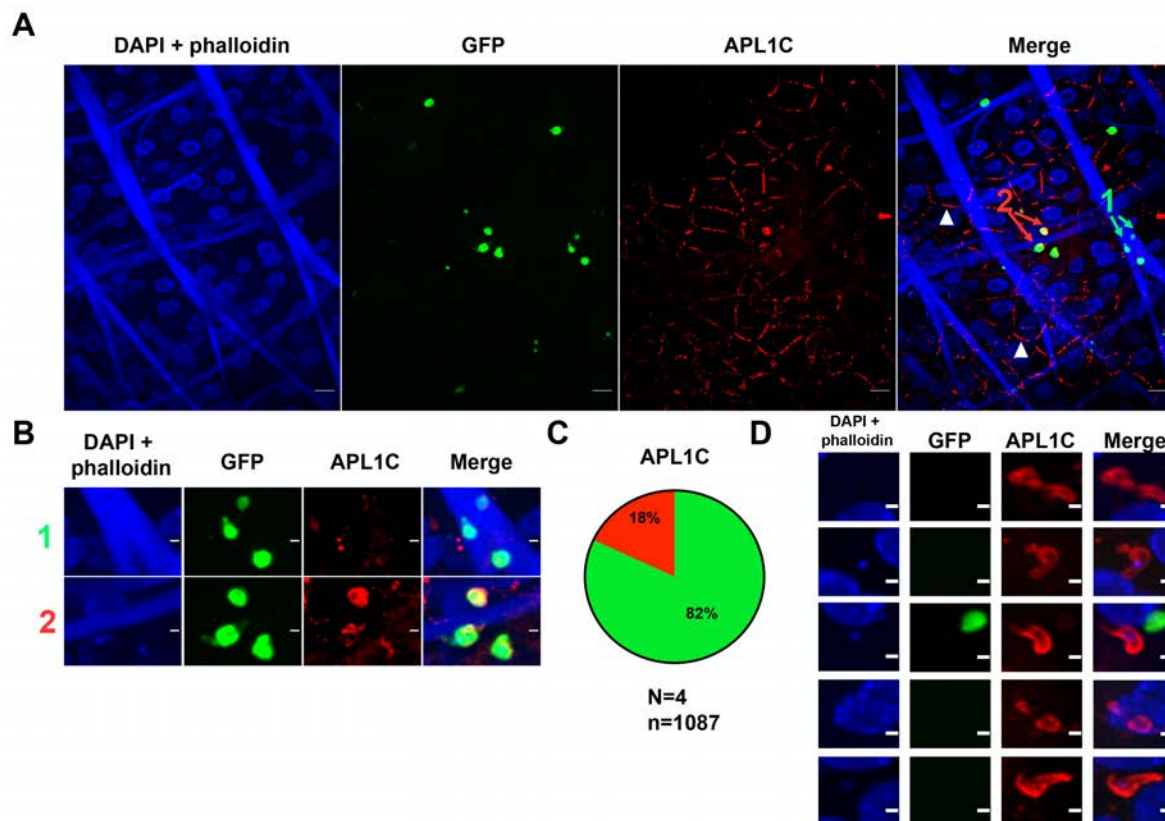
988 Therefore, I assessed whether APL1C binds to *Plasmodium* ookinetes. IFA analysis

989 on midguts collected 24 h post-IBM displayed APL1C signal on some parasites, but

990 the majority of parasites were negative for APL1C (**Figure 10 A and B "Merge"**)

## RESULTS

991 column). Overall, 18% of live (GFP) ookinetes present in the infected midguts were  
 992 positive for APL1C (**Figure 10C**). In addition, APL1C occasionally labelled structures  
 993 previously reported as stalk-form ookinetes<sup>124</sup>, but these structures were negative for  
 994 GFP (**Figure 10C**). It suggests that these GFP-negative forms were probably  
 995 undergoing lysis, and GFP protein leaked from dead parasites. This is consistent with  
 996 the description of mosquito complement leading to parasite lysis and previous  
 997 observations, in which parasite ookinetes labelled by TEP1 at 24 h post-infection were  
 998 found both, live (GFP-positive) and dead (GFP-negative)<sup>115 125</sup>.



999

1000 **Figure 10. APL1C binds to ookinetes in *A. coluzzii* midguts.** **A.** Immunostaining analysis  
 1001 of IBM mosquito midguts 24h post-infection shows GFP-expressing ookinetes (green)  
 1002 associated with APL1C protein (red arrows on the “Merge” picture) and APL1C-negative  
 1003 parasites (green arrow on the “Merge” picture). Scale bar is 10  $\mu$ m. **B.** Enlarged projections of  
 1004 APL1C-negative parasites (1) and APL1C-positive parasites (2) from (A). The scale bar is 2  
 1005  $\mu$ m. **C.** Depending on APL1C binding, GFP-positive parasite numbers (n=1087 number of the  
 1006 parasites scored across the replicates) were scored and are shown in the pie chart (red:  
 1007 APL1C-positive parasites, green: APL1C-negative parasites) as the mean percentage  
 1008 obtained from four independent experiments (N=4, 4 to 7 midguts per each replicate). **D.**  
 1009 APL1C labelling localises with dead ookinetes in mosquito midgut. Immunostaining analysis

## RESULTS

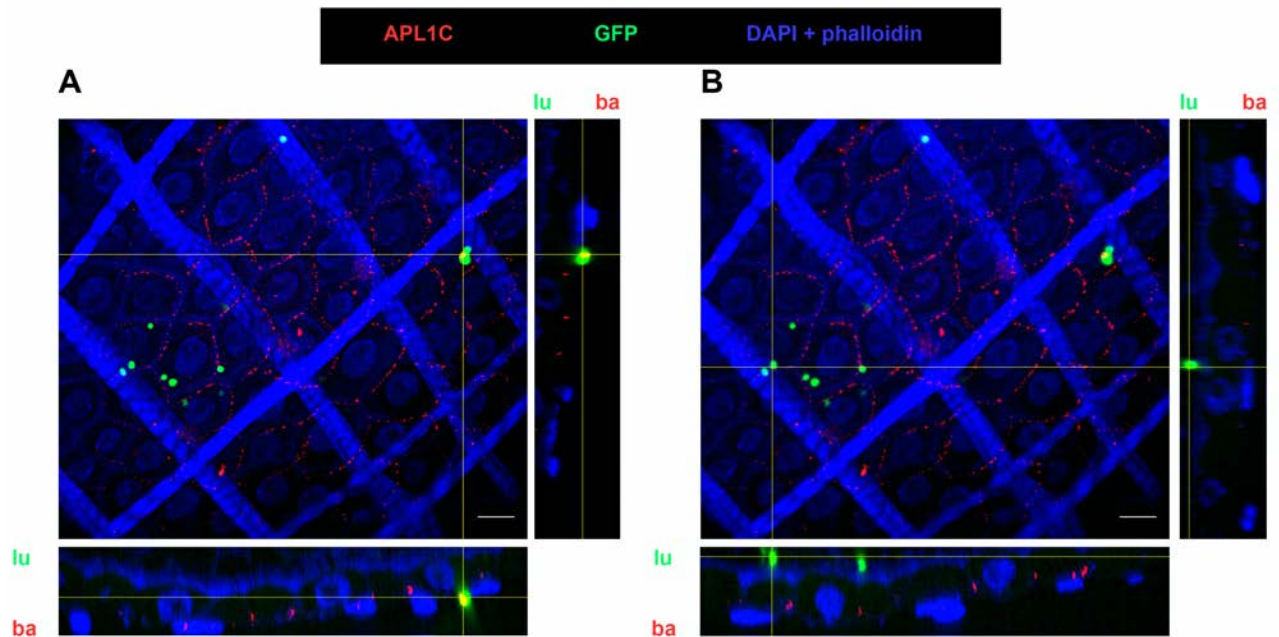
1010 of mosquito midgut infected with GFP-expressing parasites (green) at 24 h post-infection. The  
1011 APL1C staining reveals APL1C protein signal (red) with a shape related to ookinete structure  
1012 but where GFP signal from the parasite is lacking, probably due to parasite lysis. Note nucleus  
1013 staining in the DAPI + phalloidin channel. The scale bar is 2  $\mu$ m.

1014

1015 Ookinete traversal through the midgut epithelium is a dynamic process, and 24 h after  
1016 IBM the majority of parasites are still found in the blood bolus, while some are  
1017 intracellular and only a minority have exited from epithelial cells<sup>126</sup>. Despite  
1018 permeabilization of the midgut tissues during the IFA analysis, only 18% of GFP-  
1019 positive parasites were APL1C-labeled. An additional control was performed using  
1020 anti-GFP antibody which labelled 93% of parasites in the midguts, confirming efficient  
1021 permeabilization and antibody access within the tissue (Annex, **Figure S3**). Therefore,  
1022 I determined the localization (apical and/or basal) of APL1C-labelled parasites on the  
1023 midgut epithelium. Reconstructed optical sections of the confocal stack images  
1024 revealed that APL1C-labeled parasites were strictly located at the basal side of the  
1025 midgut epithelium (**Figure 11A**). In contrast, the ookinetes that were crossing or that  
1026 remained on the apical side of the midgut epithelium lumen were not labelled with  
1027 APL1C (**Figure 11B**). This result is consistent with the 18% APL1C-labeled parasites,  
1028 as it was reported that between 24 and 25 h post-infection, around 15% of them reach  
1029 the basal lamina<sup>126</sup>.

## RESULTS

1030



1031

1032 **Figure 11. APL1C binds to the ookinetes located on the basal side of the midgut.** XZ and  
1033 YZ orthogonal views of the stack pictures links APL1C protein binding with parasite spatial  
1034 localisation in mosquito midguts. Yellow lines depict the parasite, which spatial localisation is  
1035 presented on the sides of the stack picture. **A.** Illustration of a parasite traversing the midgut  
1036 epithelium that is labelled with APL1C on the basal side of the mosquito midgut. **B.** The  
1037 parasites that did not traverse the midgut epithelium and remained on the apical side of the  
1038 mosquito midgut are not labelled with APL1C. The midgut lumen (*lu*) and basal site (*ba*) of the  
1039 midgut are depicted on the orthogonal view pictures. The scale bar is 10  $\mu$ m.

1040

1041

### 1042 **3.1.4 LRR proteins bind autonomously to the ookinete surface (coll.:** 1043 **C. Lavazec, Institut Cochin, Paris)**

1044

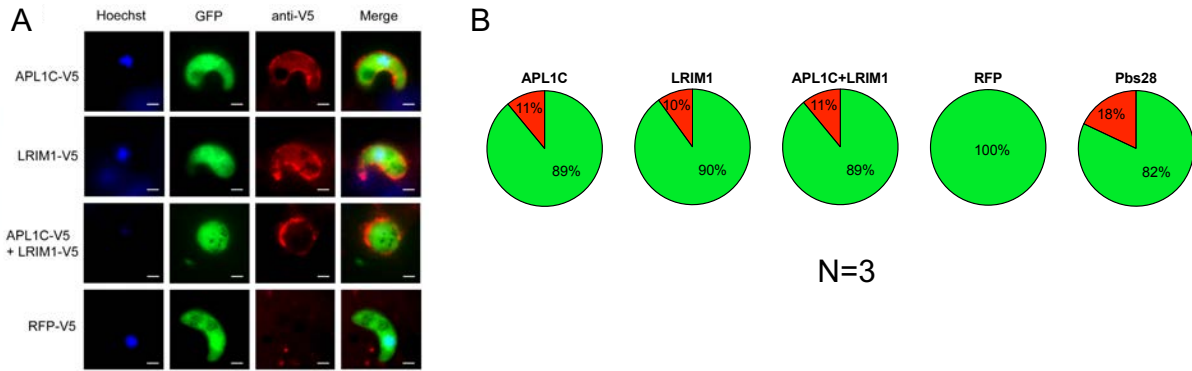
1045 As APL1C has been shown to act in a heterodimeric complex with LRIM1<sup>109</sup>, we  
1046 wished to determine whether LRIM1 could also be detected on the ookinete surface.  
1047 We developed a powerful *ex vivo* model of mosquito hemocoel immunity using the  
1048 4a3A hemocyte-like cell line. Cells were transfected to express, either separately or

## RESULTS

1049 simultaneously, V5-tagged LRIM1 (V5-LRIM1) or V5-APL1C (used as positive control  
1050 for the binding). After verifying that both V5-APL1C and V5-LRIM1 protein constructs  
1051 were secreted at day 3 post transfection from hemocyte-transfected cells in the  
1052 medium (Annex, **Figure S4**) we investigated their labelling on fluorescent-GFP  
1053 parasites. To validate the *ex vivo* system, we phenocopied APL1C binding to the  
1054 parasites on the midgut epithelium observed *in vivo*. Transfected cells were exposed  
1055 to *A. coluzzii* midguts infected with *P. berghei* (24 h post-IBM) and we found that V5-  
1056 APL1C as well as V5-LRIM1 were co-localized with parasites (**Figure 12A**). The V5-  
1057 RFP-transfected proteins, used as negative control, displayed no labelling on the GFP  
1058 parasites (**Figure 12A**), although as expected it was secreted into the medium (Annex,  
1059 **Figure S4**). In addition to validation of the *in vivo* system, these results indicated that  
1060 as for APL1C, the subunit partner LRIM1 also binds to the parasite, and in similar  
1061 proportions for V5-APL1C or V5-LRIM1-labelled parasites (12% and 10% respectively)  
1062 (**Figure 12B**). Interestingly, the proportion of labelled parasites did not increase in cells  
1063 co-transfected with both tagged-LRRs (11%) as compared to the single transfected  
1064 treatments (**Figure 12B**). This observation is consistent with but does not prove that  
1065 labelled parasites are bound with LRIM1/APL1C heterodimers.

1066 To distinguish extracellular ookinetes from those that were still within the lumen or  
1067 epithelial cells, we immunostained non-permeabilized infected midguts using *P.*  
1068 *berghei*-specific Pbs28 antibody. Only 18% of the GFP ookinetes were labelled with  
1069 the Pbs28 antibody (**Figure 12B**), indicating those that reached the basal side of the  
1070 midgut at 24 h post-IBM.

## RESULTS



1071

1072 **Figure 12. APL1C and LRIM1 both bind to *P. berghei*.**

1073 **A.** Immunostaining analysis of mosquito midguts infected with GFP parasites (green) after 2 h  
1074 incubation in culture medium from transfected 4a3A cells. Staining with an anti-V5 monoclonal  
1075 antibody (mAb) followed by fluorophore-conjugated IgG shows parasites associated with V5-  
1076 APL1C and V5-LRIM1 tagged proteins (red). Nuclei were stained with Hoechst 33342 (blue).  
1077 The scale bar is 2  $\mu$ m. **B.** V5-positive parasites and V5-negative parasites after staining with  
1078 anti-V5 antibodies in each condition were scored and illustrated in pie charts (red: V5-positive  
1079 parasites, green: V5-negative parasites) as the mean percentage obtained from three  
1080 independent experiments (N=3, 3–10 midguts per experimental point). The proportion of  
1081 parasites reaching the basal side of the midgut epithelium at 24 h post-infection was scored  
1082 using the anti-Pbs28 antibody (red: Pbs28-positive parasites, green: Pbs28-negative  
1083 parasites). *Data from collaboration: C. Lavazec, Institut Cochin, Paris.*

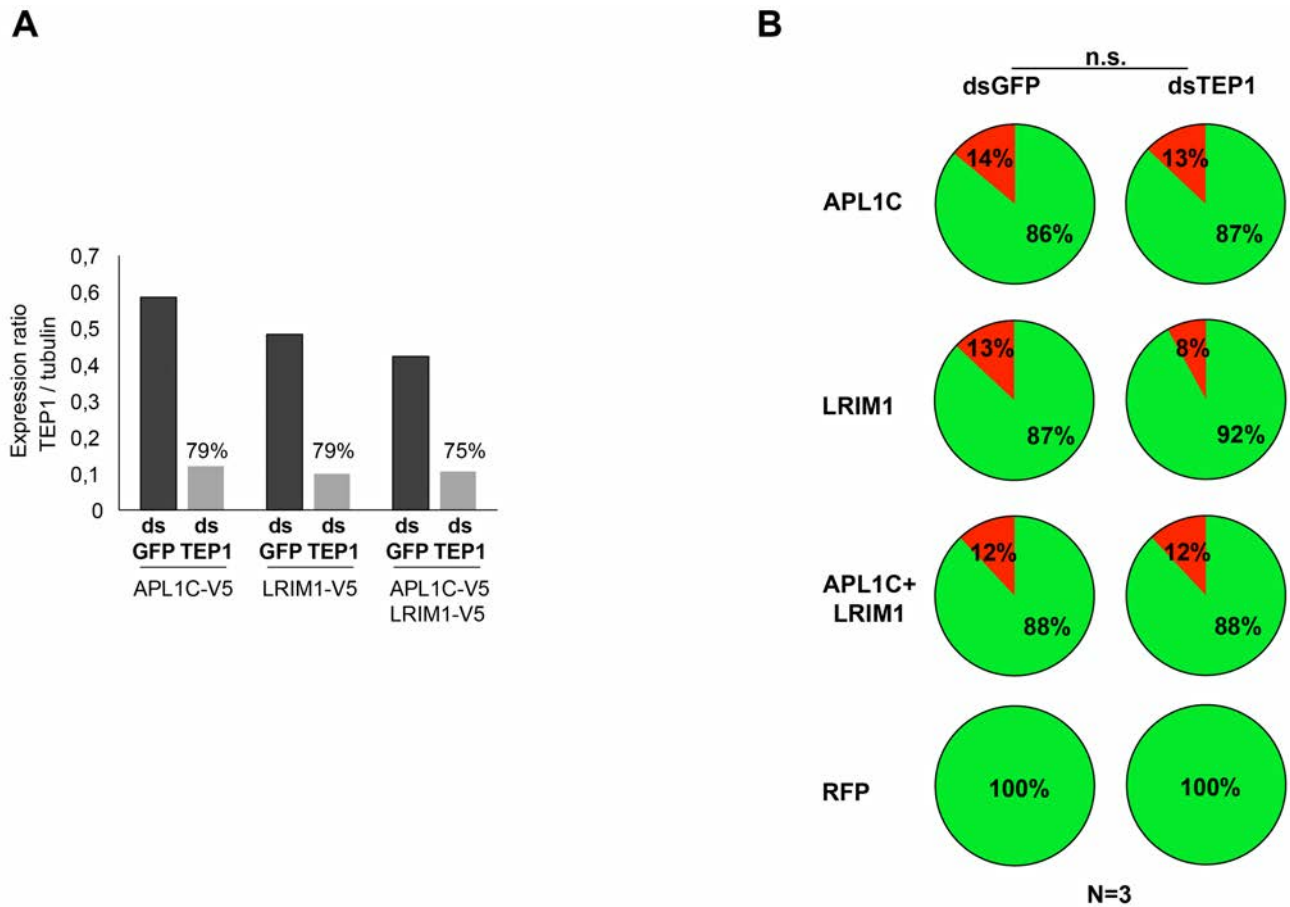
1084

1085 We then tested whether APL1C and LRIM1 parasite binding is dependent on the  
1086 presence of TEP1 protein. Cells were transfected to express V5-tagged APL1C and/or  
1087 V5-tagged LRIM1 as above, but with TEP1 gene expression silenced (dsTEP1 or  
1088 dsGFP as control). Depletion of TEP1 protein was confirmed by Western blot using an  
1089 anti-TEP1 antibody (**Figure 13A**). *A. coluzzii* midguts collected 24 h post-infection  
1090 were incubated in medium from transfected cells. Depletion of TEP1 did not alter the  
1091 proportion of V5-LRR-labelled parasites as compared to the dsGFP controls, whether  
1092 APL1C and LRIM1 were expressed individually or simultaneously (**Figure 13B**). These  
1093 results indicate that while TEP1 requires at least APL1C or LRIM1 for specific parasite  
1094 binding, the reciprocal is not true, suggesting that TEP1 binding on the parasite must  
1095 occur indirectly and secondarily, mediated by LRR binding to the parasite surface.



## RESULTS

1096



1097

1098 **Figure 13. APL1C and LRIM1 both bind to *P. berghei* ookinetes independently of TEP1**  
 1099 **presence.**

1100 **A.** TEP1 protein is successfully depleted by dsTEP1 in 4a3A cells. Expression ratios of protein  
 1101 levels TEP1/tubulin were quantified by densitometry. The efficiency of the gene silencing effect  
 1102 was monitored at day 6 by immunoblotting analysis of cells and culture medium of the 4a3A  
 1103 cell line using an anti-TEP1 antibody and an anti-tubulin antibody. For each condition, the ratio  
 1104 of protein levels TEP1/tubulin was calculated and the percentage of TEP1 signal reduction  
 1105 relative to tubulin was quantified in dsTEP1-transfected cells as compared to dsGFP  
 1106 transfected cells. **B.** APL1C and LRIM1 binding to the parasites are TEP1 independent.  
 1107 Proportions of parasites associated with V5-tagged proteins in cells depleted for TEP1, after  
 1108 staining with anti-V5 antibodies. Parasite numbers in each condition were scored and are  
 1109 shown in the pie chart (red: V5-positive parasites, green: V5-negative parasites) as the mean  
 1110 percentage obtained from three independent experiments (3–10 midguts per experimental  
 1111 point). DsGFP transfected cells were used as a control. Numbers of V5-positive and V5-  
 1112 negative parasites from each replicate were compared between the two conditions by chi-  
 1113 square test (significance level of chi-square: n.s.=not significant). *Data from collaboration: C.*  
 1114 *Lavazec, Institut Cochin, Paris.*

1115

## RESULTS

### 1116 **3.1.5 Phagocytic hemocytes are required for wildtype APL1C levels** 1117 **in the hemolymph**

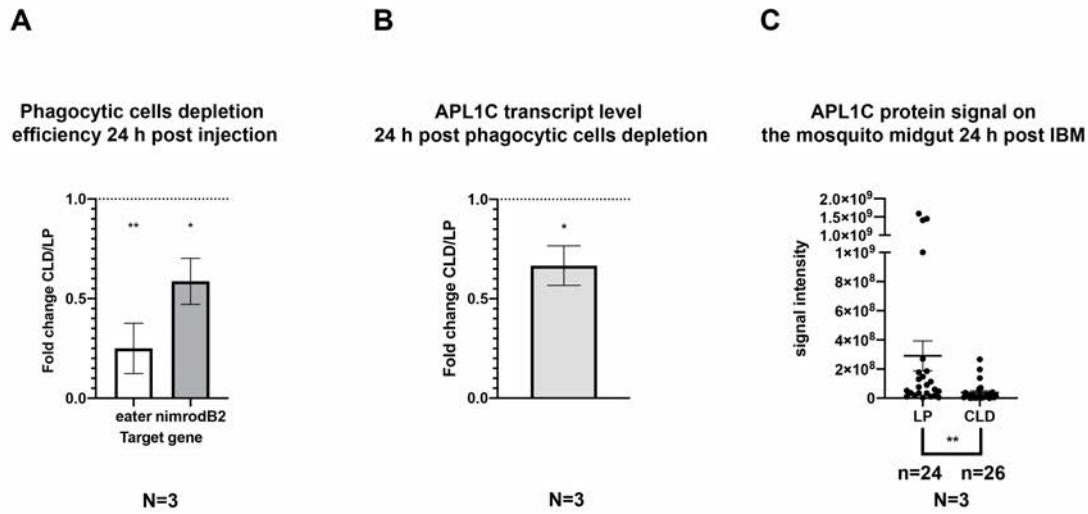
1118

1119 Recent results demonstrated the specific contribution of the phagocytic hemocyte cells  
1120 for the binding of invading ookinetes by the mosquito complement system<sup>127</sup>.  
1121 Therefore, I investigated the contribution of the phagocytic cells for APL1C protein  
1122 abundance in the hemolymph, measured as above using IFA to detect APL1C levels  
1123 present on dissected midguts. Phagocytic cells were chemically depleted by injecting  
1124 liposome-encapsulated clodronate (CLD) as described<sup>127</sup> prior to an IBM, with empty  
1125 liposomes (LP) as a control. The liposome capsule prevents systemic toxicity of CLD  
1126 to the mosquito, while phagocytosis of the encapsulated CLD and subsequent release  
1127 of CLD specifically leads to the death of the phagocytic hemocytes. The efficiency of  
1128 phagocyte depletion was measured by quantifying the expression of two phagocyte  
1129 markers 24 h post treatment, eater and nimrodB2, in pools of eight whole mosquitoes.  
1130 Expression of these two markers was significantly reduced by CLD treatment as  
1131 compared to LP controls (**Figure 14A**).

1132 Depletion of phagocytic hemocytes by CLD treatment decreased APL1C gene  
1133 expression (**Figure 14B**). Depletion of phagocytic hemocytes by CLD treatment also  
1134 reduced the abundance of APL1C protein in the hemolymph, as measured by IFA  
1135 signal intensity on dissected midguts from CLD and LP treated mosquitoes at 24 h  
1136 post-IBM ( $p < 0.01$  **Figure 14B** and Annex **Table S1**). These results indicate that  
1137 phagocytic hemocytes are required both for APL1C gene expression as well as normal  
1138 APL1C protein levels in the mosquito hemolymph.



## RESULTS



1139

### 1140 **Figure 14. Full APL1C gene expression and protein abundance in the hemolymph** 1141 **requires phagocytic hemocytes.**

1142 **A.** Phagocytic hemocytes were depleted by clodronate treatment (CLD). CLD-mediated  
 1143 phagocyte depletion was verified by qPCR measurement of phagocyte markers, eater and  
 1144 nimrodB2 between CLD-treated mosquitoes and liposome controls (LP, indicated as dotted  
 1145 line), 24 h post-treatment prior to IBM. The ratios of normalized eater, nimrodB2 transcript  
 1146 detection were computed using triplicates from the same cDNA dilution, and graph represents  
 1147 mean with  $\pm$ SEM of the expression fold change from three biological replicates (N=3). qPCR  
 1148 was analysed by unpaired t-test (\* p-value<0.05; \*\* p-value<0.01). **B.** APL1C expression was  
 1149 reduced in CLD-treated mosquitoes. Analysis of APL1C expression was performed as  
 1150 described in **A.** **C.** Hemolymph APL1C protein abundance measured as IFA signal intensity on  
 1151 midguts dissected from mosquitoes treated with CLD prior to IBM. Data points on graph  
 1152 indicate single midguts, bars indicate mean with  $\pm$ SEM. The number of independent replicate  
 1153 experiments (N) the total number of midguts dissected across replicates (n) are indicated. Data  
 1154 were analysed by measuring APL1C signal intensity (RawIntDen) using ImageJ v1.52p and  
 1155 the intensity from each midgut was compared between the two conditions by Mann-Whitney  
 1156 test. All statistical differences were first tested independently within replicates (individual p-  
 1157 values in Annex **Table S1**), and if individual replicates showed a common trend of change,  
 1158 individual p-values were combined using the meta-analytical approach of Fisher (significance  
 1159 level of Fisher-combined p-value: \*\* p-value <0.01).

1160

### 1161 **3.1.6 Nitration pathway activity is required for full APL1C** 1162 **abundance in hemolymph**

1163

## RESULTS

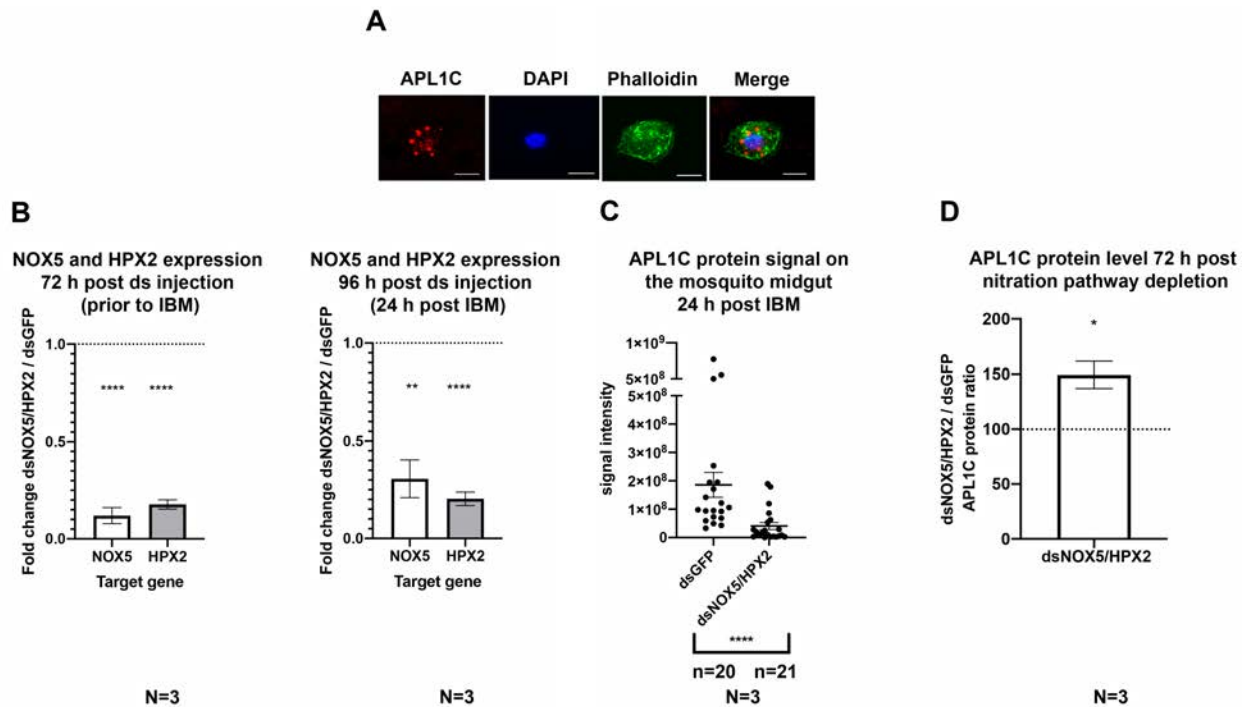
1164 The release of microvesicles by mosquito hemocytes appears to be correlated with the  
1165 deposition of TEP1 upon the ookinete surface, because depletion of the nitration  
1166 pathway simultaneously abolished release of hemocyte-derived microvesicles (HdMv)  
1167 and impaired TEP1 binding to the parasites<sup>128</sup>. This observation led to the hypothesis  
1168 that factors delivered by HdMv may promote TEP1 binding to ookinetes, although the  
1169 contents of the HdMv are unknown. APL1C is expressed in hemocytes and its  
1170 presence is required for TEP1 binding to parasites. Therefore, I evaluated whether  
1171 APL1C could be a component of the HdMv payload to deliver APL1C into the  
1172 hemolymph and ultimately to the midgut surface.

1173 First, APL1C subcellular localization was determined by IFA with APL1C antibody on  
1174 perfused hemocytes and on 4a3A hemocyte-like cells, revealing that APL1C protein is  
1175 found in vesicle-like structures in perfused hemocytes (**Figure 15A**). This observation  
1176 is consistent with our previous report of localization of APL1 family gene APL1A within  
1177 hemocyte vesicles<sup>106</sup> and with the hypothesis that APL1C may be secreted from  
1178 hemocytes in microvesicles. Next, the nitration pathway was inactivated by co-  
1179 silencing the genes for heme peroxidase (HPX2) and NADPH oxidase 5 (NOX5), with  
1180 the aim to inhibit the release of HdMv. Mosquitoes were treated with dsNOX5/HPX2  
1181 or control dsGFP prior to an IBM, and the efficiency of HPX2 and NOX5 depletion was  
1182 confirmed by RT-qPCR. (**Figure 15B**).

1183 Midguts from each group were collected 24 h after *Plasmodium* infection and APL1C  
1184 abundance on the midgut was quantified using IFA with APL1C antibody. Midguts from  
1185 mosquitoes depleted for NOX5/HPX2 displayed decreased APL1C abundance as  
1186 compared to dsGFP treated controls ( $p < 0.0001$ ) (**Figure 15C** and Annex **Table S1**).  
1187 Western blot analysis of extract from whole mosquitoes confirmed that the observed  
1188 phenotype did not result from APL1C protein decrease after inhibition of the nitration

## RESULTS

1189 pathway, and in fact APL1C protein was slightly increased in silenced mosquitoes as  
 1190 compared to dsGFP controls (**Figure 15D**), which could be explained if APL1C protein  
 1191 is retained in non-released HdMv. Overall, these results indicate that an active nitration  
 1192 pathway is necessary for wild-type levels and localization of free APL1C in the  
 1193 hemolymph, as measured by APL1C abundance captured on the midgut.



1194

1195 **Figure 15. Nitration pathway activity mediates APL1C protein presence in mosquito**  
 1196 **hemolymph.** **A.** Immunostaining analysis of perfused hemocytes from mosquitoes indicates  
 1197 that APL1C is localized in subcellular vesicles (red). Cells were stained with Hoechst 33342 to  
 1198 label nuclei (blue) and phalloidin for actin (green). Scale bar, 5  $\mu$ m. *Data from collaboration:*  
 1199 *C. Lavazec, Institut Cochin, Paris.* **B.** Expression of NOX5 and HPX2 genes is decreased after  
 1200 treatment with dsNOX5/HPX2. NOX5 and HPX2 silencing was verified by qPCR measurement  
 1201 between dsNOX5/HPX2 and dsGFP (dotted line) 72 h post-treatment (before IBM). The ratio  
 1202 of normalized NOX5 or HPX2 cDNA detection in “dsNOX5/HPX2” versus “dsGFP” was  
 1203 computed using triplicates from the same cDNA dilution. Graph represents mean with  $\pm$ SEM  
 1204 of the expression fold change between “dsNOX5/HPX2” and “dsGFP” control from three  
 1205 biological replicates (N=3). Data for qPCR analysis was analysed by unpaired t-test  
 1206 (significance levels of t-test p-values: \*\* p-value <0.01; \*\*\*\* p-value <0.0001). **B.** NOX5 and  
 1207 HPX2 silencing was verified accordingly to **(A)** 96 h post treatment (24 h after IBM). **C.** Nitration  
 1208 pathway elicits APL1C presence in the hemolymph, measured by detection of protein captured  
 1209 in the extracellular space of dissected midguts. Quantitative analysis of the APL1C signal on  
 1210 the midguts from the mosquitoes in which the nitration pathway was depleted as compared to  
 1211 controls reveals decreased APL1C protein in mosquitoes depleted for nitration activity. Graph  
 1212 labels and statistical tests for this figure: tests of APL1C protein signal intensity on the midgut

## RESULTS

1213 indicate each midgut subjected for the analysis as a dot, bars represent mean with  $\pm$ SEM.  
1214 Sample sizes (N) show the number of independent replicate experiments, (n) the total number  
1215 of midguts dissected across replicates. Data were analysed by measuring APL1C signal  
1216 intensity (RawIntDen) using ImageJ v1.52p and the intensity from each midgut was compared  
1217 between the two conditions by Mann-Whitney test. All statistical differences were first tested  
1218 independently within replicates (individual p-values in Annex **Table S1**), and if individual  
1219 replicates showed a common trend of change, individual p-values were combined using the  
1220 meta-analytical approach of Fisher (significance level of Fisher-combined p-value: \*\*\*\* p-value  
1221 <0.0001). **D.** The APL1C protein level in whole mosquitoes increases upon nitration pathway  
1222 depletion. Immunoblot analysis of APL1C protein level 72 h after dsNOX5/HPX2 treatment.  
1223 Protein extracts from mosquitoes injected with dsNOX5/HPX2 and dsGFP (control) were  
1224 subjected to a Western blot with an anti-APL1C pAb and anti-GAPDH antibody as a loading  
1225 control. The normalized APL1C signal density was compared between “dsNOX5/HPX2” and  
1226 “dsGFP” mosquitoes (dotted line). Graph represents mean with  $\pm$ SEM of the protein ratio  
1227 between “dsNOX5/HPX2” and “dsGFP” from three biological replicates (N=3). Data were  
1228 analysed by Image Lab 6.0.1 and statistically analysed by unpaired t-test (significance level of  
1229 t-test \* p-value<0.05).

1230

### 1231 **3.2 APL1C protein is a pathogen binding factor for the**

### 1232 **hemocoel sporozoite stage of *Plasmodium***

1233

#### 1234 **3.2.1 Sporozoites affect APL1C protein abundance**

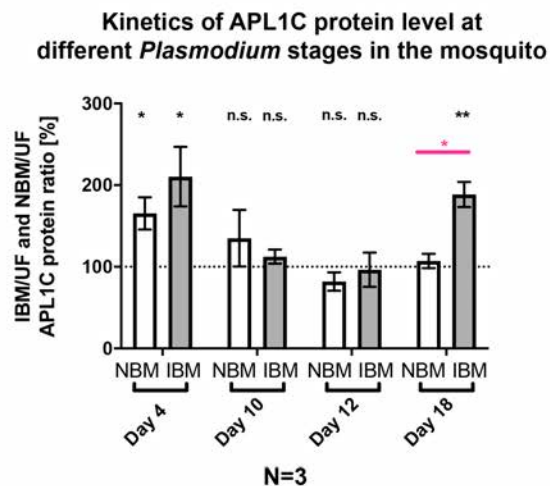
1235

1236 Only a small proportion of the thousands of sporozoites released by each oocyst  
1237 invade the salivary glands, while the majority of sporozoites are destroyed in the  
1238 mosquito hemocoel by an unknown mechanism<sup>46</sup>. Phagocytosis of sporozoites by  
1239 hemocytes was observed, but was numerically insufficient in itself to explain sporozoite  
1240 clearance from the hemocoel. Thus, the mechanism of sporozoite destruction in the  
1241 hemocoel remains unknown.

1242 To examine the activity of APL1C against released sporozoites in the hemocoel, the  
1243 abundance of APL1C protein was first measured by Western blot at four time points  
1244 during *Plasmodium* development in the vector: young oocysts at 4 d post-bloodmeal,  
1245 late oocysts before sporozoite release at 10 d post-bloodmeal, early sporozoite release

## RESULTS

1246 at 12 d post-bloodmeal, and late sporozoite release at 18 d post-bloodmeal. Unfed  
1247 mosquitoes (UF) were used as controls. At 4 d post-bloodmeal, APL1C protein was  
1248 increased in both IBM and NBM mosquitoes as compared to UF, but there was no  
1249 difference between IBM and NBM, indicating that the early stage induction of APL1C  
1250 protein was dependent only on the bloodmeal and not infection (**Figure 16**). At 10 d  
1251 and 12 d post-bloodmeal, APL1C abundance in both IBM and NBM mosquitoes was  
1252 equivalent to the UF controls, indicating that APL1C levels return to baseline after the  
1253 effect of the bloodmeal decays. Notably, at 18 d post-bloodmeal during sporozoite  
1254 release, APL1C protein abundance was significantly higher in IBM as compared to  
1255 either NBM or UF mosquitoes, which were not different from each other (**Figure 16**).  
1256 The APL1C increase at 18 d post-IBM indicates that APL1C induction at this time point  
1257 was dependent only on the presence of circulating sporozoites. This result suggests  
1258 that APL1C is a component of immune signalling responding to free sporozoites, and  
1259 could help to limit salivary gland invasion and thus transmission of *Plasmodium*.



1260

1261 **Figure 16. APL1C protein level is increased by blood feeding and sporozoite release in**  
1262 ***A. coluzzii*.** Immunoblot analysis of APL1C protein levels after IBM. Protein extract from whole  
1263 mosquitoes collected at day 4 (early oocyst), day 10 (mature oocyst, before sporozoite  
1264 release), day 12 (early sporozoite release) and day 18 (late sporozoite release) were probed  
1265 with an anti-APL1C pAb and anti-GAPDH antibody as a loading control. NBM and UF samples

## RESULTS

1266 were used as controls for each time point. The normalized APL1C signal was compared  
1267 between NBM or IBM mosquitoes versus UF mosquitoes (dotted line). APL1C protein level  
1268 increases at day 4 after NBM or IBM and at day 18 after IBM as compared to controls. Graph  
1269 represents mean with  $\pm$ SEM of the protein ratio between "IBM" or "NBM" and "UF" control from  
1270 three biological replicates (N=3). Data were analysed by Image Lab 6.0.1 and statistically  
1271 analysed by unpaired t-test (significance levels of t-test p-values: n.s., not significant; \* p-  
1272 value<0.05; \*\* p-value <0.01).

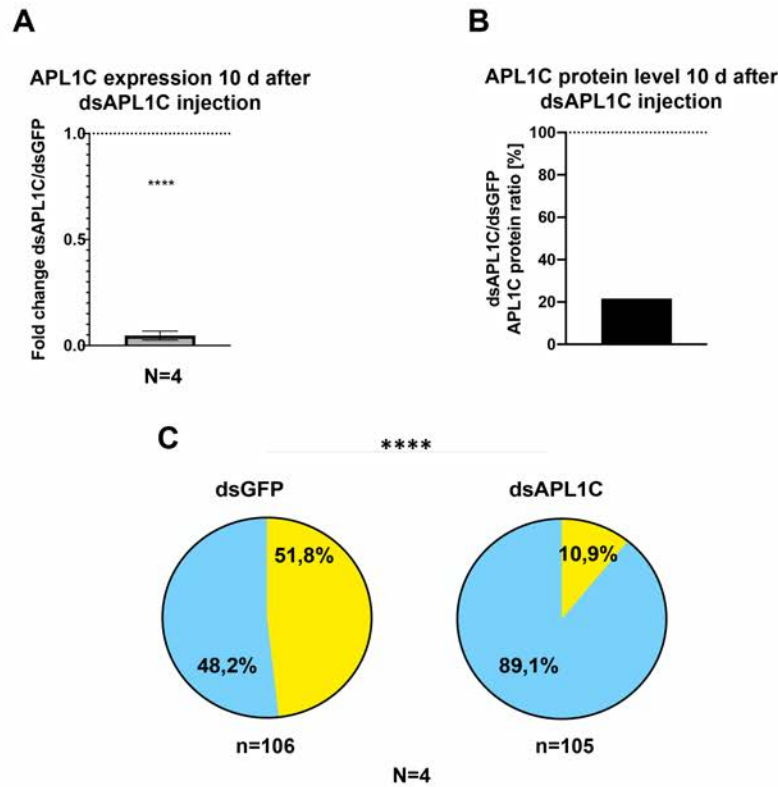
1273

### 1274 **3.2.2 APL1C activity in the hemocoel limits sporozoite invasion of** 1275 **salivary glands**

1276

1277 Based on the APL1C protein increase in mosquitoes at the time of high sporozoite  
1278 release at 18 d post-IBM, APL1C was tested for a protective role against sporozoites.  
1279 Mosquitoes were treated with dsAPL1C or control dsGFP at 9 d post-IBM,  
1280 approximately 3 d before the beginning of sporozoite release (experimental design,  
1281 **Figure S5**). Silencing efficiency was confirmed for APL1C transcript and protein, and  
1282 was persistent until at least 10 d post-treatment (**Figure 17A and B**). Salivary glands  
1283 were dissected at 19 d post-IBM and sporozoites were counted using fluorescence  
1284 microscopy for each individual mosquito. APL1C-depleted mosquitoes displayed  
1285 increased sporozoite infection prevalence, that is, the proportion of positive versus  
1286 negative glands, as compared to dsGFP controls ( $p<0.0001$ ) (**Figure 17C**, Annex,  
1287 **Table S2**). There was no effect on infection intensity, that is, the numbers of  
1288 sporozoites in positive glands (Annex, **Table S2**). Thus, APL1C activity controls the  
1289 efficiency of sporozoite invasion of salivary glands, and therefore should influence  
1290 vector competence for parasite transmission.

## RESULTS



1291

1292 **Figure 17. APL1C activity limits sporozoite invasion of salivary glands.** **A.** Silencing of  
 1293 APL1C gene expression is efficient until at least 10 d after dsAPL1C treatment, the time point  
 1294 when salivary glands were dissected. APL1C silencing was verified by qPCR measurement  
 1295 between dsAPL1C and dsGFP (dotted line) treatments. The ratio of the normalized APL1C  
 1296 cDNA detection in “dsAPL1C” versus “dsGFP” treatments was computed using triplicates from  
 1297 the same cDNA dilution. Graph represents mean with  $\pm$ SEM of the expression fold change  
 1298 between “dsAPL1C” and “dsGFP” control from four biological replicates (N=4). Data for qPCR  
 1299 analysis was analysed by unpaired t-test (significance level of t-test \*\*\*\* p-value <0.0001). **B.**  
 1300 APL1C protein level is efficiently decreased 10 d after dsAPL1C treatment. Graph represents  
 1301 the protein ratio between “dsAPL1C” and “dsGFP” control (depicted as a dotted line),  
 1302 measured by densitometry. GAPDH density signal was used to normalize APL1C signal in  
 1303 both dsGFP and dsAPL1C treatments, and the normalized APL1C signals were then  
 1304 compared between these two treatments. **C.** APL1C depletion led to increased sporozoite  
 1305 infection prevalence of the salivary glands. The proportions of the sporozoite-infected salivary  
 1306 glands (blue) or non-infected (yellow) in dsAPL1C and dsGFP treated mosquitoes were scored  
 1307 (n=total number of salivary glands dissected across the replicates) and are illustrated in the  
 1308 pie chart as the mean percentage obtained from four independent experiments (N=4).  
 1309 Prevalence from each replicate was compared between the two conditions by chi-square test.  
 1310 All statistical differences were first tested independently within replicates (individual p-values  
 1311 in Annex **Table S2**), and if individual replicates showed a common trend of change, individual  
 1312 p-values were combined using the meta-analytical approach of Fisher (significance level of  
 1313 chi-square \*\*\*\* p-value <0.0001).

1314

1315

## RESULTS

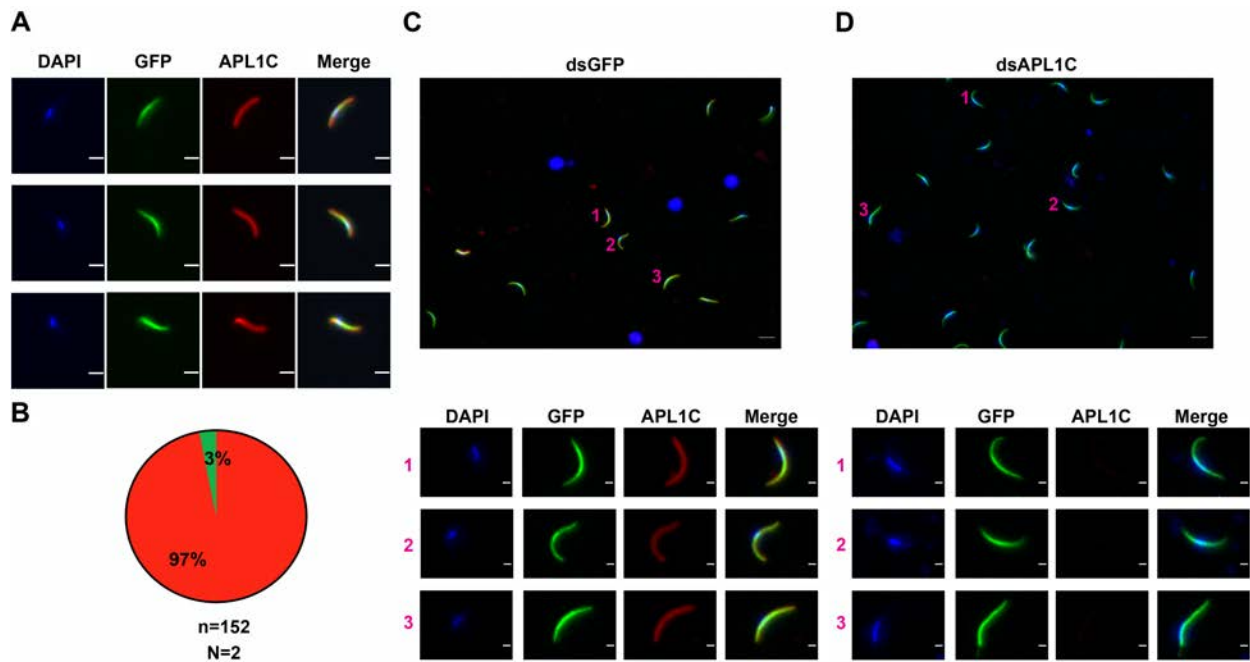
### 1316 **3.2.3 APL1C binds to free *Plasmodium* sporozoites in the mosquito** 1317 **hemolymph**

1318

1319 The simplest model to explain APL1C inhibition of sporozoite invasion into the salivary  
1320 glands is similar to the above proposed model for ookinete immunity, where LRR  
1321 proteins recognize and bind to the pathogen, then recruiting the TEP1 effector for lysis  
1322 of the cell. Therefore, we tested whether APL1C binds directly to sporozoites.  
1323 Sporozoites were perfused from IBM mosquitoes at 17 d post-IBM and IFA was  
1324 performed using APL1C antibody. The majority (97%) of sporozoites were labelled with  
1325 APL1C antibody (**Figure 18A and B**), indicating that APL1C protein binds to  
1326 sporozoites in the mosquito hemocoel. As a control, 2 of 210 sporozoites perfused  
1327 from APL1C-silenced mosquitoes were detectably labelled with APL1C antibody  
1328 (**Figure 18C and D**). These results demonstrate that APL1C protein binds to free  
1329 sporozoites in the hemolymph, which is consistent with its protective activity against  
1330 salivary gland infection shown above.



## RESULTS



1331

1332 **Figure 18. APL1C protein binds to free sporozoites in mosquito hemolymph.**  
1333 **A.** Immunostained GFP sporozoites (green) from mosquitoes perfused 17 days post-IBM are  
1334 labelled with APL1C protein (red). Nuclei were stained with DAPI (blue). Images are  
1335 representative of two independent biological replicates. The scale bar is 5  $\mu$ m. **B.** The  
1336 proportion of circulating sporozoites (n=152) labelled with APL1C protein was scored and  
1337 illustrated in a pie chart as a mean percentage from two independent experiments (red:  
1338 APL1C-positive sporozoites, green: APL1C-negative sporozoites). **C.** Sporozoites perfused  
1339 from dsGFP-injected mosquitoes are labelled by APL1C protein. The main picture depicts  
1340 merge (scale bar 10  $\mu$ m), whereas pictures below show 3 channels of enlarged projections of  
1341 sporozoites (scale bar 2  $\mu$ m), indicated by the numbers on the merge picture. **D.** Sporozoites  
1342 perfused from dsAPL1C-treated controls are not labelled by APL1C protein. The main picture  
1343 depicts merge (scale bar 10  $\mu$ m), whereas pictures below 3 channels of enlarged projections  
1344 of sporozoites (scale bar 2  $\mu$ m), indicated by the numbers on the merge picture.

1345

1346

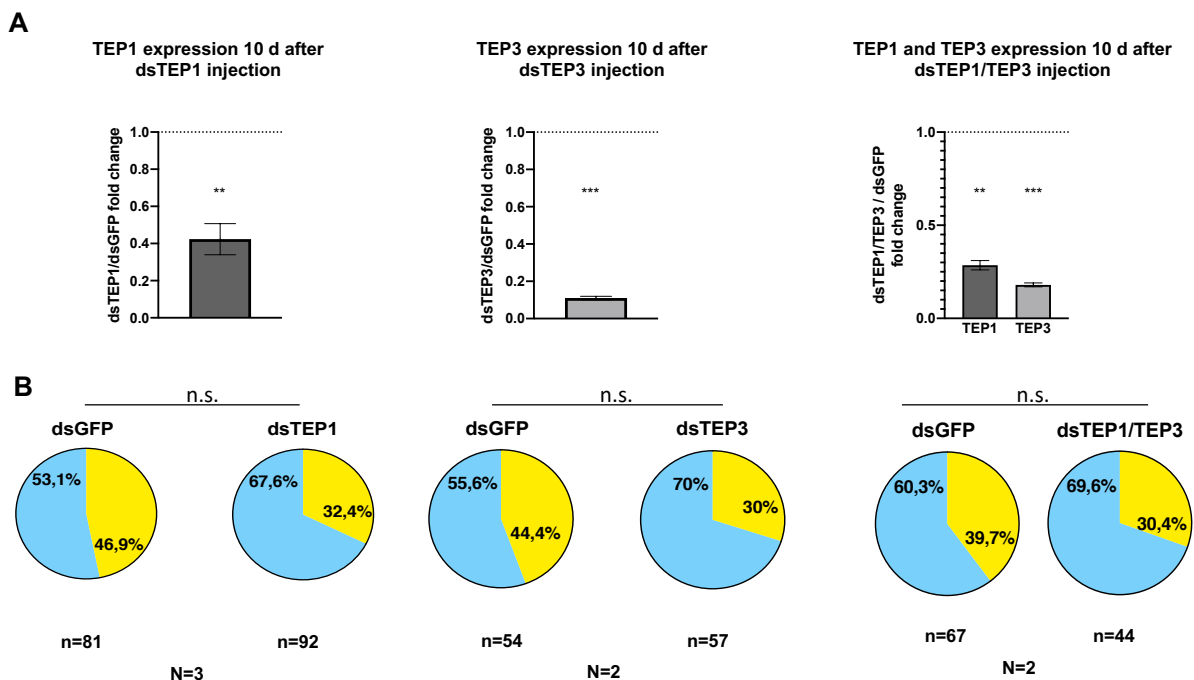
### 1347 **3.2.4 APL1C protective activity against sporozoites does not require** 1348 **the complement proteins TEP1 or TEP3**

1349

1350 For killing of rodent *Plasmodium* midgut ookinetes, it was previously shown that a  
1351 physical complex of the APL1C-LRIM1 heterodimer with either one of the thioester  
1352 proteins TEP1 or TEP3 is required<sup>104,111</sup>. Therefore, the role of TEP1 and TEP3 in  
1353 APL1C-mediated sporozoite protection was queried. Expression of TEP1 and TEP3

## RESULTS

1354 was silenced individually or simultaneously by treatment with dsRNA at 9 d post-IBM,  
 1355 and sporozoite infection prevalence and intensity in the salivary glands was measured  
 1356 at 19 d post-IBM (experimental design, **Figure S5**). Silencing efficiency was confirmed  
 1357 at 10 d after dsRNA treatment (**Figure 19A**). Neither TEP gene, alone or together,  
 1358 displayed an influence on sporozoite infection of salivary glands (**Figure 19B**). These  
 1359 results indicate that while TEP1 and TEP3 both display anti-ookinete function in the  
 1360 mosquito midguts<sup>104,111,129 115</sup>, they do not contribute to APL1C-mediated anti-  
 1361 sporozoite function. APL1C may recruit different TEP family members for sporozoite  
 1362 protection, or sporozoite destruction could be mediated by a distinct mechanism that  
 1363 does not include a TEP effector. Further work will be required to distinguish between  
 1364 these models.



1365

1366 **Figure 19. APL1C protective activity against sporozoites does not require the**  
 1367 **complement proteins TEP1 or TEP3. A.** TEP1 and TEP3 expression are efficiently silenced  
 1368 at the time of the salivary gland dissection (10 days post-dsTEP1, dsTEP3 or dsTEP1/TEP3  
 1369 treatments). TEP1 and TEP3 silencing was verified by qPCR. The ratio of normalized TEP1 or  
 1370 TEP3 cDNA detection in “dsTEP1”, “dsTEP3” or “dsTEP1/TEP3” versus “dsGFP” treatments  
 1371 was computed using triplicates from the same cDNA dilution. Graph represents mean with  
 1372  $\pm$ SEM of the expression fold change between “dsTEP1”, “dsTEP3” or “dsTEP1/TEP3” as

## RESULTS

1373 compared to the “dsGFP” control from all biological replicates (N) depicted in **B**. Data for qPCR  
1374 analysis was analysed by unpaired t-test (significance levels of t-test p-values: \*\* p-value  
1375 <0.01; \*\*\* p-value <0.001). **B**. TEP1, TEP3 or simultaneous TEP1/TEP3 depletion does not  
1376 affect sporozoite salivary gland infection. The salivary glands dissected among all the  
1377 replicates were scored (n=total number of salivary glands dissected in all replicates) and  
1378 represented in pie charts as the mean percentage of sporozoite-infected salivary glands (blue)  
1379 or non-infected (yellow) obtained from indicated number of biological replicates (N). The  
1380 percentages of dsTEP1, dsTEP3 or dsTEP1/TEP3 salivary gland infection prevalence were  
1381 compared with dsGFP injected control. Infection prevalence from each replicate was  
1382 compared between the two conditions by chi-square test. All statistical differences were first  
1383 tested independently within replicates (individual p-values in Annex **Table S2**), and if individual  
1384 replicates showed a common trend of change, individual p-values were combined using the  
1385 meta-analytical approach of Fisher (significance level of chi-square: n.s.=not significant).

1386

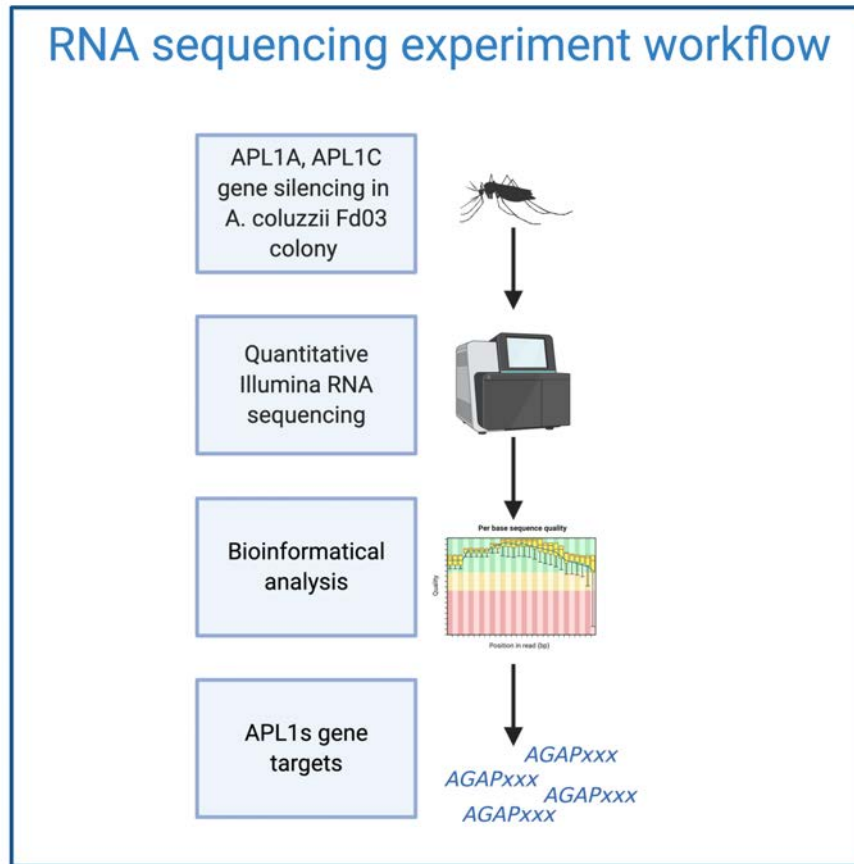
1387

### 1388 **3.3 Implication of the APL1 LRR family in immune signal**

#### 1389 **transduction**

1390 APL1C binds to both ookinete and sporozoite parasite stages, and appears to be a  
1391 component of pathogen sensing for free sporozoite presence in the hemocoel. APL1C  
1392 may generate signals leading to the expression of additional downstream immune  
1393 factors. To test the hypothesis that APL1 family proteins lead to downstream immune  
1394 signalling, a transcriptomic analysis using RNA sequencing (RNAseq) was carried out  
1395 to detect the transcriptional effects of APL1A and APL1C activity (**Figure 20**). Because  
1396 the phenotypes of each gene are largely mediated by different immune pathways<sup>103,102</sup>,  
1397 it was expected that the transcriptional effects of each gene would be different.

## RESULTS



1398

1399 **Figure 20. Overview of the RNAseq experiment.** The APL1A or APL1C genes were silenced  
1400 in *A. coluzzii*, with dsGFP treatment as the control, and RNAs were subjected to RNAseq.  
1401 Bioinformatic analysis identified differentially expressed genes for each silenced gene as  
1402 compared to the control.

1403

### 1404 **3.3.1 APL1A and APL1C gene silencing in *A. coluzzii***

1405

1406 The APL1 genes are characterized by their exceptional genetic diversity observed in  
1407 wild mosquitoes as well as in the lab-reared colonies<sup>105,106</sup>. This genetic variability  
1408 within each gene include gene insertion-deletions<sup>106</sup>, and variants in the region  
1409 targeted for gene silencing using dsRNA could affect silencing efficiency. Therefore, I  
1410 identified the potential APL1A and APL1C variants segregating in the studied mosquito  
1411 colony (Fd03) to verify that they are sensitive to the designed dsAPL1A and dsAPL1C  
1412 fragments targeting the gene 3'UTRs<sup>102,103</sup>. 3' Rapid Amplification of cDNA Ends (3'

## RESULTS

1413 RACE) was used to sequence APL1A and APL1C 3'UTR variation. Five different  
1414 APL1A 3' variants were identified, and all of them overlapped and would be sensitive  
1415 to the designed dsAPL1A fragment (**Figure 21A**). Among 22 sequenced clones, 18  
1416 represented one major structural variant with 8 bp and 9 bp insertions compared to the  
1417 reference sequence, resulting in a shifted stop codon (stop codon depicted as 'TAG'  
1418 on **Figure 21A**). Another variant (1 clone) lacked a stop codon. A third variant included  
1419 two deletions of 308 bp and 15 bp, as compared to the reference sequence, which also  
1420 shifted the stop codon. This variant also displayed a 9 bp insertion after the stop codon.  
1421 The two remaining other variants are the closest to the reference sequence, with a  
1422 stop codon located at the same position.

1423 For APL1C, over 31 sequences analysed, 4 different structural variants were identified.  
1424 All of them overlapped with the dsAPL1C fragment (**Figure 21B**). All variant cases  
1425 encoded the stop codon in the same place as the reference sequence (**Figure 21B**).  
1426 The differences between the variants were observed in the 3'UTR region. These  
1427 differences included insertion-deletion events as compared to the reference sequence  
1428 (**Figure 21B**).

1429 Thus, sequence analysis indicated that all APL1A and APL1C gene variants identified  
1430 in the Fd03 colony used in this study should be efficiently targeted by the synthesized  
1431 dsAPL1A and dsAPL1C fragments.

1432

## RESULTS

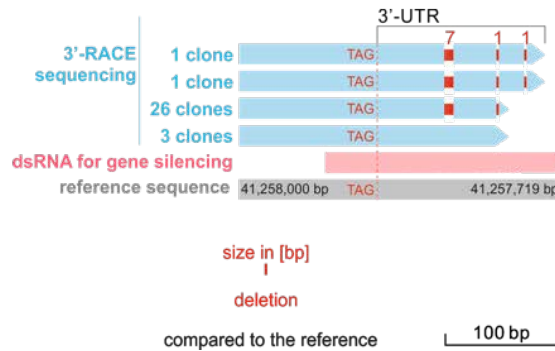
**A**

### APL1A 3'-fragment variants



**B**

### APL1C 3'-fragment variants



1433

1434 **Figure 21. 3' APL1A and APL1C fragments identified by 3' RACE overlap with dsAPL1A**  
 1435 **and dsAPL1C fragments. A.** 3'RACE of APL1A gene identified 5 different 3'APL1A fragment  
 1436 variants. **B.** 3'RACE of APL1C gene identified 4 different 3'APL1C fragment variants.

1437

1438

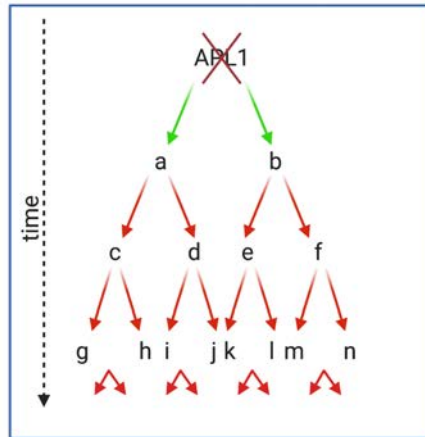
### 1439 3.3.2 Kinetics of APL1A and APL1C silencing

1440

1441 The next step was to establish the correct time point(s) after gene silencing to collect  
 1442 samples for RNAseq. Ideally, as depicted in **Figure 22**, it would be the best to capture  
 1443 the first layer of the genes affected by APL1A and APL1C silencing and avoid the  
 1444 subsequent layer(s) of indirectly affected genes and secondary effects. Collecting the  
 1445 samples at latter time points (red arrow, **Figure 22**) would increase the number of

## RESULTS

1446 differentially expressed genes to test and decrease their informativeness for identifying  
1447 the proximal target genes of putative APL1 signalling.



1448

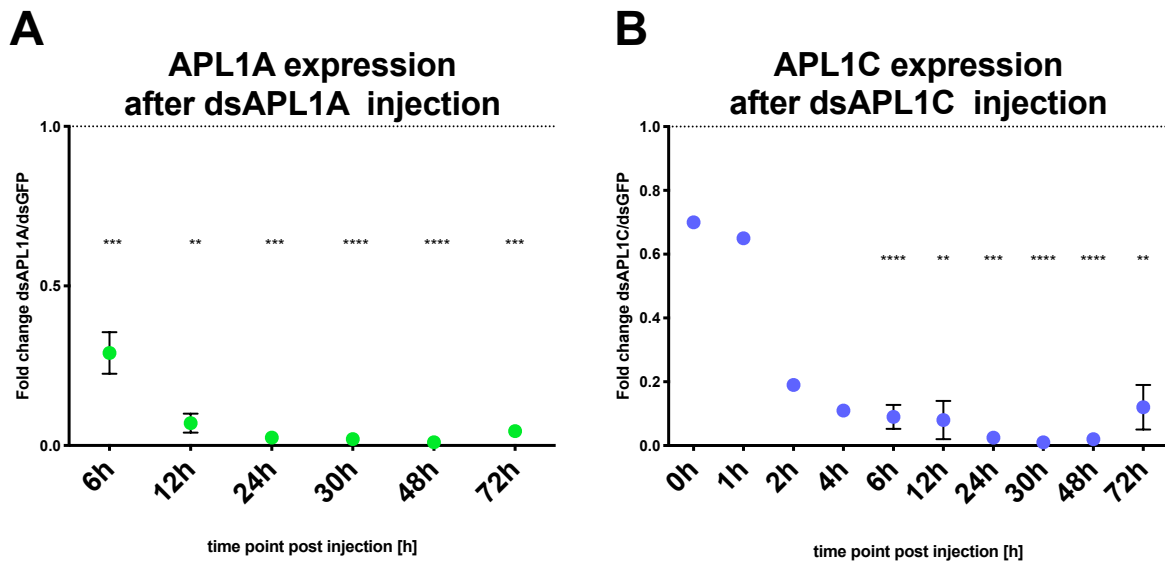
1449 **Figure 22. The potential time effect after APL1 depletion on the number of the affected**  
1450 **genes.** 'a' and 'b' correspond to the hypothetical first layer of genes directly affected by  
1451 depletion of APL1 (green arrows). The influence of APL1 on downstream expression could be  
1452 caused by any combination of transcriptional, translational and post-translational effects. The  
1453 red arrows indicate the genes that are secondary, tertiary and subsequently affected by the  
1454 time post-APL1s depletion.

1455

1456 In order to determine the optimal time to harvest RNA for collecting mosquitoes after  
1457 gene silencing, pools of eight mosquitoes were collected at 6, 12, 24, 30, 48 and 72 h  
1458 for each treatment (dsAPL1A, dsAPL1C and dsGFP). Efficiency of APL1A and APL1C  
1459 transcript silencing was measured by qPCR and determined as expression fold change  
1460 of the dsGFP control group. Silencing of APL1A was efficient already at 6 h with 70%  
1461 of transcript decrease. For the next time points (12, 24, 30, 48 and 72 h post-  
1462 treatment), APL1A transcript was further decreased to at least 90% below controls  
1463 (**Figure 23A**). For APL1C transcript levels were decreased by 90% at 6 h post-dsRNA  
1464 treatment, which persisted until 72 h post-treatment. Therefore, I screened earlier time  
1465 points post-dsAPL1C treatment (0, 1, 2 and 4 h), in order to test how the silencing  
1466 efficiency evolved. The results illustrated in **Figure 23B** showed that the major effect  
1467 of gene silencing occurred 1-2 h after dsRNA treatment.

## RESULTS

1468



1469

1470 **Figure 23. APL1 gene silencing efficiency after dsRNA treatment.** **A.** The kinetics of  
 1471 APL1A transcript depletion after dsAPL1A treatment in *A. coluzzii*. APL1A silencing was  
 1472 verified by qPCR measurement relative to dsGFP controls (dotted line). The ratio of the  
 1473 normalized APL1A cDNA detection in “dsAPL1A” versus “dsGFP” treatments was computed  
 1474 using triplicates from the same cDNA dilution. Graph represents mean with  $\pm$ SEM of the  
 1475 expression fold change between “dsAPL1A” and “dsGFP” control from at least two biological  
 1476 replicates. **B.** The kinetics of APL1C transcript depletion after dsAPL1C treatment. APL1C  
 1477 silencing was verified by qPCR and computed as above. For **A** and **B** data for qPCR analysis  
 1478 was analysed by unpaired t-test (significance levels of t-test p-values: \*\* p-value<0.01; \*\*\* p-  
 1479 value <0.001; \*\*\*\* p-value <0.0001).

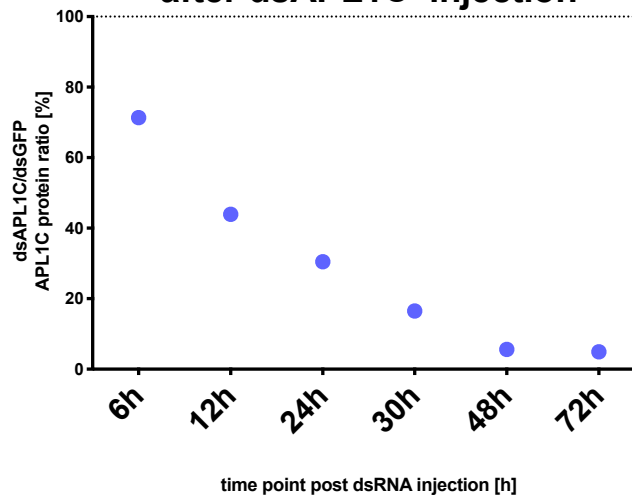
1480

1481 APL1C protein levels were also measured at the same time points. Pools of eight  
 1482 mosquitoes were collected at 6, 12, 24, 48 and 72 h post dsRNA and proteins extracted  
 1483 from mosquitoes were quantified by Western blot.



## RESULTS

### APL1C protein level after dsAPL1C injection



1484

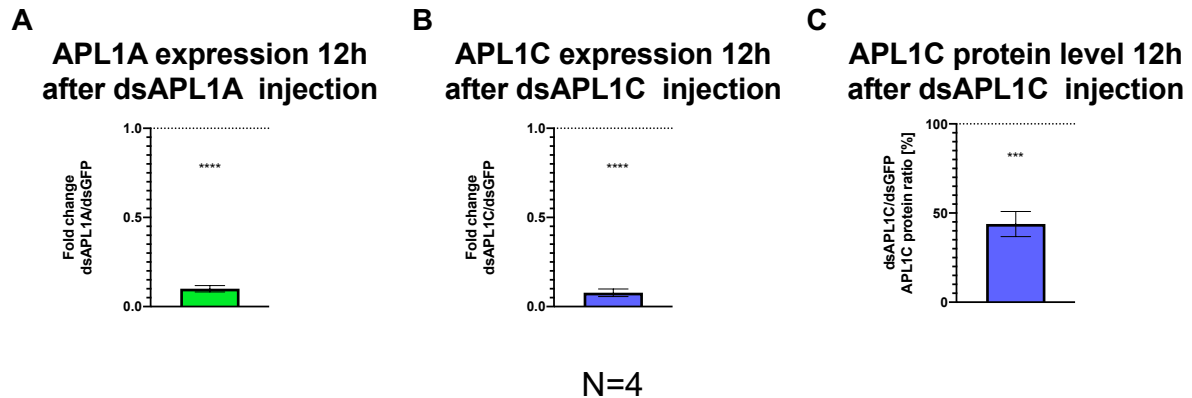
1485 **Figure 24. The kinetics of APL1C protein levels after dsAPL1C.** The graph represents the  
1486 level of APL1C protein in dsAPL1C-treated mosquitoes normalized with the dsGFP control  
1487 (depicted as a dotted line), measured by densitometry. GAPDH density signal was used to  
1488 normalize APL1C signal in both “dsGFP” and “dsAPL1C” treatments, and the normalized  
1489 APL1C signals were then compared between “dsAPL1C” and “dsGFP” conditions.

1490

1491 The amount of APL1C protein decreased by 50% at 12 h post-dsRNA and more than  
1492 90% reduction at 48 and 72 h as compared to dsGFP controls (**Figure 24**). It was  
1493 estimated that 50% protein reduction at 12 h post-dsRNA treatment should be  
1494 appropriate to identify the more proximal gene expression affected by APL1C  
1495 silencing. The protein analysis could not be conducted for APL1A protein because of  
1496 the absence of specific antigen targets for APL1A antibody production. However,  
1497 analysis of the transcript level showed that APL1A silencing was as efficient as APL1C  
1498 at 12 h post-dsRNA treatment (**Figure 23A**). Taken together, for both APL1A and  
1499 APL1C silencing, I decided to collect mosquito RNA samples at 12 h post-dsRNA  
1500 treatment, as the earliest time-point that would preferentially capture the most proximal  
1501 response genes to APL1 depletion. Four independent biological replicates were  
1502 performed, where mosquitoes were injected with dsAPL1A, dsAPL1C or dsGFP as  
1503 control. For each replicate, total RNAs were extracted from a pool of eight collected

## RESULTS

1504 females, at 12 h post-dsRNA treatment. APL1A or APL1C silencing efficiency was  
1505 confirmed in each replicate by qPCR analysis (**Figure 25A** and **B**). In addition, APL1C  
1506 protein decrease in dsAPL1C-injected mosquitoes was confirmed by Western blot  
1507 analysis (**Figure 25C**).



1508

1509 **Figure 25. APL1 genes silencing controls prior to the RNAseq analysis.** **A.** APL1A  
1510 expression is silenced 12 h post dsAPL1A treatment. APL1A silencing was verified by qPCR  
1511 measurement between dsAPL1A and dsGFP (dotted line) treatments. The ratio of the  
1512 normalized APL1A cDNA detection in “dsAPL1A” versus “dsGFP” treatments was computed  
1513 using triplicates from the same cDNA dilution. Graph represents mean with  $\pm$ SEM of the  
1514 expression fold change between “dsAPL1A” and “dsGFP” control from four (N=4) biological  
1515 replicates. **B.** APL1C expression is silenced 12 h post dsAPL1C treatment. APL1C silencing  
1516 was verified by qPCR measurement between dsAPL1C and dsGFP (dotted line) treatments.  
1517 The ratio of the normalized APL1C cDNA detection in “dsAPL1C” versus “dsGFP” treatments  
1518 was computed using triplicates from the same cDNA dilution. Graph represents mean with  
1519  $\pm$ SEM of the expression fold change between “dsAPL1C” and “dsGFP” control from four (N=4)  
1520 biological replicates. **A** and **B** data for qPCR analysis was analysed by unpaired t-test  
1521 (significance level of t-test p-value: \*\*\*\* p-value <0.0001). **C.** APL1C protein level is decreased  
1522 by more than 50% at 12 h post dsAPL1C treatment. The graph represents the level of APL1C  
1523 protein in dsAPL1C-injected mosquitoes normalized with the dsGFP control (depicted as a  
1524 dotted line), measured by densitometry. GAPDH density signal was used to normalize APL1C  
1525 signal in both dsGFP and dsAPL1C treatments. Graph represents mean with  $\pm$ SEM from four  
1526 biological replicates (N=4). Data were analysed by Image Lab 6.0.1 and statistically analysed  
1527 by unpaired t-test (significance level of t-test p-value: \*\*\* p-value <0.001).

1528

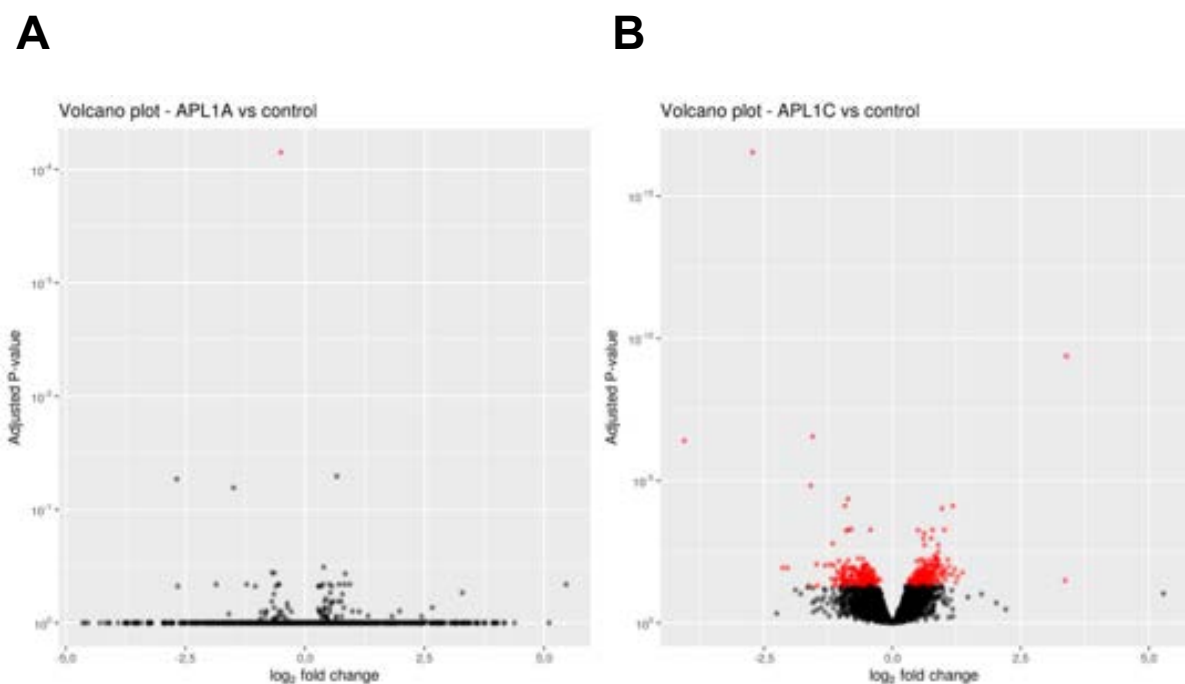
### 1529 3.3.3 RNAseq and bioinformatic analysis

1530

1531 Differentially expressed genes were identified bioinformatically (see Materials and  
1532 Methods). Results of the differential analysis are presented graphically as volcano and

## RESULTS

1533 MA plots (**Figure 26, Figure 27**) as well as summarized in **Table 2**. As depicted on the  
1534 volcano plot, the comparison between dsAPL1A and dsGFP highlighted surprisingly  
1535 only one candidate gene displaying a statistically significant modulated expression  
1536 upon APL1A gene depletion (**Figure 26A**). Overall, the mosquito transcriptome was  
1537 nearly non-affected by the lack of APL1A. In contrast, depletion of APL1C at the same  
1538 time point highly influenced mosquito transcriptome profile, with gene candidates that  
1539 were statistically significantly down- or upregulated (**Figure 26B**).



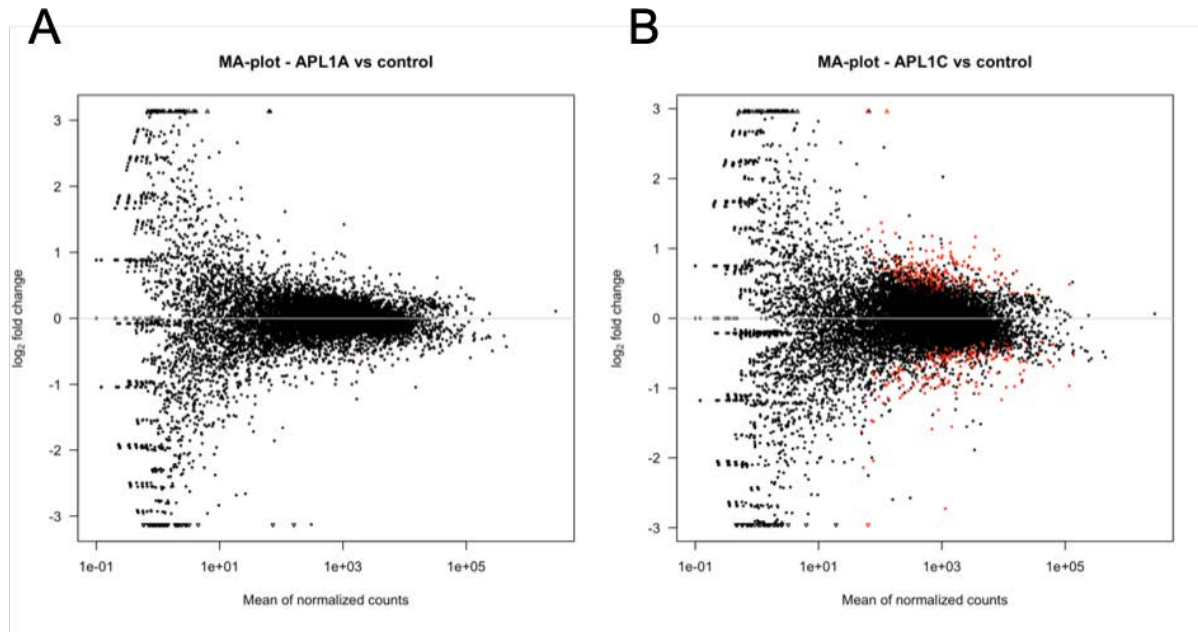
1540

1541 **Figure 26. Volcano plots visualize the differential expression analysis of RNAseq taken**  
1542 **in dsAPL1A vs dsGFP (A) and dsAPL1C vs dsGFP conditions (B).** Each dot represents a  
1543 gene, which expression was compared between the tested conditions (APL1A-depleted  
1544 mosquitoes vs dsGFP control or APL1C-depleted mosquitoes vs dsGFP control). Red dots  
1545 correspond to significantly different genes based on adjusted p-value. For each pairwise  
1546 comparison, raw p-values were adjusted for multiple testing using the Benjamini and Hochberg  
1547 procedure<sup>130</sup>. Genes with adjusted p-values below 0.05 were considered differentially  
1548 expressed

1549

1550 MA-plots for both tested conditions display similar profile among tested conditions  
1551 (**Figure 27**). Nevertheless, APL1C depletion influenced the expression of a greater  
1552 number of genes than APL1A depletion (**Figure 27B**).

## RESULTS



1553

1554 **Figure 27. MA-plots visualizing the log ratio of differential expression as a function of**  
 1555 **the mean intensity taken in dsAPL1A vs dsGFP (A) and dsAPL1C vs dsGFP conditions**  
 1556 **(B).** Each dot represents a gene which expression was compared between the tested  
 1557 conditions. Differentially expressed features are highlighted in red and the triangles correspond  
 1558 to features having a too low/high  $\log_2(\text{FC})$  to be displayed on the plot.

1559

1560 **Table 2** summarizes significantly down- or upregulated genes in mosquitoes depleted  
 1561 for APL1A or APL1C. As reported graphically, APL1A depletion caused  
 1562 downregulation of only one gene, and none were upregulated. In contrast, APL1C  
 1563 silencing significantly regulated 423 genes, 200 downregulated and 223 upregulated.

<b>Tested condition</b>	<b>Downregulated</b>	<b>Upregulated</b>	<b>Total number of differentially expressed genes</b>
dsAPL1A vs dsGFP	1	0	1
dsAPL1C vs dsGFP	200	223	423

## RESULTS

1564 **Table 2. Number of up-, down- and total number of differentially expressed genes in**  
1565 **each compared condition.** The gene was defined as differentially expressed if adjusted p-  
1566 value <0,05.

1567 The only candidate gene from APL1A depleted mosquitoes encodes “defective in cullin  
1568 neddylation protein” (AGAP002513). Among APL1C-downregulated and upregulated  
1569 gene candidates, 112 and 154 genes have an annotated function, respectively (Annex,  
1570 **Table S4** and **Table S5**).

1571

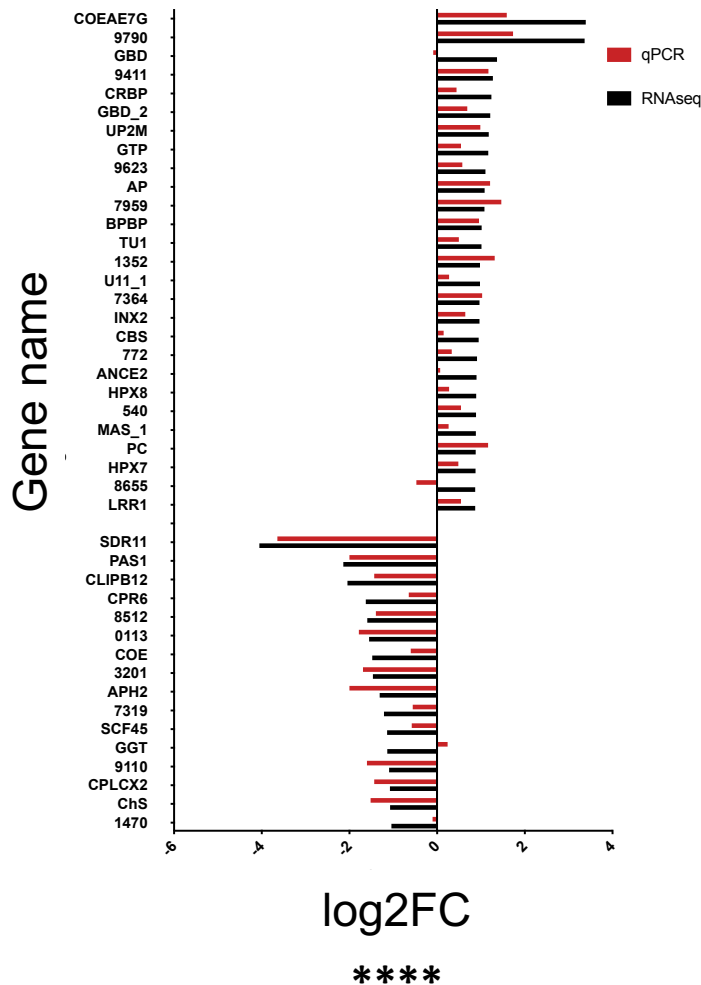
### 1572 **3.3.4 qPCR validation of the RNAseq gene candidates.**

1573

1574 The APL1C gene candidate gene targets revealed by the RNAseq experiment were  
1575 tested by qPCR analysis. Candidates with adjusted p-values<0,05 and log2  
1576 foldchange (log2FC) cut-off of -1 (16 genes) and log2FC cut-off of 0,87 (27 genes), for  
1577 down- and upregulated genes, respectively, were tested. Results showed that the  
1578 expression of tested genes correlated between RNAseq and qPCR experiments  
1579 ( $p<0.0001$  **Figure 28**), thus validating performed RNAseq analysis.

## RESULTS

### Relative expression of a subset of genes as determined by RNAseq and qPCR assays



1580

1581 **Figure 28. qPCR gene expression analysis validates RNAseq APL1C gene targets.** Gene  
 1582 expression is reported as expression log<sub>2</sub> foldchange [log<sub>2</sub>FC, x-axis] measured between  
 1583 “dsAPL1C” and “dsGFP” conditions. Expression of each tested gene is depicted as a red bar  
 1584 for qPCR analysis and black bar for RNAseq analysis. For gene calling (y-axis), annotated  
 1585 genes are called with their gene name or aberrations of gene function, whereas the genes  
 1586 without annotated function are called by the last four digits of their AGAP number. The  
 1587 correlation between qPCR and RNAseq was tested by computing Pearson correlation  
 1588 coefficients. Significance levels of Pearson correlation \*\*\*\* p-value <0.0001).

## RESULTS

### 1589 3.3.5 APL1C controls expression of immune-like genes in the 1590 mosquito

1591  
1592 The 423 genes significantly affected by APL1C depletion were each screened for  
1593 encoded protein domains displaying immune-like function, such as LRRs<sup>99,104,110</sup>,  
1594 TEPs<sup>104,111,120,131</sup>, CLIP-serine proteases<sup>132-135</sup>, serine proteases inhibitors  
1595 (serpins)<sup>136,137</sup>, ankyrin repeats<sup>138</sup>, immunoglobulins<sup>139</sup>, heme peroxidase<sup>140,141</sup> and  
1596 angiotensin-converting<sup>142-145</sup>. Among the downregulated genes, 22 genes encodes for  
1597 such immune-like domains (**Table 3**), whereas 6 genes were identified with such  
1598 domains among the up-regulated genes (**Table 4**). Interestingly, within 22  
1599 downregulated-immune-like genes, I identified nimrod B2 and Scavenger receptors  
1600 coding genes, both being implicated in phagocytosis in mosquitoes<sup>146,147</sup>.

Gene ID	Gene name	FoldChange	log2FoldChange	p-adj	Gene description	Gene description source
AGAP009217	CLIPB12	0,243	-2,043	0,012	CLIP-domain serine protease	VB
AGAP029110	N/A	0,469	-1,093	0,029	C-type lectin-like/link domain superfamily	IP
AGAP001470	N/A	0,486	-1,041	0,007	Leucine-rich repeat, typical subtype	IP
AGAP005496	LRIM12	0,487	-1,039	0,018	leucine-rich immune protein (Short)	VB
AGAP007045	LRIM15	0,514	-0,961	0,015	leucine-rich immune protein (TM)	VB
AGAP028028	LRIM16A	0,514	-0,959	0,012	Leucine-rich immune protein (TM)	UP
AGAP010814	TEP6	0,517	-0,952	0,024	thioester-containing protein 6	VB
AGAP008368	TEP14	0,527	-0,924	0,039	thioester-containing protein 14	VB
AGAP029054	NimB2	0,527	-0,923	0,018	nimrod B2	VB
AGAP028064	LRIM16B	0,535	-0,901	0,042	Leucine-rich immune protein (TM)	UP
AGAP028680	N/A	0,558	-0,842	0,025	Leucine-rich repeat, cysteine-containing subtype, F-box domain	IP
AGAP008091	CLIP1E1	0,573	-0,804	0,001	CLIP-domain serine protease	VB
AGAP028167	N/A	0,581	-0,783	0,034	Proteinase, regulatory CLIP domain, Serine proteases	IP
AGAP001375	SRPN12	0,587	-0,768	0,012	serine protease inhibitor (serpin) 12	VB
AGAP004846	SCRB9	0,667	-0,583	0,017	Class B Scavenger Receptor (CD36 domain)	VB
AGAP004643	SCRB6	0,674	-0,569	0,011	Class B Scavenger Receptor (CD36 domain)	VB
AGAP007061	N/A	0,69	-0,535	0,035	Leucine-rich repeat, typical subtype	IP
AGAP029564	N/A	0,694	-0,527	0,037	Immunoglobulin-like domain, CD80-like, immunoglobulin C2-set	IP
AGAP008366	TEP2	0,72	-0,475	0,018	thioester-containing protein 2	VB
AGAP007556	N/A	0,724	-0,465	0,035	Immunoglobulin subtype 2, Fibronectin type III	IP
AGAP008914	N/A	0,757	-0,401	0,026	Basic-leucine zipper domain	IP
AGAP006645	N/A	0,809	-0,307	0,034	Leucine-rich repeat, typical subtype	IP

1601  
1602 **Table 3. List of the immune-like genes, which expression is downregulated upon APL1C**  
1603 **depletion.** Genes are enlisted according to the descending expression fold change between  
1604 “dsAPL1C” and “dsGFP” (control) treatments. The columns include gene ID referred as the  
1605 AGAP number; gene name, N/A stands for not annotated gene; expression foldchange and  
1606 log2foldchange, adjusted p-value (p-adj), gene description including encoded protein function  
1607 or predicted protein domain based on the databases: VB-VectorBase Community Annotation,  
1608 IP-Interpro, UP-UniPro.

1609

## RESULTS

Gene ID	Gene name	FoldChange	log2FoldChange	p-adj	Gene description	Gene description source
AGAP009751	ANCE2	1,868	0,901	0,00605	angiotensin-converting enzyme 2	VB
AGAP004038	HPX8	1,858	0,893	0,0032	heme peroxidase 8	VB
AGAP004036	HPX7	1,838	0,878	0,02739	heme peroxidase 7	VB
AGAP003591	N/A	1,826	0,869	0,00944	LRR-repeat protein 1	VB
AGAP011534	N/A	1,51	0,594	0,02302	Ankyrin repeat Ankyrin repeat and LEM domain-containing protein	IP
AGAP004563	ANCE9	1,313	0,393	0,04951	angiotensin-converting enzyme 9	VB

1610

1611 **Table 4. List of the immune-like genes, which expression is upregulated upon APL1C**  
 1612 **depletion.** Genes are enlisted according to the ascending expression fold change between  
 1613 “dsAPL1C” and “dsGFP” (control) treatments. The columns include gene ID referred as the  
 1614 AGAP number; gene name, where N/A stands for not annotated gene; expression foldchange  
 1615 and log2foldchange, adjusted p-value (p-adj), gene description including encoded protein  
 1616 function or predicted protein domain based on the databases: VB- VectorBase Community  
 1617 Annotation, IP-Interpro, UP-UniPro.

1618



## DISCUSSION

1619

## **IV. DISCUSSION**

1620

1621 **4.1 Dual functionality for APL1C as guard and PRR-like**1622 **factor**

1623 The LRR APL1 proteins are the key components of the mosquito defence against  
1624 *Plasmodium* infection<sup>99,102,103</sup>. However, little was known about the mechanism of  
1625 APL1 function in mosquito anti-*Plasmodium* defence. This thesis work showed that  
1626 APL1C can bind to both *Plasmodium* ookinetes and sporozoites and thus acts as a  
1627 pathogen binding factor. Therefore, I propose that APL1C could behave as a PRR in  
1628 heterocomplex with LRIM1 for initiating complement-like cascade via TEP1<sub>cut</sub>  
1629 stabilization and binding to the parasites at least in mosquito midgut, which probably  
1630 elicit a signalling pathway, as shown that APL1C controls the expression of several  
1631 genes in *A. coluzzii*. In addition, it was observed that APL1C presence on the mosquito  
1632 midguts was not restricted to the parasite surface, as it could form hemolymph-  
1633 originated, extracellular layer on the basal side of the midgut epithelial cells. The  
1634 APL1C protein level is induced by the bloodmeal independently from the presence of  
1635 the parasite, which suggests that APL1C acts as a guard protein in mosquito  
1636 hemolymph, where beneath basal lamina of the midgut, it could form a state of immune  
1637 readiness to face *Plasmodium* and probably other potential microbes that may emerge  
1638 from the blood-engorged midguts. The bloodmeal thus constitutes a generalized signal  
1639 that induces APL1C and other immune factors that may recognize a large range of  
1640 blood borne potentially pathogenic microbes. In the case of the malaria parasites,  
1641 when the ookinetes traverse the midgut epithelium and reach the basal side of the cells  
1642 into the extracellular space, APL1C appears ready to “catch” the parasites through a  
1643 binding process for blocking their further development. It has been recently reported

## DISCUSSION

1644 that mosquito infection with Gram-negative, but not Gram-positive bacteria specifically  
1645 induces APL1C cleavage by a still unidentified protease<sup>122</sup>. My thesis work showed  
1646 that APL1C binds to the *Plasmodium* ookinetes but we still do not know if this binding  
1647 requires or not the cleavage process, or if the cleavage of APL1C occurs posteriori of  
1648 the parasite binding. Another possibility would be that previously reported cleavage of  
1649 APL1C is bacteria-specific, and thus does not occur in the case of *Plasmodium*  
1650 parasites. PRR have been described mainly as transmembrane but can be also  
1651 present as soluble receptors. The peptidoglycan recognition proteins (PGRPs) are  
1652 conserved in both vertebrates and insects. In *Drosophila*, transmembrane, cytosolic  
1653 and soluble forms of PGRPs display recognition and catalytic properties linked to their  
1654 capacity to bind PGN<sup>148-150</sup>. C-type lectins (CTLs) are another example of PRRs that  
1655 occur extracellularly and are either secreted or membrane-bound. C-type lectins are  
1656 pattern recognition receptors that function in both mammalian and insect innate  
1657 immunity through the carbohydrate recognition domain. In *A. gambiae* two CTLs  
1658 cooperate to defend the mosquito against Gram-negative bacterial infections. CTL  
1659 function was also assessed towards mosquito *Plasmodium* infection, where two CTL,  
1660 CTL4 and CTLMA2 protect the ookinetes against a potent mosquito innate immune  
1661 response, melanization<sup>151,152</sup>. However, the most intriguing resemblance with the  
1662 APL1C-LRIM1 LRR protein complex is the jawless vertebrate VLRB proteins. They are  
1663 secreted LRR-antibody-like receptors inducing the classical-complement cascade<sup>92</sup>  
1664 and circulate as disulphide-bridged multimers<sup>153</sup>. Therefore, the combinatorial immune  
1665 LRR protein complexes could constitute a repertoire of antibody-like structures in  
1666 *Anopheles* mosquitoes for facing distinct microbial infections.

1667

1668 **4.2 The HdMvs, a source of secreted APL1C, mediated by**  
1669 **the nitration pathway?**

1670 Phagocytic hemocytes represent more than 80% of total hemocytes in  
1671 mosquitoes<sup>154,155</sup> and were reported to affect TEP1 binding to the ookinete on infected  
1672 midguts<sup>127</sup>. Presented results demonstrated the important role of the phagocytic cells  
1673 as a source of APL1C expression. APL1C proteins, which are expressed in hemocytes,  
1674 and especially in phagocytic hemocytes, are localized in cytoplasmic vesicles. This  
1675 suggests that, consistent with the presence of a signal peptide on its protein sequence,  
1676 APL1C is secreted through this vesicle structure. Nitration has been shown to control  
1677 release of hemocyte-derived microvesicles (HdMv) that were proposed to transport  
1678 key factors mediating subsequent TEP1 binding and lysis of ookinetes<sup>128</sup>. It was found  
1679 that silencing the nitration pathway, and thus inhibiting the release of HdMv, also  
1680 reduced APL1C abundance in the midgut extracellular space, strongly suggesting that  
1681 APL1C may be part of the payload of HdMv. To demonstrate this assumption, one  
1682 possibility would be to physically isolate and collect these HdMvs and perform either a  
1683 proteomic study or, more directly, a Western blot analysis using anti-APL1C antibody.  
1684

1685 **4.3 APL1C presence in hemolymph at the basal side of the**  
1686 **midgut epithelium is not PAMP dependent**

1687 Nitric oxide (NO) production was correlated with *Plasmodium* ookinete invasion of the  
1688 midgut epithelium<sup>156</sup> and was described as mediating protein nitration of ookinete-  
1689 invading epithelial cells ('Time Bomb' model)<sup>157</sup>. Activation of the nitration pathway

## DISCUSSION

1690 leads to ookinete tagging as a danger (or non-self) signal, which promotes the  
1691 activation of the mosquito complement cascade<sup>140</sup>. This suggests that activation of the  
1692 immune response in mosquito is not initiated by *Plasmodium* PAMP but by parasite-  
1693 induced tissue damage signals (DAMP) such as nitration. In line with this assumption,  
1694 a study showed that TEP1 binds damaged sperm cells in a male mosquito in nitration  
1695 pathway-dependent manner<sup>158</sup>. In the case of APL1C, presented results showed that  
1696 a non-infected bloodmeal (NBM) also induced APL1C protein levels in mosquito.  
1697 Because bloodmeals increase the abundance of the bacterial flora<sup>39</sup>, I tested whether  
1698 the enteric microbiome can induce elevated APL1C levels in the hemolymph, the  
1699 source of the APL1C protein measured by IFA on dissected midguts. However, a  
1700 similar abundance of APL1C protein between ATB-treated and control midguts groups,  
1701 ruled out the microbiome as potential signal elicitor leading to increased APL1C protein  
1702 level. Therefore, an alternative hypothesis for this signal would be related to the  
1703 mechanical stretching of the midgut induced by the bloodmeal, which could damage  
1704 the cellular integrity of the epithelium. To test this hypothesis, we could perform a  
1705 blood-free feeding that would mimic the mechanical distension of the midgut epithelium  
1706 and assess APL1C presence on the midgut epithelium.

1707 Another possibility for the induction signal would be related to the physiological stress  
1708 induced by bloodmeal chemical components or metabolites. Digestion of vertebrate  
1709 haemoglobin results in the production of large amounts of free heme, a cytotoxic  
1710 molecule that generates reactive oxygen species, which can disrupt phospholipid  
1711 bilayers of cell<sup>159</sup>. Thus, the oxidative stress induced by the bloodmeal could serve as  
1712 a danger signal, inducing the presence of LRR protein in the hemolymph, which  
1713 diffuses through the porous basal lamina into the extracellular space around the  
1714 midgut.

## DISCUSSION

1715 Taken together, knowing that APL1C presence on the midgut in the absence of  
1716 *Plasmodium* is even not microbiome-related strengthens the hypothesis of a PAMP-  
1717 independent source of signal to induce APL1C secretion in the hemolymph.

1718

### 1719 **4.4 The sophisticated sporogony control by APL1C**

1720 Little attention has been paid on the mosquito immune response against the last stage  
1721 of *Plasmodium* development, which are the sporozoites. My results showed that  
1722 APL1C protein binds to the released sporozoites in the hemolymph, and more  
1723 importantly, it also controls sporozoite infection prevalence in the salivary glands. The  
1724 temporal behaviour of APL1C protein during sporogonic development appeared  
1725 consistent with its immune functionality against the ookinete and the late sporozoite  
1726 stages. APL1C protein levels were unchanged at the time corresponding to the oocyst  
1727 stage, suggesting that it is not active against this parasite stage. This is consistent with  
1728 other reports that the oocyst stage does not seem to stimulate immune responses.  
1729 Indeed, oocyst surface proteins such as laminin are in fact mosquito host-derived  
1730 factors<sup>35</sup>, and these could hide the oocyst from the immune system recognition. The  
1731 oocyst acquires metabolic resources from the mosquito for its enormous growth and  
1732 internal cell divisions to generate thousands of sporozoites, however this appears to  
1733 occur quietly from an immune perspective, until oocyst rupture. Therefore, potential  
1734 inflammatory responses or cell apoptosis leading to DAMPs production are probably  
1735 not prominent during the oocyst stage.

1736 In addition, thesis results showed that APL1C protective function against the  
1737 sporozoites does not involve the same factors as operate against the ookinete stage.  
1738 TEP1 and TEP3 that were both reported to control *P. berghei* ookinetes and to interact

## DISCUSSION

1739 with the APL1C/LRIM1 complex<sup>104,111</sup> are not involved in protection against sporozoite  
1740 salivary gland invasion. However, I cannot rule out that the control of sporozoites in  
1741 collaboration with APL1C would require other TEP factors. In addition, it has been  
1742 reported that unlike to the ookinete that harms the midgut epithelial cells, this is not  
1743 likely the case for the sporozoites on the salivary gland epithelium as neither NOS  
1744 increase, nor apoptosis has been reported<sup>137</sup>. Finally, a low proportion of phagocytized  
1745 sporozoites was observed<sup>46</sup>, suggesting that the phagocytic cells are activated during  
1746 sporozoite release, and this activation might contribute to APL1C release in the  
1747 hemolymph. Taken together, additional work is still needed to fully elucidate the killing  
1748 mechanism of the sporozoites induced by APL1C.

1749

### 1750 **4.5 Whole transcriptome analysis reveals differential** 1751 **signalling function of APL1 genes**

1752 Genomic and evolutionary studies revealed that immune-related genes and especially  
1753 microbial-sensor coding genes such as PRR evolve faster than other insect genes<sup>160-</sup>  
1754 <sup>162</sup>. The APL1 gene family underwent an expansion from a single ancestral gene to  
1755 three genes in the lineage of African *Anopheles*<sup>163</sup>. The expanded APL1 genes  
1756 acquired distinct immune activities, as exemplified by functional differences in anti-  
1757 *Plasmodium* responses<sup>103</sup>. My RNAseq data analysis of APL1A and APL1C silencing  
1758 treatments revealed distinct profiles of downstream-modulated genes. While depletion  
1759 of APL1A expression did not induce differences in overall transcriptome, APL1C  
1760 depleted treatment led to high number of modulated genes. Interestingly, APL1C  
1761 appears to be orthologous to the ancestral single APL1 gene of *Anopheles*<sup>163</sup>.

## DISCUSSION

1762 Therefore, it should likely still display the more ancestral APL1 functions, while the  
1763 APL1A and APL1B paralogs are derived forms and evolved new functions. APL1C  
1764 might protect against a broader spectrum of microbes as compared to, for example,  
1765 APL1A, which could respond to a more specialized microbial profile. Consistently,  
1766 functional studies revealed that, in addition to the rodent malaria parasites, APL1C  
1767 protects against *E. coli* bacteria<sup>122</sup>.

1768

1769 Even though my thesis did not concentrate on the allelic variation of each gene, my 3'  
1770 RACE sequencing study of APL1A and APL1C identified allelic variants. It was  
1771 previously reported that among distinct APL1A alleles segregating in the Ngousso *A.*  
1772 *coluzzii* colony, only the APL1A<sup>2</sup> allele was associated with the protective phenotype  
1773 of APL1A in *P. falciparum* challenge experiments<sup>106</sup>. My RNAseq experiment used the  
1774 Fd03 colony, which has not been tested in a similar association study. Therefore, it  
1775 cannot be excluded that current Fd03 colony may not carry the APL1A<sup>2</sup> allele, or  
1776 another protective APL1A allele. The underlying hypothesis would be that protective  
1777 variants would lead to a stronger transcriptional response, detectable by RNAseq, as  
1778 compared to non-protective alleles. Further genomic and functional experiments would  
1779 be necessary to test this hypothesis.

1780



1781  
1782

**V. CONCLUDING REMARKS AND FUTURE  
PERSPECTIVES**

## CONCLUDING REMARKS AND FUTURE PERSPECTIVES

1783

1784 My thesis work gave insight in the function of LRR proteins in *A. coluzzii* mosquito.

1785 Focusing on the antimalarial APL1C gene as a representative of immune LRR family

1786 member, my results showed that it acts as pathogen binding factor in the mosquito

1787 hemolymph. I present evidence that APL1C and LRIM1 are directly associated to the

1788 *Plasmodium* ookinete surface, that APL1C binding to the parasite requires activity of

1789 the nitration pathway and the participation of phagocytic hemocytes. In addition to

1790 limiting the midgut stage of infection, presented results revealed that APL1C binds to

1791 the surface of free sporozoites in the hemocoel, and is required to limit salivary gland

1792 infection by sporozoites. These results are consistent with a model in which APL1C

1793 protein acts as a pathogen recognition factor released from hemocytes at elevated

1794 levels following a bloodmeal (**Figure 29**), diffuses in the hemolymph to form a

1795 protective layer in the midgut basal extracellular space and also protects the open

1796 systemic compartment from microbes (such as sporozoites) escaping the epithelial

1797 barrier of the midgut. APL1C release after a bloodmeal is triggered by nitration that

1798 could function as a DAMP. The bloodmeal induction of APL1C creates a state of

1799 immune readiness on the midgut epithelium against microbes ingested with the blood

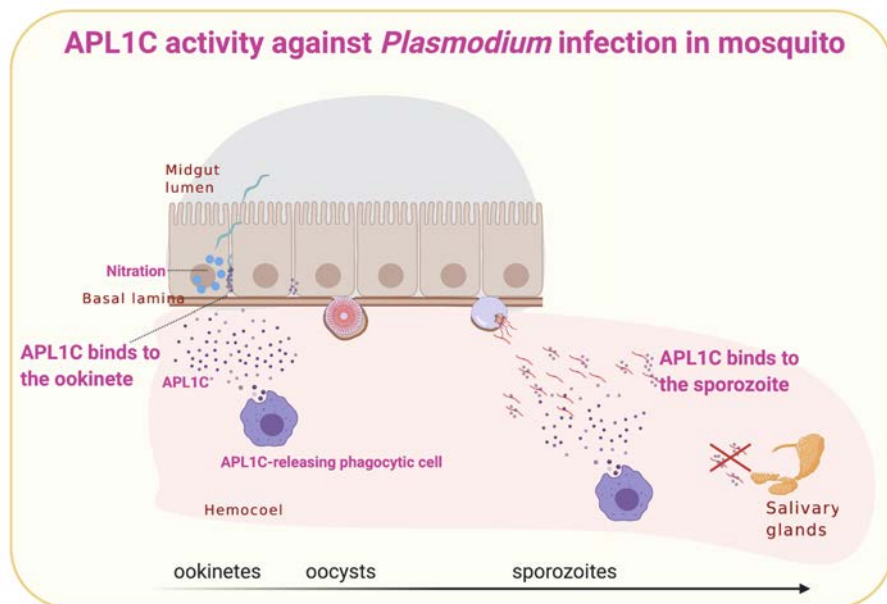
1800 that could infect the mosquito. Furthermore, APL1C appears to be an active

1801 recognition factor in the hemolymph, where it binds to sporozoites and apparently tags

1802 them for killing by unknown effectors.

1803

1804



1805 **Figure 29. Proposed model of APL1C function in anti-*Plasmodium* response.** Detailed  
 1806 description in text.

1807

1808 Assuming that as PRR-like factors APL1 genes might elicit a signal transduction  
 1809 pathway, I performed a RNAseq analysis to identify the downstream factors of APL1A  
 1810 and APL1C. Consistently with their distinct activity, the results revealed distinct  
 1811 transcriptomic profiles between both genes. These results are consistent with the  
 1812 hypothesis that mosquito immune LRR proteins are functional counterparts of the  
 1813 immunoglobulin-based receptors that vertebrates use for antigen recognition, although  
 1814 the parasite surface ligands remain unknown.

1815 Lastly, it would be interesting to enlarge these findings to other members of the LRR  
 1816 family and assess the pathogen-specificity of their combinatorial binding property, as  
 1817 this family could represent the first secreted LRR pathogen recognition factors in the  
 1818 invertebrates.

1819

1820

## **VI. MATERIALS AND METHODS**

## MATERIALS AND METHODS

### 1821 **Ethic statement**

1822 This study was conducted in strict accordance with the recommendations from the  
1823 Guide for the Care and Use of Laboratory Animals of the European Union (European  
1824 Directive 2010/63/UE) and the French Government. The protocol was approved by the  
1825 “Comité d’éthique en expérimentation animale de l’Institut Pasteur” CETEA 89 (Permit  
1826 number: 2013–0129), by the French Ministry of Scientific Research (Permit number:  
1827 202195.02) and undertaken in compliance with Institut Pasteur Biosafety Committee  
1828 (protocol CHSCT 14.114).

1829

### 1830 **Mosquitoes**

1831 Mosquitoes *Anopheles coluzzii* colony Fd03 initiated in Mali <sup>164</sup> was reared at 26°C and  
1832 80% humidity, on a 12 h light/dark cycle with access to cotton soaked in 10% sucrose  
1833 solution, in Genetics and Genomics of Insect Vectors Unit insectary, Institut Pasteur,  
1834 Paris, France.

1835

### 1836 **Mouse infections with *Plasmodium berghei***

1837 All experiments were performed with three-week-old female Swiss *Mus musculus* mice  
1838 (Janvier, France). Mice were inoculated with 10<sup>5</sup> red blood cells (RBCs) infected with  
1839 GFP-expressing under the *hsp70* regulatory regions *P. berghei* parasite<sup>165</sup>. Four days  
1840 post-treatment, blood samples were taken from the tail, and parasitemia was  
1841 determined by flow cytometry. Furthermore, male gametocyte maturity was verified by  
1842 performing an exflagellation test as previously described<sup>166</sup>.

1843

## MATERIALS AND METHODS

### 1844 **Mosquito feeding on non-infected mice or *P. berghei*-infected mice**

1845 For all the experiments, *P. berghei*-infected mice were used at 4-8% parasitemia, with  
1846 mature gametocytes. Exflagellation tests were performed to evaluate the maturity of  
1847 male gametocytes. First, both non-infected and infected mice were anesthetized by  
1848 intraperitoneal (IP) treatment of ketamine (Imalgene 1000 at 125 mg/kg) and xylazine  
1849 (Rompun 2% at 12.5 mg/kg), then mosquitoes were allowed to feed on the  
1850 anesthetized mice for 20 min. Unfed mosquitoes were discarded and fed mosquitoes  
1851 were maintained at 21°C and 70% relative humidity on 10% sucrose solution as  
1852 previously described <sup>103</sup>.

### 1853 **Midguts immunofluorescence assay (IFA)**

1854 Immunostainings were performed as previously described <sup>128</sup>. Briefly, mosquitoes  
1855 were subjected to *P. berghei*-infected bloodmeal (IBM), non-infected bloodmeal (NBM)  
1856 or 10% sucrose solution (unfed, UF) and incubated at 21°C. 24 h after, midguts from  
1857 each condition were dissected in sterile 1× PBS, fixed for 40s in 4% paraformaldehyde  
1858 for easier removal of the blood and opened longitudinally in sterile PBS to remove the  
1859 blood bolus. Cleaned midguts were then fixed in 4% paraformaldehyde in 1× PBS for  
1860 1 hour at room temperature and rinsed five times in sterile PBS for 5 minutes each. In  
1861 a next step, midguts were permeabilized, and blocked in PBT (PBS supplemented with  
1862 1% bovine serum albumin (BSA) and 0.1% Triton X-100) for 2 h at room temperature.  
1863 Midguts were then incubated overnight with a rabbit polyclonal anti-APL1C antibody  
1864 (pAb) <sup>102</sup> (1:600 in PBT) at 4°C. The next day, midguts were rinsed five times (10 min  
1865 each) with PBT, followed by a 2-hour incubation at room temperature with Alexa Fluor  
1866 647-conjugated anti-rabbit IgG secondary antibody (Thermofisher) (1:500 in PBT) at  
1867 room temperature. Then, midguts were rinsed five times in PBT (10 min each).

## MATERIALS AND METHODS

1868 Immunostained tissues were counterstained for 10 min with 405 phalloidin-iFluor  
1869 (actin, 1:1,000 abcam) and subsequently for 10 minutes with DAPI (nuclei, 10 µg/ml,  
1870 Thermofisher). All incubations were performed with 12 rpm rocking. Lastly, tissues  
1871 were rinsed in sterile PBS and mounted in Vectashield (Vector Laboratories). Antibody  
1872 accessibility to whole midgut tissue (permeabilization control) was performed as  
1873 described above, using anti-GFP Alexa Fluor 647-conjugated primary antibody  
1874 (Thermofisher).

### 1875 **Confocal Microscopy**

1876 Midguts where APL1C signal was subjected for quantitative analysis were imaged on  
1877 a laser-scanning confocal microscope (LSM700, Carl Zeiss Jena). All acquisitions  
1878 were configured as follow: pinhole size 1 Airy unit; DAPI/Phalloidin channel: 405nm  
1879 laser, SP490 filter; APL1C channel: LP640 filter, laser 633nm; images are 1024 x 1024,  
1880 digitised over 16-bit; Z step interval: 0.33 µm; objective 40x, oil, NA=1.3, Plan-Neofluar.  
1881 Ookinetes were subjected for microscope analysis using High-speed spinning-disk  
1882 confocal system [(UltraVIEW ERS; Perkin Elmer, mounted on an inverted Axiovert 200  
1883 microscope (Carl Zeiss)].

### 1884 **Image analysis**

1885 Images were analysed with Fiji <sup>167</sup>. 3D Images were filtered first with a 3x3 median filter  
1886 then with a Gaussian filter with sigma=0.7 pixels. The 3D stacks were then projected  
1887 using Maximum Intensity Projection. A selection was created using an intensity  
1888 threshold using ImageJ 'default' method. Finally, **Figure 7B**, **Figure 14C** and **Figure**  
1889 **15C** report the mean intensity for the channel APL1C in the selection.

### 1890 **Dextran injection**

## MATERIALS AND METHODS

1891 Using a Nanoject II injector (Drummond Scientific), cold-anesthetized *A. coluzzii*  
1892 females were injected intrathoracically with 138 nl of 70 kD dextran 488-conjugate of  
1893 concentration 10 mg/ml. 1,5 h post-injection, midguts were dissected, fixed and  
1894 counterstained with DAPI and phalloidin as described above. Injections were  
1895 performed with a glass capillary needle as previously described<sup>102,168</sup>.

### 1896 **Plasmids construction**

1897 APL1C and LRIM1 alleles were amplified by PCR from *A. coluzzii* Ngousso mosquitoes  
1898 using the primers flanking the coding regions of each gene (**Table S6**) and cloned into  
1899 a dual His- and V5-tag insect expression vector (pAc5.1 V5-His, Invitrogen). To  
1900 generate a plasmid control, RFP (Red Fluorescent protein) coding sequence from the  
1901 mammalian expression vector pTagRFP (Evrogen) was cloned in the pAc5.1 V5-His  
1902 expression vector using EcoRI and NotI restriction enzymes. To ensure RFP protein  
1903 secretion in the culture medium, the signal sequence from APL1C was amplified from  
1904 *A. coluzzii* Ngousso mosquitoes with the primers (**Table S6**) and cloned in the pAc5.1  
1905 V5-His-RFP expression vector.

### 1906 **Insect cell culture and transfection**

1907 *A. coluzii* derived 4a-3A hemocyte-like cells were cultured in monolayer at 27°C in  
1908 Insect X press medium (Lonza) supplemented with 5% fetal bovine serum (GIBCO  
1909 BRL) and 50µg/ml gentamycin (Sigma). Before transfection,  $1 \cdot 10^5$  cells were allowed  
1910 to grow in an 8-wells Lab-Tek chamber slide system (Thermo Scientific) for one hour.  
1911 Cells were transfected with the expression vector using cellfectin II reagent (Invitrogen)  
1912 according to the manufacturer's protocol.

### 1913 **Ex vivo midgut assay**



## MATERIALS AND METHODS

1914 24 h post *P.berghei*-infected blood feeding (IBM), *A. coluzzii* Ngousso midguts were  
1915 dissected in PBS and placed for 2 h at 21°C in culture medium of 4a-3A hemocyte-like  
1916 cells that were transfected 3 days before with pAc5.1 V5-His-RFP, pAc5.1 V5-His-  
1917 APL1C or pAc5.1 V5-His-LRIM1 expression vectors. After incubation, transfected cells  
1918 and dissected midguts treated for immunostaining analysis with a mouse V5  
1919 monoclonal antibody (mAb, Invitrogen) at 1:500 or a mouse Pbs28 mAb (without  
1920 permeabilization)<sup>169</sup>, followed by 1 h with Alexa Fluor 594-conjugated anti-mouse IgG  
1921 secondary antibody at 1:2,000. Nuclei were stained with Hoechst 33342. After  
1922 mounting in SlowFade Gold antifade reagent (Molecular Probes), samples were  
1923 observed using a Leica DM 5000 B fluorescent microscope.

### 1924 **Double-stranded RNA synthesis and gene silencing**

1925 Double-stranded RNAs were synthesized from PCR amplicons using the T7  
1926 Megascript Kit (Ambion) as described previously<sup>102</sup>. The sequences of the primers  
1927 used for synthesis of dsRNA for each gene are enlisted in Annex **Table S7**. 500 ng of  
1928 dsRNA were injected into the thorax of cold-anesthetized *A. coluzzii* females using a  
1929 nano-injector (Nanoject II, Drummond Scientific). dsGFP was used as a control.  
1930 Treatments were performed with a glass capillary needle as previously  
1931 described<sup>102,168</sup>. The efficiency of the gene silencing was monitored by quantitative  
1932 Polymerase Chain Reaction (qPCR) and western blot analysis where specific antibody  
1933 was available.

### 1934 **RNA extraction and RT**

1935 Total RNA was extracted with TRIzol reagent (Invitrogen). cDNAs were generated  
1936 using M-MLV reverse transcriptase (Invitrogen) from 1 µg of total RNAs in 40 µL.

### 1937 **Quantitative polymerase chain reaction (qPCR)**

## MATERIALS AND METHODS

1938 Using cDNA or genomic DNA, all qPCRs were performed as described in <sup>104</sup>, using  
1939 SYBR green supermix (KAPA SYBR FAST ABI, Sigma-Aldrich) and the CFX96 Touch  
1940 Real-Time PCR Detection System (Biorad). Ribosomal protein *rpS7* gene was used  
1941 as an internal control. The quantification of each gene was obtained as a ratio of  
1942 the *rpS7*. All the primers sequences are enlisted in Annex **Table S7**. Analysis of the  
1943 expression of transcript relative to *rpS7* was performed according to the  $2^{-\Delta\Delta Ct}$   
1944 method <sup>170</sup>. PCR conditions of the run were: 95°C for 10min, then 40 cycle of [95°C for  
1945 15 sec, 60°C for 1min (plateread), melting curve analysis].

### 1946 **Protein extraction and western blotting from whole mosquitoes**

1947 Pools of 8 mosquitoes were placed in tubes containing glass beads with 150µl of  
1948 freshly prepared 1X RIPA buffer (Cell Signalling) and 1X protease inhibitor cocktail  
1949 (Roche). Samples were grinded in FastPrep using 3 cycles of 30s with 1400 rpm  
1950 shaking each (with 30s breaks between each cycle to avoid sample heating). Grinded  
1951 samples were centrifuged for 5 min, 18,000g, 4°C. Collected supernatant was  
1952 sonicated in ice and centrifuged for 15 min, 18,000g, 4°C. Reduced protein samples  
1953 were prepared by adding XT sample buffer (Bio-Rad), 1X DTT and heating at 95°C for  
1954 5 min. Samples were loaded and separated on 4-12% Criterion SDS-PAGE gels (Bio-  
1955 Rad). Protein from equivalent of 0.2 mosquito were transferred to the membrane using  
1956 Trans-Blot Turbo Transfer System (Bio-Rad), after which immunoblots were blocked  
1957 for 1 hour in 5% BSA in TBST (Tris-buffered saline, 0.1% Tween 20). Immunoblots  
1958 were probed overnight with a rabbit pAb anti-APLC antibody at 1:5,000 with 5% BSA  
1959 in TBST. Next day, membranes were rinsed three times in TBST for 10 minutes each,  
1960 stained by a mouse anti-GAPDH antibody (Thermofisher) at 1:4,000 and followed by  
1961 horseradish peroxidase-conjugated anti-mouse IgG secondary antibody  
1962 (Thermofisher) at 1:8,000, which was used as a loading control. Detection step was

## MATERIALS AND METHODS

1963 performed using the Enhanced chemiluminescence (ECL) system (Biorad) following  
1964 the manufacturer's instructions. The levels of APL1C were quantified by densitometry  
1965 using the Image Lab analysis software (BioRad). To normalize expression, the ratio of  
1966 the GAPDH density from the control sample to the GAPDH density from sample of  
1967 interest was determinate as a normalization factor. APL1C band intensity was  
1968 normalized using GAPDH normalization factor.

### 1969 **Mosquito antibiotic treatment (ATB)**

1970 Mosquitoes were treated with an antibiotic (ATB) cocktail: penicillin 62.5 µg/mL,  
1971 streptomycin 100 µg/mL and gentamicin 50 µg/mL in 10% sucrose solution directly  
1972 after emergence and were kept under ATB pressure for three days, and ATB solution  
1973 were changed every day. Mosquitoes treated (+ATB) or non-treated (-ATB) (for  
1974 control) were fed on naive mouse for 20 min. Unfed mosquitoes were discarded and  
1975 fed mosquitoes were maintained at 21°C and 70% relative humidity. For the (+ATB)  
1976 group the antibiotic pressure was maintained after the feeding until their collection for  
1977 midgut dissection.

### 1978 **Microbiota analysis**

1979 Dissection was performed as described previously <sup>171</sup>. Briefly, before dissection,  
1980 mosquitoes were washed in 75% ethanol for 5 min, and then three times in sterile PBS  
1981 to eliminate non-attached bacteria, thus preventing bacteria contamination during  
1982 dissection. For each biological replicate, midguts were collected from each mosquito  
1983 group (-ATB) and (+ATB), frozen immediately on dry ice and stored at -80°C until  
1984 processing. To assess antibiotic effectiveness, mosquito midguts from each group (-  
1985 ATB) and (+ATB) were collected 24h post-feeding. Bacterial DNA was extracted with  
1986 DNeasy PowerSoil Kit (QIAGEN). The V4 region of the 16S rDNA (Annex, **Table S7**)

## MATERIALS AND METHODS

1987 was used for measuring the total bacterial abundance by quantitative PCR. DNA  
1988 samples from each independent biological replicate were used to perform distinct  
1989 qPCR in triplicate and fold changes obtained between (+ATB) and (-ATB) were  
1990 combined as a mean and illustrated graphically.

### 1991 **Clodronate-chemical depletion of phagocytic cells**

1992 Phagocyte Depletion Using Clodronate Liposomes was performed as previously  
1993 described <sup>172</sup>. Briefly, female mosquitoes were injected intrathoracically with either  
1994 101.2nL of control liposomes (LPs) or CLDs in concentration 1.25mg/ml (standard  
1995 macrophage depletion kit; Encapsula NanoSciences) using a Nanoject II injector  
1996 (Drummond Scientific). Liposome treatments were performed 24 h before mosquito  
1997 exposure to *P. berghei* IBM. The efficiency of phagocytic cell depletions using CLD  
1998 was tested by measuring the expression level of two phagocytic markers: eater and  
1999 nirmrodB2 by qPCR comparing CLD and LP injected mosquitoes, using pool of 8  
2000 mosquitoes from each batch.

### 2001 **Nitration pathway depletion**

2002 Nitration pathway depletion was performed as previously described <sup>128</sup>. Briefly, female  
2003 mosquitoes were injected intrathoracically with either 1µg of dsNOX5/HPX2 (500ng for  
2004 each) or dsGFP as a control using a Nanoject II injector (Drummond Scientific). Ds  
2005 treatments were performed 72 h before mosquito exposure to *P. berghei* IBM. The  
2006 efficiency of nitration pathway depletion was tested by measuring the expression level  
2007 of NOX5 and HPX2 by RT-qPCR between dsNOX5/HPX2 and dsGFP injected  
2008 mosquitoes prior to IBM (72h post treatment) and 24h post IBM (96h post treatment)  
2009 from pool of 8 mosquitoes from each batch.

### 2010 **Sporozoite salivary glands collection**

## MATERIALS AND METHODS

2011 The salivary glands (SG) from infected mosquitoes were collected at day 19 post  
2012 infection (10 days post ds GFP and dsAPL1C treatment), as previously reported with  
2013 minor modifications<sup>173</sup>. Briefly, mosquitoes were rinsed one time in ethanol and two  
2014 times in PBS. Then, each mosquito was placed on the glass side with PBS. Mosquito  
2015 dissection and SG collections were gently performed under the stereomicroscope  
2016 using the needles of the insulin syringes. SG were collected separately for each  
2017 mosquito. Briefly, each pair of isolated SG was placed separately and individually in  
2018 a 1.5-ml microcentrifuge tube containing 20µl of 1× PBS on ice. Then each tube was  
2019 gently crushed with a pestle for microtubes in order to release the sporozoites from  
2020 the salivary glands. The solution was homogenized by vigorous pipetting.  
2021 Sporozoites were shortly centrifuged on the 35µm filter to remove remaining salivary  
2022 glands debris. 10 µl of the solution was placed in a Kova slide. Once the parasites  
2023 were settled, presence and number of sporozoites were measured with an  
2024 epifluorescence microscope using a GFP filter.

### 2025 **Hemolymph sporozoites IFA**

2026 Mosquitoes were perfused with 4% PFA in 1xPBS straight to poly-L-lysine microscopy  
2027 chamber (Ibidi GmbH). Slides were centrifuged for 2 min at 500g and fixed for 30min  
2028 with subsequent centrifugation for 2min at 500g centrifugation in order to settle down  
2029 the sporozoites. Fixed samples were rinsed three times in PBS and blocked for one  
2030 hour in 1% BSA and 1X PBS prior to overnight incubation with anti-APL1C pAb (1:300  
2031 in 1% BSA in 1X PBS) at 4°C with 12 rpm rocking. The next day, samples were rinsed  
2032 three times in 1% BSA in 1X PBS for 10 minutes each and incubated for two h with the  
2033 secondary anti-rabbit Alexa Fluor 647 antibody (Thermofisher) diluted 1:500 in 1%  
2034 BSA and 1X PBS at room temperature with 12 rpm rocking. Then, samples were  
2035 rinsed three times in 1% BSA and 1X PBS for 10 minutes each and stained with DAPI.

## MATERIALS AND METHODS

2036 Immunostained samples were rinsed in PBS and observed under the inverted Axio  
2037 Observer fluorescence microscope (ZEISS).

### 2038 **Statistical analyses**

2039 Differences in APL1C protein signal intensity in midguts were tested using Wilcoxon–  
2040 Mann–Whitney test. All statistical differences were first tested independently within  
2041 replicates (individual p-values in Annex, **Table S1**), and if individual replicates showed  
2042 a common trend of change, then individual p-values were combined using the meta-  
2043 analytical approach of Fisher<sup>174</sup>. For qPCR analysis using the 2– $\Delta\Delta$ Ct method,  
2044 difference in deltaCt distribution across the independent biological replicates was  
2045 statistically tested using Student t-test. For western blot analysis, normalized band  
2046 intensity was compared by Student t-test. Differences in salivary glands infection  
2047 prevalence were tested using the Chi-Square test. All statistical differences were first  
2048 tested independently within replicates (individual p-values in Annex, **Table S2**), and if  
2049 individual replicates showed a common trend of change, individual p-values were  
2050 combined using the meta-analytical approach of Fisher<sup>174</sup>.

### 2051 **3' Rapid Amplification of cDNA Ends (3'RACE)**

2052 Total RNA from pool of 8 females were subjected to a first strand cDNA synthesis at  
2053 the poly(A) tail of mRNA using the adapter primer (AP) according to the manufacturer  
2054 (Invitrogen). Amplification of 3' fragments of APL1A and APL1C were performed by  
2055 PCR. Two rounds of PCR were performed: first, using the with Gene-Specific Primers  
2056 (GSPs) designed for each APL1A and APL1C and producer provided Abridged  
2057 Universal Amplification Primer (AUAP) that targets the cDNA complementary to the 3'  
2058 end of mRNA. Second PCR round was performed with a second, nested-GSP for  
2059 APL1A and APL1C (Annex **Table S7**) and AUAP primers. For each gene, GST and

## MATERIALS AND METHODS

2060 nested-GST were located before stop codon of the reference gene sequences. The  
2061 PCR amplicons were subsequently cloned using TOPO cloning system (Invitrogen)  
2062 and plasmids were collected, purified (QIAGEN) and subjected for sequencing.

### 2063 **RNA sequencing experiment and bioinformatic analysis**

#### 2064 **RNA sequencing**

2065 *Anopheles* mosquitoes were injected with dsGFP, dsAPL1A or dsAPL1C. For each  
2066 condition, a pool of 8 mosquitoes was collected at 12 h post-treatment. Total RNA  
2067 extractions from intact mosquitoes were performed on each pool using TRIzol reagent  
2068 (Invitrogen). This experiment was repeated four times, yielding to a total number of 12  
2069 samples. All RNA-seq was performed at the University of Minnesota Genomics Center  
2070 ([genomics.umn.edu](http://genomics.umn.edu)).

#### 2071 **Details for RNA sequencing experiment and analysis**

2072 Prior to the library composition, RNA from each sample was quantified using  
2073 fluorimetry (RiboGreen assay) and RNA integrity was assessed using capillary  
2074 electrophoresis (Agilent BioAnalyzer), generating an RNA Integrity Number (RIN).  
2075 After passing quality check, libraries were prepared according to TruSeq Stranded  
2076 mRNA library creation. Samples were sequenced using NextSeq 150SR High-Output  
2077 (400M reads/lane)  $\pm$ 33M reads/sample.

2078 Differentially expressed genes were identified through bioinformatic analysis,  
2079 performed according to the pipeline depicted in Annex, **Figure S6**. The quality of the  
2080 raw reads was checked with FastQC version 0.10.1<sup>175</sup> and multiqc version 1.7<sup>176</sup>. All  
2081 the sequenced samples contained more than 33M reads, which were assumed by the  
2082 sequencing run (Annex, **Figure S7**). The FastQC analysis proved overall high quality  
2083 of the analysed sequences within each sample file. Performed analysis detected

## MATERIALS AND METHODS

2084 presence of the adapter sequence among all of the sample files. Therefore, remaining  
2085 adapter sequences were trimmed with cutadapt<sup>177</sup> version 1.9.1 and remaining  
2086 dsAPL1A and dsAPL1C fragments were removed using samtools version 1.9<sup>178</sup>.  
2087 Spliced Transcripts Alignment to a Reference (STAR) alignment tool <sup>179</sup> version 2.5.0a  
2088 was used for alignment against the reference genome of *Anopheles gambiae* str.  
2089 PEST version AgamP4 (from VectorBase). Reads summarization was performed  
2090 using featureCounts version 1.6.1<sup>180</sup>. Counts data were analysed with SARTools  
2091 package<sup>181</sup>, using R version 3.6.1 (R Core Team (2018). R: A language and  
2092 environment for statistical computing. R Foundation for Statistical Computing, Vienna,  
2093 (Austria), and the Bioconductor package DESeq2 version 1.24.0<sup>182</sup> using default  
2094 parameters. Due to the batch effect caused by time difference of RNA collection  
2095 between replicate nr 1 and replicates nr 2, 3 and 4, the replicate nr 1 was excluded  
2096 from the final RNAseq analysis. A generalized linear model was applied in order to test  
2097 for inter-condition differences. For each pairwise comparison, raw p-values were  
2098 adjusted for multiple testing using the Benjamini and Hochberg procedure<sup>130</sup>. Genes  
2099 with adjusted p-values below 0.05 were considered differentially expressed.

2100 **Figures: 1, 2, 6, 20, 22, 29, S1, S5, S6 were created with BioRender.com**



## REFERENCES

### 2101 REFERENCES

2102

- 2103 1 Chandrasegaran, K., Lahondere, C., Escobar, L. E. & Vinauger, C. Linking Mosquito  
2104 Ecology, Traits, Behavior, and Disease Transmission. *Trends Parasitol* **36**, 393-403,  
2105 (2020).
- 2106 2 Rosenberg, R. & Beard, C. B. Vector-borne infections. *Emerg Infect Dis* **17**, 769-770,  
2107 (2011).
- 2108 3 WHO\_World\_Malaria\_Report\_2019. WHO\_World\_Malaria\_Report\_2019.
- 2109 4 Cox, F. E. G. History of the discovery of the malaria parasites and their vectors.  
2110 *Parasites & Vectors* **3**, 5, (2010).
- 2111 5 Chernin, E. Sir Patrick Manson: an annotated bibliography and a note on a collected  
2112 set of his writings. *Rev Infect Dis* **5**, 353-386, (1983).
- 2113 6 Hagan, P. & Chauhan, V. Ronald Ross and the problem of malaria. *Parasitology Today*  
2114 **13**, 290-295, (1997).
- 2115 7 Patterson, G. M. Looking Backward, Looking Forward: The Long, Torturous Struggle  
2116 with Mosquitoes. *Insects* **7**, (2016).
- 2117 8 Becker, N. *et al.* in *Mosquitoes and Their Control* 433-439 (Springer Berlin  
2118 Heidelberg, 2010).
- 2119 9 Rieckmann, K. H. The chequered history of malaria control: are new and better tools  
2120 the ultimate answer? *Ann Trop Med Parasitol* **100**, 647-662, (2006).
- 2121 10 Raghavendra, K., Barik, T. K., Reddy, B. P. N., Sharma, P. & Dash, A. P. Malaria vector  
2122 control: from past to future. *Parasitology Research* **108**, 757-779, (2011).
- 2123 11 Bhatt, S. *et al.* The effect of malaria control on Plasmodium falciparum in Africa  
2124 between 2000 and 2015. *Nature* **526**, 207-211, (2015).
- 2125 12 Keiser, J., Singer, B. H. & Utzinger, J. Reducing the burden of malaria in different eco-  
2126 epidemiological settings with environmental management: a systematic review. *The*  
2127 *Lancet Infectious Diseases* **5**, 695-708, (2005).
- 2128 13 Benelli, G. & Beier, J. C. Current vector control challenges in the fight against malaria.  
2129 *Acta Tropica* **174**, 91-96, (2017).
- 2130 14 Liu, N. Insecticide Resistance in Mosquitoes: Impact, Mechanisms, and Research  
2131 Directions. *Annual Review of Entomology* **60**, 537-559, (2015).
- 2132 15 Wilson, A. L. *et al.* The importance of vector control for the control and elimination of  
2133 vector-borne diseases. *PLOS Neglected Tropical Diseases* **14**, e0007831, (2020).
- 2134 16 Jayaraj, R., Megha, P. & Sreedev, P. Organochlorine pesticides, their toxic effects on  
2135 living organisms and their fate in the environment. *Interdiscip Toxicol* **9**, 90-100,  
2136 (2016).
- 2137 17 Ogunah, J. A., Lalah, J. O. & Schramm, K.-W. Malaria vector control strategies. What  
2138 is appropriate towards sustainable global eradication? *Sustainable Chemistry and*  
2139 *Pharmacy* **18**, 100339, (2020).

## REFERENCES

- 2140 18 HELI.
- 2141 19 Cohuet, A., Harris, C., Robert, V. & Fontenille, D. Evolutionary forces on Anopheles:  
2142 what makes a malaria vector? *Trends Parasitol* **26**, 130-136, (2010).
- 2143 20 Beerntsen, B. T., James, A. A. & Christensen, B. M. Genetics of mosquito vector  
2144 competence. *Microbiol Mol Biol Rev* **64**, 115-137, (2000).
- 2145 21 Powell, J. R. An evolutionary perspective on vector-borne diseases. *Frontiers in*  
2146 *Genetics* **10**, (2019).
- 2147 22 Carter, R. & Mendis, K. N. Evolutionary and historical aspects of the burden of  
2148 malaria. *Clin Microbiol Rev* **15**, 564-594, (2002).
- 2149 23 Tuteja, R. Malaria – an overview. *The FEBS Journal* **274**, 4670-4679, (2007).
- 2150 24 White, N. J. Plasmodium knowlesi: the fifth human malaria parasite. *Clin Infect Dis*  
2151 **46**, 172-173, (2008).
- 2152 25 Davidson, G. Anopheles gambiae Complex. *Nature* **196**, 907-907, (1962).
- 2153 26 Coluzzi, M., Sabatini, A., della Torre, A., Di Deco, M. A. & Petrarca, V. A Polytene  
2154 Chromosome Analysis of the *Anopheles gambiae* Species Complex.  
2155 *Science* **298**, 1415-1418, (2002).
- 2156 27 Riehle, M. M. *et al.* A cryptic subgroup of *Anopheles gambiae* is highly susceptible to  
2157 human malaria parasites. *Science* **331**, 596-598, (2011).
- 2158 28 Aly, A. S. I., Vaughan, A. M. & Kappe, S. H. I. Malaria Parasite Development in the  
2159 Mosquito and Infection of the Mammalian Host. *Annual Review of Microbiology* **63**,  
2160 195-221, (2009).
- 2161 29 Siciliano, G. & Alano, P. Enlightening the malaria parasite life cycle: bioluminescent  
2162 Plasmodium in fundamental and applied research. *Frontiers in microbiology* **6**,  
2163 (2015).
- 2164 30 Josling, G. A. & Llinás, M. Sexual development in Plasmodium parasites: knowing  
2165 when it's time to commit. *Nature Reviews Microbiology* **13**, 573-587, (2015).
- 2166 31 Vlachou, D., Schlegelmilch, T., Runn, E., Mendes, A. & Kafatos, F. C. The  
2167 developmental migration of Plasmodium in mosquitoes. *Current opinion in genetics*  
2168 *& development* **16**, 384-391, (2006).
- 2169 32 Baton, L. A. & Ranford-Cartwright, L. C. Spreading the seeds of million-murdering  
2170 death: metamorphoses of malaria in the mosquito. *Trends Parasitol* **21**, 573-580,  
2171 (2005).
- 2172 33 Shahabuddin, M. Plasmodium ookinete development in the mosquito midgut: a case  
2173 of reciprocal manipulation. *Parasitology* **116 Suppl**, S83-93, (1998).
- 2174 34 Guttery, D. S., Holder, A. A. & Tewari, R. Sexual Development in Plasmodium: Lessons  
2175 from Functional Analyses. *PLOS Pathogens* **8**, e1002404, (2012).
- 2176 35 Smith, R. C. & Barillas-Mury, C. Plasmodium oocysts: overlooked targets of mosquito  
2177 immunity. *Trends in parasitology* **32**, 979-990, (2016).
- 2178 36 Siciliano, G. *et al.* Critical Steps of Plasmodium falciparum Ookinete Maturation.  
2179 *Frontiers in microbiology* **11**, 269, (2020).

## REFERENCES

- 2180 37 White, B. J., Collins, F. H. & Besansky, N. J. Evolution of *Anopheles gambiae* in  
2181 Relation to Humans and Malaria. *Annual Review of Ecology, Evolution, and*  
2182 *Systematics* **42**, 111-132, (2011).
- 2183 38 Smith, R. C., Vega-Rodriguez, J. & Jacobs-Lorena, M. The Plasmodium bottleneck:  
2184 malaria parasite losses in the mosquito vector. *Mem Inst Oswaldo Cruz* **109**, 644-661,  
2185 (2014).
- 2186 39 Pumpuni, C. B., Demaio, J., Kent, M., Davis, J. R. & Beier, J. C. Bacterial population  
2187 dynamics in three anopheline species: the impact on Plasmodium sporogonic  
2188 development. *The American journal of tropical medicine and hygiene* **54**, 214-218,  
2189 (1996).
- 2190 40 Rodgers, F. H., Gendrin, M., Wyer, C. A. S. & Christophides, G. K. Microbiota-induced  
2191 peritrophic matrix regulates midgut homeostasis and prevents systemic infection of  
2192 malaria vector mosquitoes. *PLoS Pathog* **13**, e1006391, (2017).
- 2193 41 Jupatanakul, N., Sim, S. & Dimopoulos, G. The insect microbiome modulates vector  
2194 competence for arboviruses. *Viruses* **6**, 4294-4313, (2014).
- 2195 42 Dong, Y., Manfredini, F. & Dimopoulos, G. Implication of the mosquito midgut  
2196 microbiota in the defense against malaria parasites. *PLoS Pathog* **5**, e1000423,  
2197 (2009).
- 2198 43 Wang, S. & Jacobs-Lorena, M. Genetic approaches to interfere with malaria  
2199 transmission by vector mosquitoes. *Trends in Biotechnology* **31**, 185-193, (2013).
- 2200 44 Sinden, R. E. Plasmodium differentiation in the mosquito. *Parassitologia* **41**, 139-148,  
2201 (1999).
- 2202 45 Rosenberg, R. & Rungsiwongse, J. The Number of Sporozoites Produced by Individual  
2203 Malaria Oocysts. *The American Journal of Tropical Medicine and Hygiene* **45**, 574-  
2204 577, (1991).
- 2205 46 Hillyer, J. F., Barreau, C. & Vernick, K. D. Efficiency of salivary gland invasion by  
2206 malaria sporozoites is controlled by rapid sporozoite destruction in the mosquito  
2207 haemocoel. *International journal for parasitology* **37**, 673-681, (2007).
- 2208 47 Graumans, W., Jacobs, E., Bousema, T. & Sinnis, P. When Is a Plasmodium-Infected  
2209 Mosquito an Infectious Mosquito? *Trends in Parasitology* **36**, 705-716, (2020).
- 2210 48 Barrangou, R. & Marraffini, L. A. CRISPR-Cas systems: Prokaryotes upgrade to  
2211 adaptive immunity. *Mol Cell* **54**, 234-244, (2014).
- 2212 49 Charles A. Janeway, J. & Medzhitov, R. Innate Immune Recognition. *Annual Review of*  
2213 *Immunology* **20**, 197-216, (2002).
- 2214 50 Cooper, M. D. & Herrin, B. R. How did our complex immune system evolve? *Nature*  
2215 *Reviews Immunology* **10**, 2-3, (2010).
- 2216 51 Buchmann, K. Evolution of Innate Immunity: Clues from Invertebrates via Fish to  
2217 Mammals. *Frontiers in Immunology* **5**, (2014).
- 2218 52 Govind, S. Innate immunity in *Drosophila*: Pathogens and pathways. *Insect Sci* **15**, 29-  
2219 43, (2008).

## REFERENCES

- 2220 53 Buchon, N., Silverman, N. & Cherry, S. Immunity in *Drosophila melanogaster*--from  
2221 microbial recognition to whole-organism physiology. *Nat Rev Immunol* **14**, 796-810,  
2222 (2014).
- 2223 54 Enkhbayar, P., Kamiya, M., Osaki, M., Matsumoto, T. & Matsushima, N. Structural  
2224 principles of leucine-rich repeat (LRR) proteins. *Proteins* **54**, 394-403, (2004).
- 2225 55 Govind, S. Control of development and immunity by rel transcription factors in  
2226 *Drosophila*. *Oncogene* **18**, 6875-6887, (1999).
- 2227 56 Anderson, K. V., Jurgens, G. & Nusslein-Volhard, C. Establishment of dorsal-ventral  
2228 polarity in the *Drosophila* embryo: genetic studies on the role of the Toll gene  
2229 product. *Cell* **42**, 779-789, (1985).
- 2230 57 Valanne, S., Wang, J.-H. & Rämét, M. The *Drosophila* Toll Signaling  
2231 Pathway. *The Journal of Immunology* **186**, 649-656, (2011).
- 2232 58 Rutschmann, S., Kilinc, A. & Ferrandon, D. Cutting edge: the toll pathway is required  
2233 for resistance to gram-positive bacterial infections in *Drosophila*. *J Immunol* **168**,  
2234 1542-1546, (2002).
- 2235 59 Anderson, K. V. Toll signaling pathways in the innate immune response. *Curr Opin*  
2236 *Immunol* **12**, 13-19, (2000).
- 2237 60 Arnot, C. J., Gay, N. J. & Gangloff, M. Molecular mechanism that induces activation of  
2238 Spätzle, the ligand for the *Drosophila* Toll receptor. *The Journal of biological*  
2239 *chemistry* **285**, 19502-19509, (2010).
- 2240 61 Hu, X., Yagi, Y., Tanji, T., Zhou, S. & Ip, Y. T. Multimerization and interaction of Toll  
2241 and Spätzle in *Drosophila*. *Proceedings of the National Academy of*  
2242 *Sciences of the United States of America* **101**, 9369-9374, (2004).
- 2243 62 Ooi, J. Y., Yagi, Y., Hu, X. & Ip, Y. T. The *Drosophila* Toll-9 activates a constitutive  
2244 antimicrobial defense. *EMBO Rep* **3**, 82-87, (2002).
- 2245 63 Nakamoto, M. *et al.* Virus recognition by Toll-7 activates antiviral autophagy in  
2246 *Drosophila*. *Immunity* **36**, 658-667, (2012).
- 2247 64 Xu, Y. *et al.* Structural basis for signal transduction by the Toll/interleukin-1 receptor  
2248 domains. *Nature* **408**, 111-115, (2000).
- 2249 65 Brennan, C. A. & Anderson, K. V. *Drosophila*: The Genetics of Innate Immune  
2250 Recognition and Response. *Annual Review of Immunology* **22**, 457-483, (2004).
- 2251 66 Leulier, F. & Lemaitre, B. Toll-like receptors--taking an evolutionary approach. *Nat*  
2252 *Rev Genet* **9**, 165-178, (2008).
- 2253 67 Abreu, M. T. & Arditi, M. Innate immunity and toll-like receptors: clinical implications  
2254 of basic science research. *The Journal of Pediatrics* **144**, 421-429, (2004).
- 2255 68 Kawai, T. & Akira, S. TLR signaling. *Seminars in Immunology* **19**, 24-32, (2007).
- 2256 69 Jin, M. S. *et al.* Crystal structure of the TLR1-TLR2 heterodimer induced by binding of  
2257 a tri-acylated lipopeptide. *Cell* **130**, 1071-1082, (2007).
- 2258 70 Medzhitov, R. Toll-like receptors and innate immunity. *Nature Reviews Immunology*  
2259 **1**, 135-145, (2001).

## REFERENCES

- 2260 71 Oliveira-Nascimento, L., Massari, P. & Wetzler, L. M. The Role of TLR2 in Infection and  
2261 Immunity. *Front Immunol* **3**, 79, (2012).
- 2262 72 Lee, B. L. & Barton, G. M. Trafficking of endosomal Toll-like receptors. *Trends Cell Biol*  
2263 **24**, 360-369, (2014).
- 2264 73 Langjahr, P. *et al.* Metalloproteinase-dependent TLR2 ectodomain shedding is  
2265 involved in soluble toll-like receptor 2 (sTLR2) production. *PLoS One* **9**, e104624,  
2266 (2014).
- 2267 74 LeBouder, E. *et al.* Soluble Forms of Toll-Like Receptor (TLR)2 Capable of Modulating  
2268 TLR2 Signaling Are Present in Human Plasma and Breast Milk. *The Journal of*  
2269 *Immunology* **171**, 6680-6689, (2003).
- 2270 75 Cattaneo, C. *et al.* Analysis of Toll-Like Receptors in Human Milk: Detection of  
2271 Membrane-Bound and Soluble Forms. *J Immunol Res* **2019**, 4078671, (2019).
- 2272 76 Akira, S., Takeda, K. & Kaisho, T. Toll-like receptors: critical proteins linking innate and  
2273 acquired immunity. *Nature Immunology* **2**, 675-680, (2001).
- 2274 77 Medzhitov, R. & Janeway, C., Jr. The Toll receptor family and microbial recognition.  
2275 *Trends Microbiol* **8**, 452-456, (2000).
- 2276 78 Takeda, K. & Akira, S. TLR signaling pathways. *Seminars in Immunology* **16**, 3-9,  
2277 (2004).
- 2278 79 El-Zayat, S. R., Sibaii, H. & Mannaa, F. A. Toll-like receptors activation, signaling, and  
2279 targeting: an overview. *Bulletin of the National Research Centre* **43**, 187, (2019).
- 2280 80 Viala, J., Sansonetti, P. & Philpott, D. J. Nods and 'intracellular' innate immunity.  
2281 *Comptes Rendus Biologies* **327**, 551-555, (2004).
- 2282 81 Schroder, K. & Tschopp, J. The inflammasomes. *Cell* **140**, 821-832, (2010).
- 2283 82 Wilmanski, J. M., Petnicki-Ocwieja, T. & Kobayashi, K. S. NLR proteins: integral  
2284 members of innate immunity and mediators of inflammatory diseases. *Journal of*  
2285 *Leukocyte Biology* **83**, 13-30, (2008).
- 2286 83 Franchi, L., Warner, N., Viani, K. & Nunez, G. Function of Nod-like receptors in  
2287 microbial recognition and host defense. *Immunol Rev* **227**, 106-128, (2009).
- 2288 84 Athman, R. & Philpott, D. Innate immunity via Toll-like receptors and Nod proteins.  
2289 *Current opinion in microbiology* **7**, 25-32, (2004).
- 2290 85 Bonardi, V., Cherkis, K., Nishimura, M. T. & Dangl, J. L. A new eye on NLR proteins:  
2291 focused on clarity or diffused by complexity? *Current Opinion in Immunology* **24**, 41-  
2292 50, (2012).
- 2293 86 Padmanabhan, M., Cournoyer, P. & Dinesh-Kumar, S. The leucine-rich repeat domain  
2294 in plant innate immunity: a wealth of possibilities. *Cellular microbiology* **11**, 191-198,  
2295 (2009).
- 2296 87 Chisholm, S. T., Coaker, G., Day, B. & Staskawicz, B. J. Host-microbe interactions:  
2297 shaping the evolution of the plant immune response. *Cell* **124**, 803-814, (2006).
- 2298 88 DeYoung, B. J. & Innes, R. W. Plant NBS-LRR proteins in pathogen sensing and host  
2299 defense. *Nature immunology* **7**, 1243-1249, (2006).

## REFERENCES

- 2300 89 Takken, F. L. & Govere, A. How to build a pathogen detector: structural basis of NB-  
2301 LRR function. *Current opinion in plant biology* **15**, 375-384, (2012).
- 2302 90 Belkhadir, Y., Subramaniam, R. & Dangl, J. L. Plant disease resistance protein  
2303 signaling: NBS–LRR proteins and their partners. *Current opinion in plant biology* **7**,  
2304 391-399, (2004).
- 2305 91 Pancer, Z. *et al.* Somatic diversification of variable lymphocyte receptors in the  
2306 agnathan sea lamprey. *Nature* **430**, 174-180, (2004).
- 2307 92 Sutoh, Y. & Kasahara, M. The immune system of jawless vertebrates: insights into the  
2308 prototype of the adaptive immune system. *Immunogenetics*, (2020).
- 2309 93 Herrin, B. R. *et al.* Structure and specificity of lamprey monoclonal antibodies.  
2310 *Proceedings of the National Academy of Sciences* **105**, 2040-2045, (2008).
- 2311 94 Alder, M. N. *et al.* Diversity and function of adaptive immune receptors in a jawless  
2312 vertebrate. *Science* **310**, 1970-1973, (2005).
- 2313 95 Herrin, B. R. & Cooper, M. D. Alternative adaptive immunity in jawless vertebrates.  
2314 *The Journal of Immunology* **185**, 1367-1374, (2010).
- 2315 96 Christophides, G. K. *et al.* Immunity-Related Genes and Gene Families in  
2316 *Anopheles gambiae*. *Science* **298**, 159-165, (2002).
- 2317 97 Nene, V. *et al.* Genome sequence of *Aedes aegypti*, a major arbovirus vector. *Science*  
2318 **316**, 1718-1723, (2007).
- 2319 98 Mitri, C. & Vernick, K. D. *Anopheles gambiae* pathogen susceptibility: the intersection  
2320 of genetics, immunity and ecology. *Current opinion in microbiology* **15**, 285-291,  
2321 (2012).
- 2322 99 Riehle, M. M. *et al.* Natural malaria infection in *Anopheles gambiae* is regulated by a  
2323 single genomic control region. *Science* **312**, 577-579, (2006).
- 2324 100 Menge, D. M. *et al.* Quantitative trait loci controlling refractoriness to *Plasmodium*  
2325 *falciparum* in natural *Anopheles gambiae* mosquitoes from a malaria-endemic region  
2326 in western Kenya. *Genetics* **173**, 235-241, (2006).
- 2327 101 Riehle, M. M. *et al.* A major genetic locus controlling natural *Plasmodium falciparum*  
2328 infection is shared by East and West African *Anopheles gambiae*. *Malar J* **6**, 87,  
2329 (2007).
- 2330 102 Riehle, M. M. *et al.* *Anopheles gambiae* APL1 is a family of variable LRR proteins  
2331 required for Rel1-mediated protection from the malaria parasite, *Plasmodium*  
2332 *berghei*. *PLoS One* **3**, e3672, (2008).
- 2333 103 Mitri, C. *et al.* Fine pathogen discrimination within the APL1 gene family protects  
2334 *Anopheles gambiae* against human and rodent malaria species. *PLoS Pathog* **5**,  
2335 e1000576, (2009).
- 2336 104 Mitri, C. *et al.* An evolution-based screen for genetic differentiation between  
2337 *Anopheles* sister taxa enriches for detection of functional immune factors. *PLoS*  
2338 *pathogens* **11**, (2015).

## REFERENCES

- 2339 105 Rottschaefer, S. M. *et al.* Exceptional diversity, maintenance of polymorphism, and  
2340 recent directional selection on the APL1 malaria resistance genes of *Anopheles*  
2341 *gambiae*. *PLoS biology* **9**, (2011).
- 2342 106 Holm, I. *et al.* Diverged alleles of the *Anopheles gambiae* leucine-rich repeat gene  
2343 APL1A display distinct protective profiles against *Plasmodium falciparum*. *PLoS one* **7**,  
2344 (2012).
- 2345 107 Waterhouse, R. M., Povelones, M. & Christophides, G. K. Sequence-structure-  
2346 function relations of the mosquito leucine-rich repeat immune proteins. *BMC*  
2347 *Genomics* **11**, 531, (2010).
- 2348 108 Baxter, R. H. *et al.* A heterodimeric complex of the LRR proteins LRIM1 and APL1C  
2349 regulates complement-like immunity in *Anopheles gambiae*. *Proceedings of the*  
2350 *National Academy of Sciences of the United States of America* **107**, 16817-16822,  
2351 (2010).
- 2352 109 Povelones, M., Waterhouse, R. M., Kafatos, F. C. & Christophides, G. K. Leucine-rich  
2353 repeat protein complex activates mosquito complement in defense against  
2354 *Plasmodium* parasites. *Science* **324**, 258-261, (2009).
- 2355 110 Fraiture, M. *et al.* Two Mosquito LRR Proteins Function as Complement Control  
2356 Factors in the TEP1-Mediated Killing of *Plasmodium*. *Cell host & microbe* **5**, 273-284,  
2357 (2009).
- 2358 111 Povelones, M., Upton, L. M., Sala, K. A. & Christophides, G. K. Structure-function  
2359 analysis of the *Anopheles gambiae* LRIM1/APL1C complex and its interaction with  
2360 complement C3-like protein TEP1. *PLoS Pathog* **7**, e1002023, (2011).
- 2361 112 Christophides, G. K. *et al.* Immunity-related genes and gene families in *Anopheles*  
2362 *gambiae*. *Science* **298**, 159-165, (2002).
- 2363 113 Osta, M. A., Christophides, G. K. & Kafatos, F. C. Effects of mosquito genes on  
2364 *Plasmodium* development. *Science* **303**, 2030-2032, (2004).
- 2365 114 Moita, L. F. *et al.* In vivo identification of novel regulators and conserved pathways of  
2366 phagocytosis in *A. gambiae*. *Immunity* **23**, 65-73, (2005).
- 2367 115 Blandin, S. *et al.* Complement-like protein TEP1 is a determinant of vectorial capacity  
2368 in the malaria vector *Anopheles gambiae*. *Cell* **116**, 661-670, (2004).
- 2369 116 Baxter, R. H. *et al.* Structural basis for conserved complement factor-like function in  
2370 the antimalarial protein TEP1. *Proceedings of the National Academy of Sciences* **104**,  
2371 11615-11620, (2007).
- 2372 117 Blandin, S. A., Marois, E. & Levashina, E. A. Antimalarial responses in *Anopheles*  
2373 *gambiae*: from a complement-like protein to a complement-like pathway. *Cell host &*  
2374 *microbe* **3**, 364-374, (2008).
- 2375 118 Ricklin, D., Hajishengallis, G., Yang, K. & Lambris, J. D. Complement: a key system for  
2376 immune surveillance and homeostasis. *Nat Immunol* **11**, 785-797, (2010).
- 2377 119 Le, B. V., Williams, M., Logarajah, S. & Baxter, R. H. Molecular basis for genetic  
2378 resistance of *Anopheles gambiae* to *Plasmodium*: structural analysis of TEP1  
2379 susceptible and resistant alleles. *PLoS Pathog* **8**, e1002958, (2012).

## REFERENCES

- 2380 120 Levashina, E. A. *et al.* Conserved role of a complement-like protein in phagocytosis  
2381 revealed by dsRNA knockout in cultured cells of the mosquito, *Anopheles gambiae*.  
2382 *Cell* **104**, 709-718, (2001).
- 2383 121 Blandin, S. A. *et al.* Dissecting the genetic basis of resistance to malaria parasites in  
2384 *Anopheles gambiae*. *Science* **326**, 147-150, (2009).
- 2385 122 Reyes Ruiz, V. M. *et al.* Stimulation of a protease targeting the LRIM1/APL1C complex  
2386 reveals specificity in complement-like pathway activation in *Anopheles gambiae*. *PLoS*  
2387 *one* **14**, (2019).
- 2388 123 Reddy, J. T. & Locke, M. The size limited penetration of gold particles through insect  
2389 basal laminae. *Journal of Insect Physiology* **36**, 397-407, (1990).
- 2390 124 Vernick, K. D., Fujioka, H. & Aikawa, M. *Plasmodium gallinaceum*: a novel morphology  
2391 of malaria ookinetes in the midgut of the mosquito vector. *Exp Parasitol* **91**, 362-366,  
2392 (1999).
- 2393 125 Ukegbu, C. V. *et al.* *Plasmodium berghei* P47 is essential for ookinete protection from  
2394 the *Anopheles gambiae* complement-like response. *Scientific reports* **7**, 6026, (2017).
- 2395 126 Volohonsky, G. *et al.* Kinetics of *Plasmodium* midgut invasion in  
2396 *Anopheles* mosquitoes. *bioRxiv*, 2020.2002.2006.937276, (2020).
- 2397 127 Kwon, H. & Smith, R. C. Chemical depletion of phagocytic immune cells in  
2398 *Anopheles gambiae* reveals dual roles of mosquito hemocytes in anti-  
2399 *Plasmodium* immunity. *Proceedings of the National Academy of Sciences*  
2400 **116**, 14119-14128, (2019).
- 2401 128 Castillo, J. C., Ferreira, A. B. B., Trisnadi, N. & Barillas-Mury, C. Activation of mosquito  
2402 complement antiplasmodial response requires cellular immunity. *Science*  
2403 *immunology* **2**, (2017).
- 2404 129 Blandin, S. *et al.* Complement-like protein TEP1 is a determinant of vectorial capacity  
2405 in the malaria vector *Anopheles gambiae*. *Cell* **116**, 661-670, (2004).
- 2406 130 Benjamini, Y. & Hochberg, Y. Controlling the false discovery rate: A practical and  
2407 powerful approach to multiple testing. *JR Statist. Soc. B*, **57**, 289–300. *Find this article*  
2408 *online*, (1995).
- 2409 131 Blandin, S. & Levashina, E. A. Thioester-containing proteins and insect immunity.  
2410 *Molecular Immunology* **40**, 903-908, (2004).
- 2411 132 Povelones, M. *et al.* The CLIP-domain serine protease homolog SPCLIP1 regulates  
2412 complement recruitment to microbial surfaces in the malaria mosquito *Anopheles*  
2413 *gambiae*. *PLoS pathogens* **9**, (2013).
- 2414 133 Veillard, F., Troxler, L. & Reichhart, J.-M. *Drosophila melanogaster* clip-domain serine  
2415 proteases: Structure, function and regulation. *Biochimie* **122**, 255-269, (2016).
- 2416 134 Kanost, M. R. & Jiang, H. Clip-domain serine proteases as immune factors in insect  
2417 hemolymph. *Current opinion in insect science* **11**, 47-55, (2015).
- 2418 135 Yassine, H., Kamareddine, L., Chamat, S., Christophides, G. K. & Osta, M. A. A serine  
2419 protease homolog negatively regulates TEP1 consumption in systemic infections of  
2420 the malaria vector *Anopheles gambiae*. *J Innate Immun* **6**, 806-818, (2014).



## REFERENCES

- 2421 136 Abraham, E. G. *et al.* An immune-responsive serpin, SRPN6, mediates mosquito  
2422 defense against malaria parasites. *Proceedings of the National Academy of Sciences*  
2423 *of the United States of America* **102**, 16327-16332, (2005).
- 2424 137 Pinto, S. B., Kafatos, F. C. & Michel, K. The parasite invasion marker SRPN6 reduces  
2425 sporozoite numbers in salivary glands of *Anopheles gambiae*. *Cell Microbiol* **10**, 891-  
2426 898, (2008).
- 2427 138 Meister, S. *et al.* Immune signaling pathways regulating bacterial and malaria  
2428 parasite infection of the mosquito *Anopheles gambiae*. *Proceedings of the National*  
2429 *Academy of Sciences of the United States of America* **102**, 11420-11425, (2005).
- 2430 139 Dong, Y., Taylor, H. E. & Dimopoulos, G. AgDscam, a hypervariable immunoglobulin  
2431 domain-containing receptor of the *Anopheles gambiae* innate immune system. *PLoS*  
2432 *Biol* **4**, e229, (2006).
- 2433 140 Oliveira Gde, A., Lieberman, J. & Barillas-Mury, C. Epithelial nitration by a  
2434 peroxidase/NOX5 system mediates mosquito antiplasmodial immunity. *Science* **335**,  
2435 856-859, (2012).
- 2436 141 Kajla, M. *et al.* Silencing of *Anopheles stephensi* Heme Peroxidase HPX15 Activates  
2437 Diverse Immune Pathways to Regulate the Growth of Midgut Bacteria. *Frontiers in*  
2438 *microbiology* **7**, 1351, (2016).
- 2439 142 Salzet, M., Deloffre, L., Breton, C., Vieau, D. & Schoofs, L. The angiotensin system  
2440 elements in invertebrates. *Brain Res Brain Res Rev* **36**, 35-45, (2001).
- 2441 143 Abu Hasan, Z. *et al.* The toxicity of angiotensin converting enzyme inhibitors to larvae  
2442 of the disease vectors *Aedes aegypti* and *Anopheles gambiae*. *Scientific reports* **7**,  
2443 45409, (2017).
- 2444 144 Burnham, S., Smith, J. A., Lee, A. J., Isaac, R. E. & Shirras, A. D. The angiotensin-  
2445 converting enzyme (ACE) gene family of *Anopheles gambiae*. *BMC Genomics* **6**, 172,  
2446 (2005).
- 2447 145 Ekbote, U., Looker, M. & Isaac, R. E. ACE inhibitors reduce fecundity in the mosquito,  
2448 *Anopheles stephensi*. *Comp Biochem Physiol B Biochem Mol Biol* **134**, 593-598,  
2449 (2003).
- 2450 146 Midega, J. *et al.* Discovery and characterization of two Nimrod superfamily members  
2451 in *Anopheles gambiae*. *Pathog Glob Health* **107**, 463-474, (2013).
- 2452 147 Bi, W. J. *et al.* Scavenger receptor B protects shrimp from bacteria by enhancing  
2453 phagocytosis and regulating expression of antimicrobial peptides. *Developmental and*  
2454 *comparative immunology* **51**, 10-21, (2015).
- 2455 148 Zaidman-Remy, A. *et al.* The *Drosophila* amidase PGRP-LB modulates the immune  
2456 response to bacterial infection. *Immunity* **24**, 463-473, (2006).
- 2457 149 Kurata, S. Peptidoglycan recognition proteins in *Drosophila* immunity. *Dev Comp*  
2458 *Immunol* **42**, 36-41, (2014).
- 2459 150 Neyen, C., Poidevin, M., Roussel, A. & Lemaitre, B. Tissue- and ligand-specific sensing  
2460 of gram-negative infection in *drosophila* by PGRP-LC isoforms and PGRP-LE. *J*  
2461 *Immunol* **189**, 1886-1897, (2012).

## REFERENCES

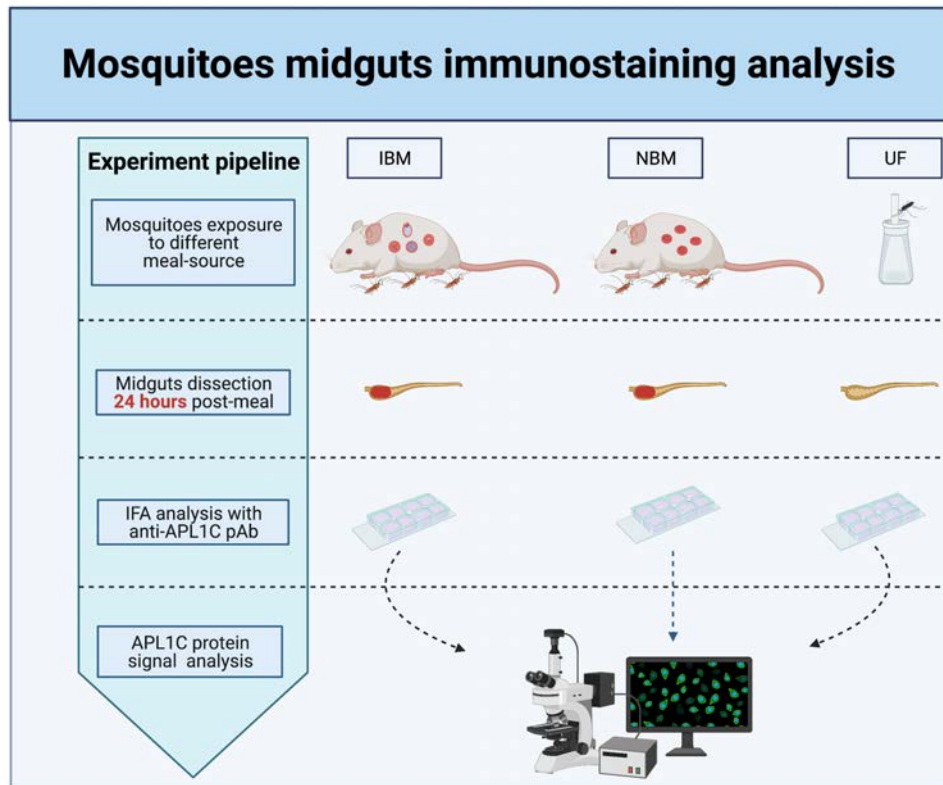
- 2462 151 Schnitger, A. K., Yassine, H., Kafatos, F. C. & Osta, M. A. Two C-type lectins cooperate  
2463 to defend *Anopheles gambiae* against Gram-negative bacteria. *J Biol Chem* **284**,  
2464 17616-17624, (2009).
- 2465 152 Bishnoi, R. *et al.* Solution structure, glycan specificity and of phenol oxidase inhibitory  
2466 activity of *Anopheles* C-type lectins CTL4 and CTLMA2. *Scientific reports* **9**, 15191,  
2467 (2019).
- 2468 153 Matsushita, M. The Complement System of Agnathans. *Front Immunol* **9**, 1405,  
2469 (2018).
- 2470 154 Castillo, J. C., Robertson, A. E. & Strand, M. R. Characterization of hemocytes from  
2471 the mosquitoes *Anopheles gambiae* and *Aedes aegypti*. *Insect Biochem Mol Biol* **36**,  
2472 891-903, (2006).
- 2473 155 Hillyer, J. F. & Strand, M. R. Mosquito hemocyte-mediated immune responses.  
2474 *Current opinion in insect science* **3**, 14-21, (2014).
- 2475 156 Luckhart, S., Vodovotz, Y., Cui, L. & Rosenberg, R. The mosquito *Anopheles stephensi*  
2476 limits malaria parasite development with inducible synthesis of nitric oxide.  
2477 *Proceedings of the National Academy of Sciences of the United States of America* **95**,  
2478 5700-5705, (1998).
- 2479 157 Kumar, S. & Barillas-Mury, C. Ookinete-induced midgut peroxidases detonate the  
2480 time bomb in anopheline mosquitoes. *Insect biochemistry and molecular biology* **35**,  
2481 721-727, (2005).
- 2482 158 Pompon, J. & Levashina, E. A. A New Role of the Mosquito Complement-like Cascade  
2483 in Male Fertility in *Anopheles gambiae*. *PLoS Biol* **13**, e1002255, (2015).
- 2484 159 Okuda, K. *et al.* Cell death and regeneration in the midgut of the mosquito, *Culex*  
2485 *quinquefasciatus*. *J Insect Physiol* **53**, 1307-1315, (2007).
- 2486 160 Viljakainen, L. Evolutionary genetics of insect innate immunity. *Brief Funct Genomics*  
2487 **14**, 407-412, (2015).
- 2488 161 Wang, X., Zhang, Y., Zhang, R. & Zhang, J. The diversity of pattern recognition  
2489 receptors (PRRs) involved with insect defense against pathogens. *Current opinion in*  
2490 *insect science* **33**, 105-110, (2019).
- 2491 162 Ghosh, J. *et al.* Invertebrate immune diversity. *Developmental and comparative*  
2492 *immunology* **35**, 959-974, (2011).
- 2493 163 Mitri, C. *et al.* Gene copy number and function of the APL1 immune factor changed  
2494 during *Anopheles* evolution. *Parasites & vectors* **13**, 18, (2020).
- 2495 164 Redmond, S. N. *et al.* Association mapping by pooled sequencing identifies TOLL 11  
2496 as a protective factor against *Plasmodium falciparum* in *Anopheles gambiae*. *BMC*  
2497 *genomics* **16**, 779, (2015).
- 2498 165 Ishino, T., Orito, Y., Chinzei, Y. & Yuda, M. A calcium-dependent protein kinase  
2499 regulates *Plasmodium* ookinete access to the midgut epithelial cell. *Mol Microbiol* **59**,  
2500 1175-1184, (2006).
- 2501 166 Bhattacharyya, M. K. & Kumar, N. Effect of xanthurenic acid on infectivity of  
2502 *Plasmodium falciparum* to *Anopheles stephensi*. *Int J Parasitol* **31**, 1129-1133, (2001).

## REFERENCES

- 2503 167 Schindelin, J. *et al.* Fiji: an open-source platform for biological-image analysis. *Nat*  
2504 *Methods* **9**, 676-682, (2012).
- 2505 168 Blandin, S. *et al.* Reverse genetics in the mosquito *Anopheles gambiae*: targeted  
2506 disruption of the Defensin gene. *EMBO Rep* **3**, 852-856, (2002).
- 2507 169 Simonetti, A. B., Billingsley, P. F., Winger, L. A. & Sinden, R. E. Kinetics of expression  
2508 of two major *Plasmodium berghei* antigens in the mosquito vector, *Anopheles*  
2509 *stephensi*. *J Eukaryot Microbiol* **40**, 569-576, (1993).
- 2510 170 Livak, K. J. & Schmittgen, T. D. Analysis of relative gene expression data using real-  
2511 time quantitative PCR and the 2-  $\Delta\Delta$ CT method. *methods* **25**, 402-408, (2001).
- 2512 171 Gendrin, M. *et al.* Antibiotics in ingested human blood affect the mosquito  
2513 microbiota and capacity to transmit malaria. *Nat Commun* **6**, 5921, (2015).
- 2514 172 Kwon, H. & Smith, R. C. Chemical depletion of phagocytic immune cells in *Anopheles*  
2515 *gambiae* reveals dual roles of mosquito hemocytes in anti-*Plasmodium* immunity.  
2516 *Proc Natl Acad Sci U S A* **116**, 14119-14128, (2019).
- 2517 173 Tavares, J., Formaglio, P., Medvinsky, A., Ménard, R. & Amino, R. in *Malaria* 401-  
2518 410 (Springer, 2012).
- 2519 174 Fisher, R. A. in *Mathematical Proceedings of the Cambridge Philosophical Society*.  
2520 700-725 (Cambridge University Press).
- 2521 175 Andrews, S. (Babraham Bioinformatics, Babraham Institute, Cambridge, United  
2522 Kingdom, 2010).
- 2523 176 Ewels, P., Magnusson, M., Lundin, S. & Kaller, M. MultiQC: summarize analysis results  
2524 for multiple tools and samples in a single report. *Bioinformatics* **32**, 3047-3048,  
2525 (2016).
- 2526 177 Martin, M. Cutadapt removes adapter sequences from high-throughput sequencing  
2527 reads. *2011* **17**, 3, (2011).
- 2528 178 Li, H. *et al.* The Sequence Alignment/Map format and SAMtools. *Bioinformatics* **25**,  
2529 2078-2079, (2009).
- 2530 179 Dobin, A. *et al.* STAR: ultrafast universal RNA-seq aligner. *Bioinformatics* **29**, 15-21,  
2531 (2013).
- 2532 180 Liao, Y., Smyth, G. K. & Shi, W. featureCounts: an efficient general purpose program  
2533 for assigning sequence reads to genomic features. *Bioinformatics* **30**, 923-930,  
2534 (2014).
- 2535 181 Varet, H., Brillet-Gueguen, L., Coppee, J. Y. & Dillies, M. A. SARTools: A DESeq2- and  
2536 EdgeR-Based R Pipeline for Comprehensive Differential Analysis of RNA-Seq Data.  
2537 *PLoS One* **11**, e0157022, (2016).
- 2538 182 Love, M. I., Huber, W. & Anders, S. Moderated estimation of fold change and  
2539 dispersion for RNA-seq data with DESeq2. *Genome Biol* **15**, 550, (2014).
- 2540

2541 ANNEX

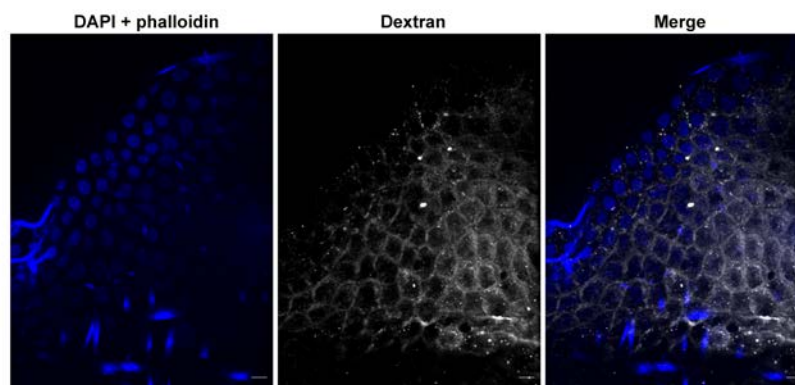
2542



2543

2544 **Figure S1. Workflow of IFA experiment.** Mosquito midguts were dissected 24 h after  
 2545 *Plasmodium* infected bloodmeal (IBM), non-infected bloodmeal (NBM) or unfed (UF) and  
 2546 immunoassayed using anti-APL1C pAb. The signal corresponding to APL1C protein  
 2547 presence on the midguts was captured by laser scanning microscope (LSM), quantified by  
 2548 ImageJ v1.52p and the signal intensity from each midgut was compared between the tested  
 2549 conditions by Mann-Whitney test.

2550

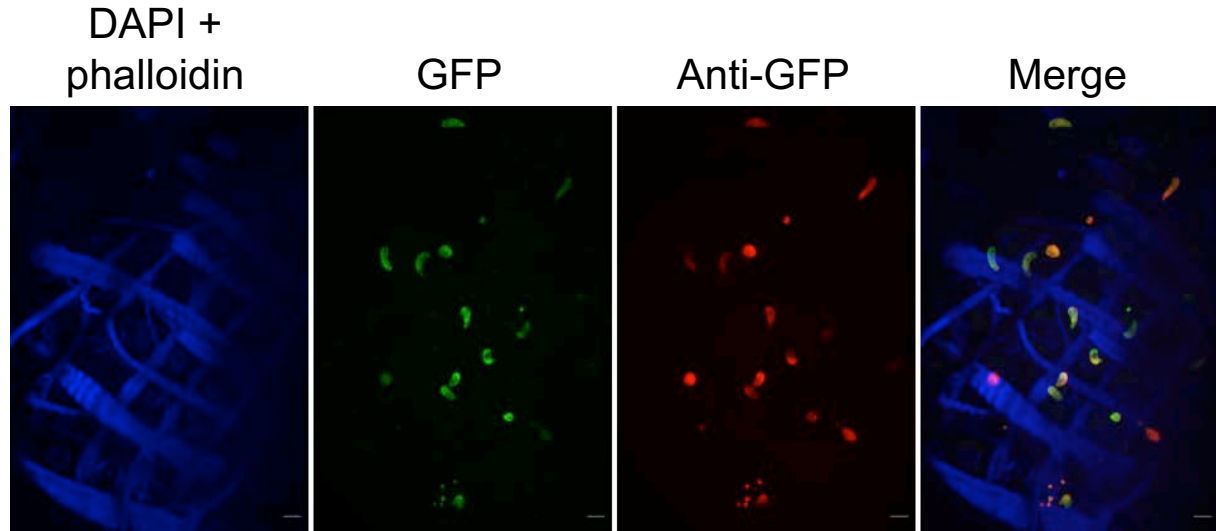


2551

ANNEX

2552 **Figure S2. Mosquito hemolymph-injected dextran diffuses on mosquito midguts in a**  
 2553 **way resembling APL1C protein deposition.** Immunostaining analysis of the midguts from  
 2554 dextran-injected mosquitoes shows that dextran diffuses through basal lamina and is deposited  
 2555 on the tissue. The scale bar is 10µm.

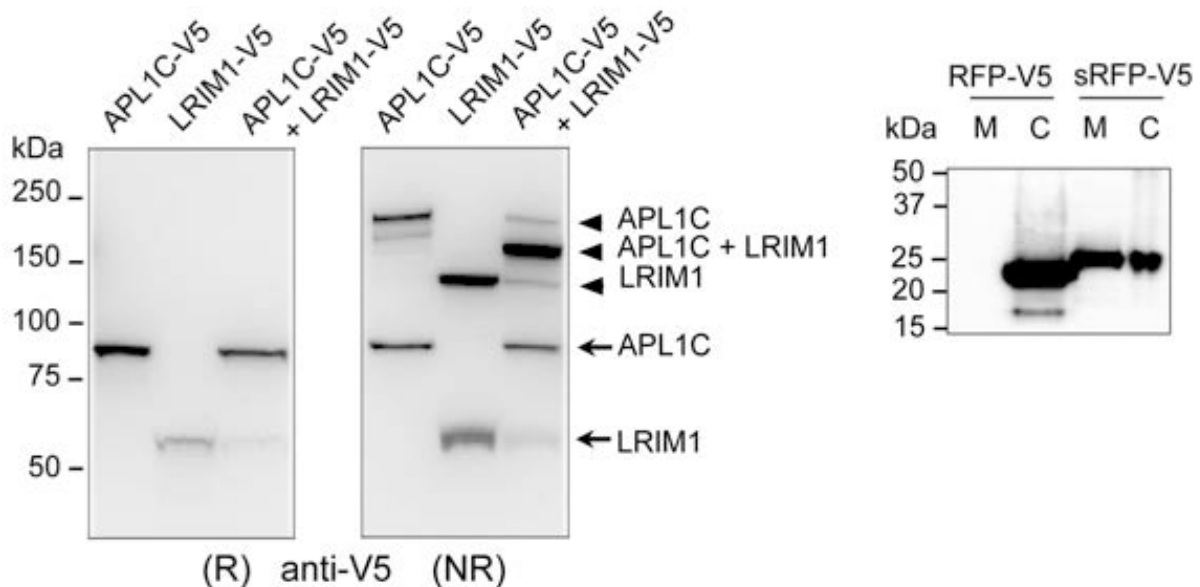
2556



2557

2558 **Figure S3. Tissue permeabilization enables antibody access to all the parasites present**  
 2559 **in mosquito midguts.** IFA of *P. berghei*-infected mosquito midguts collected 24 h post-IBM  
 2560 and immunostained with anti-GFP conjugated antibody. GFP-expressing ookinetes (green)  
 2561 are also associated with anti-GFP protein (red). Nuclei and actin were stained with DAPI and  
 2562 Phalloidin (blue). The scale bar is 10 µm.

2563



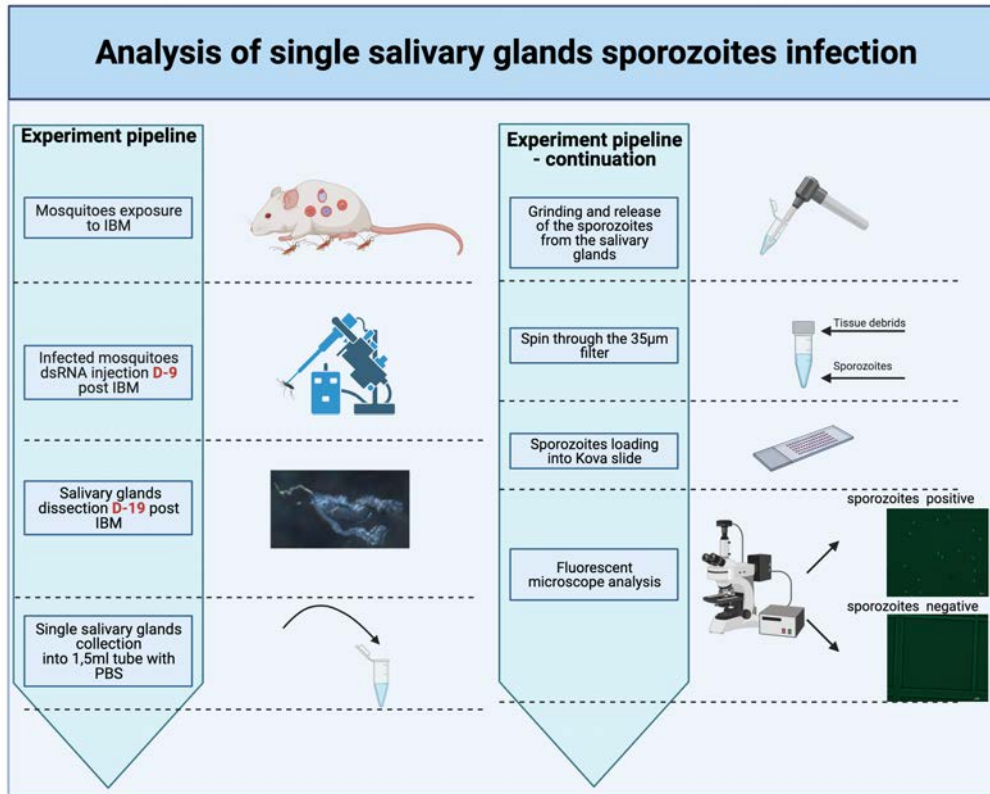
2564

2565 **Figure S4. V5-tagged protein constructs are secreted from hemocyte-like cells. A.**  
 2566 Immunoblot analysis using a mouse anti-V5 mAb and culture medium of the 4a3A hemocyte-  
 2567 like cell line transfected with plasmids encoding V5-tagged APL1C and LRIM1 under reducing  
 2568 (R) and non-reducing (NR) conditions. Estimated sizes of monomeric APL1C and LRIM1 forms  
 2569 including V5-tag are: 88 kDa (APL1C) and 60 kDa (LRIM1), respectively. The results show

## ANNEX

2570 that both V5-APL1C and V5-LRIM1 are secreted in the medium. **B.** Immunoblot analysis of  
 2571 cells (C) and culture medium (M) of the 4a3A hemocyte-like cell line transfected with plasmids  
 2572 encoding V5-tagged RFP (V5-RFP) and V5-tagged RFP fused with the signal sequence from  
 2573 APL1C (sRFP). Immunoblot was probed with a mouse anti-V5 mAb.

2574

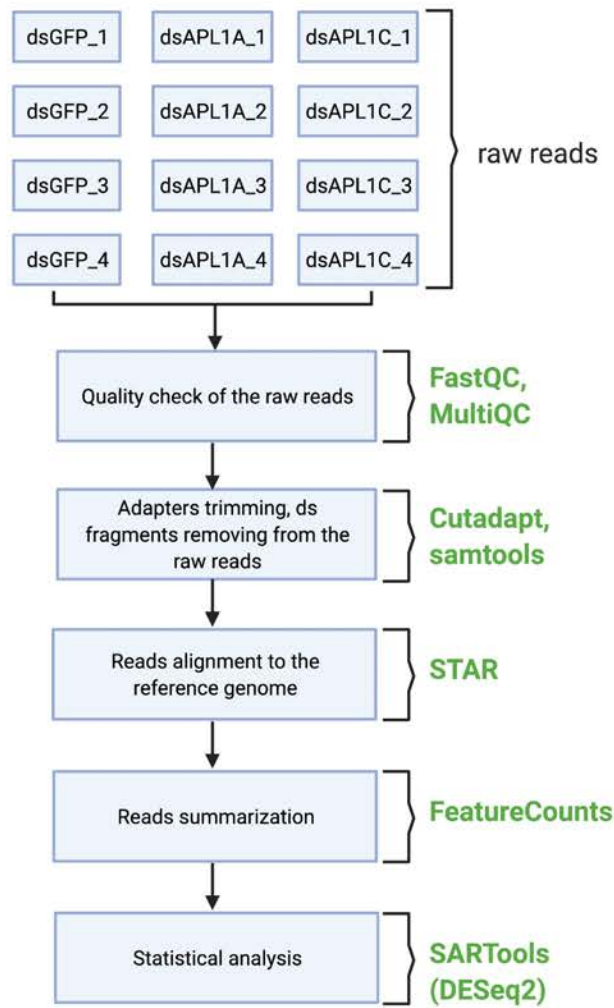


2575

2576 **Figure S5. Workflow of salivary glands infection prevalence and intensity analysis after**  
 2577 **dsRNA treatment.** Mosquitoes were injected 9 d post IBM with dsRNAs (dsAPL1C, dsSTEP1,  
 2578 dsSTEP3 or co-injected with dsSTEP1/dsSTEP3). DsGFP-injected mosquitoes served as a  
 2579 control. In the case of double treatment (dsSTEP1/dsSTEP3), a double dose of dsGFP was also  
 2580 injected in the control group. Salivary glands were dissected 19 d post IBM (10 d post dsRNA  
 2581 treatment) and collected separately in 1,5 ml tube with 20 µl of PBS. Sporozoites were released  
 2582 from the salivary gland tissues by grinding and filtering in order to get rid of tissue debris.  
 2583 The collected sporozoites containing flow-through were subsequently loaded into a Kova slide  
 2584 and screened for measuring sporozoites infection prevalence and intensity.

2585

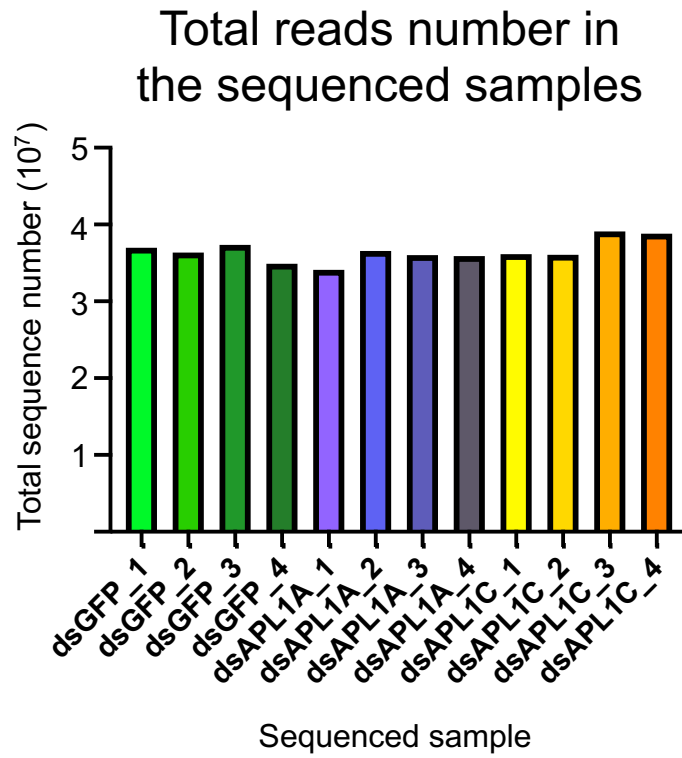
# ANNEX



2586

2587 **Figure S6. RNAseq bioinformatic analysis pipeline that identified differentially**  
2588 **expressed genes in APL1A and APL1C depleted mosquitoes.** Each step of the analysis is  
2589 presented in blue bracket, and the bioinformatic tools used for each step are indicated in green  
2590 font.

2591



2592

2593 **Figure S7. Total sequence number reported in the RNAseq analysis for all the studied**  
2594 **samples. The results show that all samples displayed a close total number of reads (30-**  
2595 **40 M reads).**



ANNEX

2596

Experiment #	Figure # <sup>a</sup>	Condition	Number of mosquitoes	Tested conditions	Median values <sup>b</sup>			individual p-value <sup>c</sup>	Fisher combined prob. <sup>d</sup>
1	7B	UF	4	-	-	-	-	-	-
2		UF	5	-	-	-	-	-	-
3		UF	4	-	-	-	-	-	-
4		UF	10	-	-	-	-	-	-
1		NBM	6	NBM vs UF	1,11E+08	vs	1,64E+07	0,0095	< 0,0001
2		NBM	4	NBM vs UF	1,20E+08	vs	3,28E+05	0,0159	
3		NBM	6	NBM vs UF	3,48E+08	vs	6,56E+06	0,0095	
4		NBM	10	NBM vs UF	1,78E+08	vs	3,03E+06	<0,0001	
3		IBM	5	IBM vs UF	6,06E+07	vs	6,56E+06	0,0317	< 0,0001
4		IBM	11	IBM vs UF	3,88E+08	vs	3,03E+06	<0,0001	
-		-	-	-	NBM vs IBM	3,48E+08	vs	6,06E+07	0,1775
-	-	-	-	NBM vs IBM	1,78E+08	vs	3,88E+08	0,5573	
5	14C	LP	8	-	-	-	-	-	-
6		LP	8	-	-	-	-	-	-
7		LP	8	-	-	-	-	-	-
5		CLD	6	CLD vs LP	9,37E+06	vs	3,52E+07	0,0813	0,0087
6		CLD	10	CLD vs LP	1,42E+07	vs	1,19E+08	0,0676	
7		CLD	10	CLD vs LP	1,80E+07	vs	1,04E+08	0,0343	
8		dsGFP	6	-	-	-	-	-	
9	dsGFP	6	-	-	-	-	-	-	
10	dsGFP	8	-	-	-	-	-	-	
8	15C	dsNOX5/HPX2	6	dsNOX5/HPX2 vs dsGFP	2,82E+07	vs	1,98E+08	0,0152	< 0,0001
9		dsNOX5/HPX2	7	dsNOX5/HPX2 vs dsGFP	3,25E+06	vs	7,05E+07	0,0012	
10		dsNOX5/HPX2	8	dsNOX5/HPX2 vs dsGFP	3,70E+07	vs	1,46E+08	0,0379	

2597

2598

2599 <sup>a</sup> The corresponding figure numbers associated with the data.

2600 <sup>b</sup> Median values of tested conditions in each replicate.

2601 <sup>c</sup> Statistical differences for APL1C signal intensity on the midguts were tested in each replicate using the Wilcoxon–Mann–Whitney test.

2603 <sup>d</sup> This p-value was calculated using the method of R.A. Fisher to combine independent tests of significance<sup>174</sup>. Notation NA (Not Applicable) indicates that the direction of phenotypic effect among experimental replicates was not consistent and therefore replicates were not statistically combined.

2606 **Table S1. Statistical analysis of APL1C signal on mosquito midguts.** Summary data for  
 2607 all experimental replicates testing the APL1C protein signal as compared to control treatment.  
 2608 Statistically significant differences are indicated by green shading. Rows indicate the treatment  
 2609 for a given replicate, the experiment number, the number of mosquitoes used for the statistical  
 2610 analysis. Individual p-values were calculated for each replicate by statistical comparison  
 2611 between the corresponding control experiment shown in the column “Experiment #” and  
 2612 indicated in "Tested conditions" column. If the different replicates of a tested condition were  
 2613 consistent (display the same phenotypic direction), then the individual p-values were combined  
 2614 by Fisher's method “Fisher combined prob”. If the replicate phenotypes were not consistent,  
 2615 the individual p-values are shown but combining of p-values is not justified.

2616

## ANNEX

Experiment #	Figure # <sup>a</sup>	Injected dsRNA	Number of mosquitoes	Salivary glands infection prevalence (%)	Infection prevalence individual p-value <sup>b</sup>	Infection prevalence Fisher combined prob. <sup>c</sup>	Salivary glands infection intensity (median)	Infection intensity individual p-value <sup>d</sup>	Infection intensity - Fisher combined prob. <sup>c</sup>
1	17C	dsGFP	30	70	-	-	-	-	-
2		dsGFP	20	25	-	-	6,40E+03	-	-
3		dsGFP	20	65	-	-	2,40E+03	-	-
4		dsGFP	36	47,22	-	-	2,90E+03	-	-
1		dsAPL1C	30	96,67	0,0056	< 0.0001	-	-	-
2		dsAPL1C	9	100	0,0002		4,44E+02	0,1898	NA
3		dsAPL1C	34	84,85	0,0942		4,25E+03	0,5657	
4		dsAPL1C	32	75	0,0195		1,90E+03	0,642	
5	19B	dsGFP	18	61,11	-	-	6,00E+03	-	-
6		dsGFP	36	50	-	-	6,39E+02	-	-
7		dsGFP	27	48,15	-	-	4,20E+03	-	-
8		dsGFP	40	72,5	-	-	5,10E+03	-	-
5		dsTEP1	23	73,91	0,382	0,21	2,90E+03	0,4233	0,3517485
6		dsTEP1	43	67,44	0,1158		6,00E+02	0,598	
7		dsTEP1	26	61,54	0,3276		6,50E+02	0,1402	
5		dsTEP3	21	76,19	0,3091	0,26	2,55E+03	0,1303	NA
6	dsTEP3	36	63,89	0,234	5,80E+03		0,0035		
7	dsTEP1/TEP3	20	60	0,4208	0,57	1900	0,6591	NA	
8	dsTEP1/TEP3	24	79,17	0,551		8,40E+03	0,4738		

2617

2618 <sup>a</sup> The corresponding figure numbers associated with data.

2619 <sup>b</sup>Statistical differences for salivary glands infection prevalence were tested for each replicate (target  
2620 gene vs dsGFP control) using the  $\chi^2$  test.

2621 <sup>c</sup> This p value was calculated using the method of R.A. Fisher to combine independent tests of  
2622 significance<sup>174</sup>. Notation NA (Not Applicable) indicates that the direction of phenotypic effect among  
2623 experimental replicates was not consistent and therefore replicates were not statistically combined.

2624 <sup>d</sup> Statistical differences for salivary glands infection intensity were tested for each replicate (dtarget  
2625 gene vs dsGFP control) using the Wilcoxon–Mann–Whitney test.

2626 **Table S2. Statistical analysis of *P. berghei* salivary glands infection prevalence**  
2627 **following gene silencing.** Summary data for all experimental replicates testing the effect of  
2628 a targeted gene silencing as compared to control treatment dsGFP. Statistically significant  
2629 differences are indicated by green shading. Rows indicate target gene tested (injected dsRNA)  
2630 for a given replicate, with the corresponding replicates indicated in the following row(s) .  
2631 Individual p-values were calculated for each replicate by statistical comparison between the  
2632 dsRNA-targeted gene and the corresponding dsGFP control showed in column “Experiment  
2633 #”. If the different replicates of a tested gene were consistent (display the same phenotypic  
2634 direction), then the individual p-values were combined by Fisher's method “Fisher combined

## ANNEX

2635 prob". If the replicate phenotypes were not consistent, the individual p-values are shown but  
2636 combining of p-values is not justified.

2637

dsGFP	dsAPL1C
2 745	6 652

2638

2639 **Table S3. APL1C controls sporozoite intensity in mosquito hemolymph.** APL1C depletion  
2640 results in higher number of sporozoites present in mosquito hemolymph. Infected mosquitoes  
2641 were perfused 17 d post-*P. berghei* infection (8 d post dsAPL1C or dsGFP treatment). The  
2642 number of perfused sporozoites was counted in both studied conditions and reported in the  
2643 table.

# ANNEX

Gene ID	Gene name	FoldChange	log2FoldChange	p-adj	Gene description	Gene description source
AGAP005503	<b>SDR11</b>	0,06	-4,055	3,95E-07	dehydrogenase/reductase SDR family member 11 precursor	VB
AGAP007033	APL1C	0,151	-2,726	2,90E-17	Anopheles Plasmodium-responsive Leucine-Rich Repeat 1C	VB
AGAP010058	<b>PAS1</b>	0,227	-2,139	0,0115104	neuronal PAS domain protein 1	VB
AGAP009217	CLIPB12	0,243	-2,043	0,01173398	CLIP-domain serine protease	VB
AGAP001669	CPR6	0,324	-1,625	0,04941549	cuticular protein RR-2 family 6	VB
AGAP008512	<b>8512</b>	0,332	-1,59	1,48E-05	NodB homology domain, Glycoside hydrolase/deacetylase, beta/alpha-barrel	IP
AGAP000113	<b>113</b>	0,341	-1,552	2,79E-07	Peptidase M14, carboxypeptidase A Cytosolic carboxypeptidase, N-terminal	IP
AGAP011509	<b>COEunkn</b>	0,359	-1,478	0,0087826	carboxylesterase	VB
AGAP013201	<b>3201</b>	0,363	-1,462	0,04950774	Ecdysteroid kinase-like, Protein kinase-like domain superfamily, CHK kinase-like	IP
AGAP006400	<b>APH2</b>	0,403	-1,31	0,00950344	alkaline phosphatase 2	VB
AGAP007319	<b>7319</b>	0,432	-1,212	0,00889759	N/A	N/A
AGAP010057		0,44	-1,186	0,02827811	nicotinic acetylcholine receptor, beta-2 subunit	VB
AGAP008444	CPLCG1	0,446	-1,164	0,00160066	cuticular protein CPLCG family (CPLCG1)	VB
AGAP010854	<b>SCF45</b>	0,453	-1,141	0,00952942	solute carrier family 45, member 1/2/4	VB
AGAP011564	<b>GGT</b>	0,455	-1,137	0,04950774	glucosyl/glucuronosyl transferases	VB
AGAP012648		0,46	-1,121	0,03487263	Sulfotransferase domain, P-loop containing nucleoside triphosphate hydrolase	IP
AGAP029110	<b>9110</b>	0,469	-1,093	0,02867749	C-type lectin-like/link domain superfamily	IP
AGAP003334	<b>CPLCX2</b>	0,474	-1,077	0,02778496	cuticular protein unclassified	VB
AGAP001748	<b>ChS</b>	0,476	-1,072	0,01545946	chitin synthase	VB
AGAP029383		0,477	-1,069	0,04961506	Pyruvate kinase, C-terminal domain superfamily	IP
AGAP001470	<b>1470</b>	0,486	-1,041	0,00667323	Leucine-rich repeat, typical subtype	IP
AGAP005496	LRIM12	0,487	-1,039	0,01760889	leucine-rich immune protein (Short)	VB
AGAP028512		0,504	-0,99	0,02877573	N/A	N/A
AGAP005459	CPR16	0,505	-0,985	0,028741	cuticular protein RR-1 family 16	VB
AGAP003600		0,508	-0,978	0,02279322	Elongation of very long chain fatty acids protein	UP
AGAP006480		0,509	-0,974	0,03487263	N/A	N/A
AGAP013507		0,509	-0,974	0,01960246	N/A	N/A
AGAP001969		0,51	-0,972	0,04950774	polyubiquitin	VB
AGAP009623	GAPDH	0,513	-0,964	0,0115104	glyceraldehyde 3-phosphate dehydrogenase	VB
AGAP007045	LRIM15	0,514	-0,961	0,01484981	leucine-rich immune protein (TM)	VB
AGAP013192		0,514	-0,96	0,04875578	venom allergen	VB
AGAP028028	LRIM16A	0,514	-0,959	0,01173009	Leucine-rich immune protein (TM)	UP
AGAP010814	TEP6	0,517	-0,952	0,0238129	thioester-containing protein 6	VB
AGAP003692		0,518	-0,948	0,03379828	Aminopeptidase	UP
AGAP007416		0,519	-0,945	0,02784284	SMAD domain, Dwarfin-type, SMAD/FHA domain superfamily	IP
AGAP005952		0,52	-0,943	0,01964634	N/A	N/A
AGAP001594		0,526	-0,926	7,71E-05	GNAT domain, Acyl-CoA N-acyltransferase	IP
AGAP008368	TEP14	0,527	-0,924	0,03909671	thioester-containing protein 14	VB
AGAP029054	NimB2	0,527	-0,923	0,01800364	nimrod B2	VB
AGAP000424		0,528	-0,921	0,03250892	N/A	N/A
AGAP004809	APN1	0,528	-0,92	0,00889759	alanine aminopeptidase N 1	VB
AGAP009871	CPR75	0,531	-0,913	0,01317147	cuticular protein RR-1 family 75	VB
AGAP012718		0,535	-0,903	0,00057426	proline oxidase	VB
AGAP028064	LRIM16B	0,535	-0,901	0,04181769	Leucine-rich immune protein (TM)	UP
AGAP006898	Ch16	0,537	-0,897	0,03680706	chitinase	VB
AGAP013245		0,538	-0,894	0,03922705	N/A	N/A
AGAP010385		0,545	-0,875	0,03182168	N/A	N/A
AGAP010562		0,546	-0,872	0,03494485	Amino acid/polyamine transporter I, Cationic amino acid transporter, C-terminal	IP
AGAP003060		0,55	-0,861	4,37E-05	N/A	N/A
AGAP008209	CYP6M1	0,552	-0,858	0,00645541	cytochrome P450	VB
AGAP004799		0,553	-0,856	0,0005304	N/A	N/A
AGAP002204	CYP325D1	0,554	-0,853	0,03487263	cytochrome P450	VB
AGAP001603		0,557	-0,845	0,0114621	mitochondrial 18 kDa protein	VB
AGAP003354	A5R3	0,557	-0,845	0,01317147	antigen 5 related protein 3	VB
AGAP028680		0,558	-0,842	0,02503675	Leucine-rich repeat, cysteine-containing subtype, F-box domain	IP
AGAP028612		0,563	-0,828	0,0330731	N/A	N/A
AGAP003695		0,571	-0,807	0,02075811	aminopeptidase N	VB
AGAP008091	CLIFE1	0,573	-0,804	0,00051006	CLIP-domain serine protease	VB
AGAP003775		0,575	-0,798	0,04464053	N/A	N/A
AGAP029055		0,577	-0,793	0,02215262	EGF-like domain	IP
AGAP005837	COEJHE5E	0,58	-0,786	0,01173009	carboxylesterase	VB
AGAP007717		0,581	-0,783	0,03161692	SH3 domain	IP
AGAP028167		0,581	-0,783	0,03383209	Proteinase, regulatory CLIP domain, Serine proteases, trypsin domain, Peptidase S1A, chymotrypsin family, Peptidase S1, PA clan	IP
AGAP013061		0,582	-0,781	0,02160989	Haemolymph juvenile hormone binding, Takeout superfamily	IP
AGAP001375	SRPN12	0,587	-0,768	0,01173398	serine protease inhibitor (serpin) 12	VB
AGAP029185		0,587	-0,768	0,03890513	CBM21 (carbohydrate binding type-21) domain	IP
AGAP001181		0,588	-0,767	0,01484981	N/A	N/A
AGAP001344		0,593	-0,754	0,04919363	Transcription factor, T-box p53-like transcription factor, DNA-binding	IP
AGAP013202		0,596	-0,746	0,03049047	N/A	N/A
AGAP011666		0,598	-0,742	0,01271696	N/A	N/A
AGAP005094		0,599	-0,74	0,0114621	Protein of unknown function DUF1397	IP
AGAP011911		0,606	-0,722	0,04405244	F-box protein 39	UP
AGAP000198	Ch15-5	0,607	-0,719	0,03823344	Chitinase	UP

## ANNEX

AGAP029297		0,611	-0,711	0,03524546	Major facilitator superfamily	IP
AGAP001486		0,612	-0,707	0,04822763	N/A	N/A
AGAP005450		0,621	-0,687	0,03414776	Kazal domain	IP
AGAP011749		0,621	-0,687	0,0179084	Major facilitator superfamily domain	IP
AGAP029820		0,621	-0,686	0,01484981	ATP:guanido phosphotransferase, Glutamine synthetase/guanido kinase, catalytic domain	IP
AGAP006835		0,627	-0,674	0,02237205	N/A	N/A
AGAP008059	CSP1	0,628	-0,671	0,0179084	chemosensory protein 1	VB
AGAP009415		0,629	-0,668	0,01811537	lysophosphatidate acyltransferase	VB
AGAP009982		0,63	-0,666	0,00551591	N/A	N/A
AGAP013459		0,632	-0,662	0,02129007	N/A	N/A
AGAP004936		0,638	-0,649	0,02341124	EGF-like domain	IP
AGAP007974	drm	0,638	-0,649	0,02362829	Protein drumstick	UP
AGAP011341		0,641	-0,641	0,02984576	Ets domain, Pointed domain, Sterile alpha motif/pointed domain superfamily, Winged helix DNA-binding domain superfamily	IP
AGAP007349		0,642	-0,64	0,02867749	N/A	IP
AGAP013106		0,644	-0,635	0,0478941	Niemann-Pick Type C-2	VB
AGAP003038		0,648	-0,625	0,02373187	Major facilitator, sugar transporter-like, Sugar transporter, conserved site	IP
AGAP005470		0,648	-0,625	0,01426511	N/A	IP
AGAP007786		0,648	-0,625	0,03909671	isocitrate dehydrogenase (NAD+)	VB
AGAP007939		0,649	-0,624	0,04375476	starch phosphorylase	VB
AGAP000986	CPAP3-D	0,652	-0,618	0,04997375	cuticular protein	VB
AGAP001417		0,658	-0,603	0,03653976	Brain protein 44-like protein	VB
AGAP001052		0,66	-0,598	0,04989706	ubiquitin carboxyl-terminal hydrolase 2/21	VB
AGAP001799		0,661	-0,598	0,03494485	tropomyosin 1	VB
AGAP006260		0,661	-0,597	0,0478941	Z band alternatively spliced PDZ-motif protein 66	VB
AGAP003860		0,665	-0,588	0,01650848	proline dehydrogenase	VB
AGAP008640		0,666	-0,586	0,02488032	N/A	N/A
AGAP010229		0,666	-0,586	0,03282155	beta-ureidopropionase	VB
AGAP004846	SCR9	0,667	-0,583	0,01746087	Class B Scavenger Receptor (CD36 domain).	VB
AGAP012515	ATPsynF	0,667	-0,584	0,01746087	F-type H+-transporting ATPase subunit f	VB
AGAP012643		0,668	-0,581	0,00979108	solute carrier family 7 (cationic amino acid transporter, y+ system)	VB
AGAP007611		0,669	-0,579	0,0179084	SEA domain, EGF-like domain, Cellulose/chitin-binding protein, N-terminal, DOMON domain, Galactose-binding-like domain superfamily	IP
AGAP009200		0,669	-0,579	0,03564164	collagen type IV alpha	VB
AGAP011355		0,671	-0,576	0,04483169	Domain of unknown function DB, Immunoglobulin subtype 2, Fibronectin type III	IP
AGAP029127	CSP5	0,672	-0,574	0,03681021	chemosensory protein 5	VB
AGAP001313		0,673	-0,571	0,02279322	Muscular protein 20	VB
AGAP005010		0,673	-0,572	0,0055572	Homeobox KN domain, Iroquois-class homeodomain protein	IP
AGAP004643	SCR6	0,674	-0,569	0,0114621	Class B Scavenger Receptor (CD36 domain).	VB
AGAP010145		0,674	-0,569	0,03681021	Yellow	UP
AGAP010230		0,675	-0,566	0,0112639	MAGUK p55 subfamily member 3	VB
AGAP001303	SAP1	0,676	-0,564	0,02450875	chemosensory protein 6	VB
AGAP003620		0,678	-0,56	0,01890109	N/A	N/A
AGAP006900		0,678	-0,56	0,04961506	Zinc finger, RING-type, Lon, substrate-binding domain, Tetratricopeptide-like helical domain superfamily, PUA-like superfamily	IP
AGAP005878		0,682	-0,552	0,04464053	POU domain transcription factor, class 3	VB
AGAP028061		0,682	-0,551	0,01261935	Protein SOGA	IP
AGAP004834		0,683	-0,549	0,02318625	PMP-22/EMP/MP20/Claudin superfamily	IP
AGAP008400		0,683	-0,551	0,00979108	Laminin G domain, Concanavalin A-like lectin/glucanase domain superfamily, CSPG repeat	IP
AGAP002859		0,684	-0,548	0,02847867	solute carrier family 8 (sodium/calcium exchanger)	VB
AGAP007876		0,684	-0,549	0,01327857	N/A	N/A
AGAP001582		0,688	-0,54	0,03681021	Ankyrin repeat-containing domain superfamily	IP
AGAP005043		0,688	-0,541	0,03187828	Inactive dipeptidyl peptidase 10	VB
AGAP004263	TO1	0,689	-0,538	0,00889759	takeout 1	VB
AGAP007061		0,69	-0,535	0,03487263	Leucine-rich repeat, typical subtype	IP
AGAP012447	GLURIIe	0,691	-0,534	0,03823344	ionotropic receptor GLURIIe	VB
AGAP010193	CTLGA3	0,693	-0,529	0,03441075	C-type lectin (CTL) - galactose binding	VB
AGAP010891		0,694	-0,528	0,03225499	ariadne-1	VB
AGAP029564		0,694	-0,527	0,03718961	Immunoglobulin-like domain, CD80-like, immunoglobulin C2-set	IP
AGAP007849		0,695	-0,526	0,03049047	laminin, alpha 1/2	VB
AGAP007963		0,695	-0,525	0,0282315	Sarcoplasmic calcium-binding protein	VB
AGAP009224		0,695	-0,525	0,02705207	GroES-like superfamily, Polyketide synthase, enoylreductase domain, NAD(P)-binding domain superfamily	IP
AGAP012256		0,695	-0,525	0,04528598	Polypeptide N-acetylgalactosaminyltransferase	UP
AGAP012399		0,696	-0,523	0,01850004	maltase	VB
AGAP028496		0,7	-0,514	0,02222595	Zinc finger C2HC domain-containing protein	IP
AGAP001369		0,701	-0,512	0,03282155	SH3 domain, IMD/I-BAR domain, AH/BAR domain superfamily, I-BAR domain containing protein	IP
AGAP006452		0,701	-0,513	0,03681021	Thioredoxin domain	IP
AGAP007373		0,701	-0,512	0,03379828	Tetraspanin	UP
AGAP010233		0,701	-0,512	0,025216	integrin beta 1	VB

## ANNEX

AGAP004457		0,702	-0,51	0,03718961	Sugar transporter ERD6-like 4	VB
AGAP004237		0,704	-0,505	0,03182418	nuclear lamin L1 alpha	VB
AGAP001413		0,705	-0,505	0,02373187	N/A	N/A
AGAP006179		0,706	-0,502	0,01271696	tropinin C	VB
AGAP002456		0,707	-0,5	0,03049047	P-loop containing nucleoside triphosphate hydrolase, Dynamin-type guanine nucleotide-binding (G) domain, EH domain-containing protein, N-terminal	IP
AGAP029561		0,708	-0,497	0,00889759	N/A	N/A
AGAP010502		0,709	-0,497	0,04372467	N/A	N/A
AGAP004642	NPF	0,71	-0,495	0,03374923	Neuropeptide F	UP
AGAP008516		0,71	-0,494	0,02561527	nicotinate phosphoribosyltransferase	VB
AGAP010750		0,71	-0,493	0,0375236	ryanodine receptor 2	VB
AGAP000305		0,711	-0,492	0,04961506	SPARC	VB
AGAP007629	LANB2	0,711	-0,492	0,00944157	laminin gamma 1	VB
AGAP001366		0,712	-0,49	0,02279322	Serine proteases, trypsin domain, Peptidase S1A, chymotrypsin family, Serine protease gd, N-terminal domain	IP
AGAP007791		0,713	-0,489	0,03201638	sodium/potassium-transporting ATPase subunit beta	VB
AGAP008702	GPRNPR2	0,713	-0,487	0,03161692	Putative neuropeptide receptor 2	UP
AGAP005247		0,714	-0,486	0,03282155	Protein of unknown function DUF4728	IP
AGAP003785		0,716	-0,482	0,03653976	glucose dehydrogenase (acceptor)	VB
AGAP003843		0,716	-0,482	0,04950774	cytochrome c heme-lyase	VB
AGAP010394		0,717	-0,481	0,0431993	Patched-related	UP
AGAP010938		0,718	-0,477	0,03909671	Trp repressor/replication initiator, Brinker DNA-binding domain	IP
AGAP000847		0,72	-0,473	0,03462588	CDC-like kinase	VB
AGAP008366	TEP2	0,72	-0,475	0,01847926	thioester-containing protein 2	VB
AGAP009693		0,721	-0,473	0,03618715	RNA pseudouridylation synthase domain-containing protein 2	VB
AGAP006063		0,722	-0,47	0,0179084	N/A	N/A
AGAP007556		0,724	-0,465	0,03487263	Immunoglobulin subtype 2, Fibronectin type III	IP
AGAP009569		0,724	-0,466	0,01271696	syntenin-1 isoform 1	VB
AGAP010436		0,729	-0,456	0,03049047	adenylate cyclase 2	VB
AGAP002492		0,733	-0,448	0,04748468	nuclear respiratory factor 1	VB
AGAP004596	PyK	0,735	-0,444	0,01800364	pyruvate kinase	VB
AGAP029620		0,735	-0,445	0,03078378	Thrombospondin type-1 (TSP1) repeat, Somatomedin B domain	IP
AGAP004335		0,736	-0,442	0,02117272	filamin	VB
AGAP009579		0,736	-0,442	0,04961506	dihydropyridine-sensitive L-type calcium channel	VB
AGAP007005		0,738	-0,438	0,02650679	N/A	N/A
AGAP007788		0,739	-0,437	0,02925868	E3 ubiquitin-protein ligase RNF31	VB
AGAP002513		0,743	-0,428	0,0005304	Defective in cullin neddylation protein	UP
AGAP007833		0,744	-0,426	0,02143453	FAM21/CAPZIP domain	IP
AGAP029606		0,747	-0,421	0,01261935	IQ motif, EF-hand binding site, Myosin head, motor domain, P-loop containing nucleoside triphosphate hydrolase,	IP
AGAP002010		0,751	-0,413	0,04981479	N/A	N/A
AGAP000331		0,752	-0,412	0,02741115	low density lipoprotein-related protein 2]	VB
AGAP000815	INTB	0,753	-0,409	0,02318625	integrin beta subunit	VB
AGAP001590		0,754	-0,408	0,03618715	DBB domain, Ankyrin repeat-containing domain superfamily	IP
AGAP005057		0,754	-0,408	0,01569346	Resistance to inhibitors of cholinesterase protein 3, N-terminal	IP
AGAP008914		0,757	-0,401	0,02561527	Basic-leucine zipper domain	IP
AGAP007198		0,759	-0,398	0,02535361	N/A	N/A
AGAP002898	GSTZ1	0,761	-0,393	0,03489037	glutathione S-transferase zeta class 1	VB
AGAP009638		0,761	-0,394	0,03494485	Stabilizer of axonemal microtubules 1/2	IP
AGAP000304		0,765	-0,386	0,04989706	Protein twisted gastrulation	VB
AGAP008254		0,766	-0,384	0,03379828	VWFC domain	IP
AGAP005032		0,767	-0,384	0,04950774	calnexin	VB
AGAP012305		0,773	-0,372	0,04950774	Muscle LIM protein at 84B	VB
AGAP007049		0,774	-0,369	0,01441075	N/A	N/A
AGAP006389		0,776	-0,366	0,01944884	ribonuclease kappa	VB
AGAP005400		0,78	-0,358	0,03057533	four and a half LIM domains	VB
AGAP008763		0,78	-0,359	0,01932857	MIF4G-like, type 3, Armadillo-type fold	IP
AGAP001023		0,793	-0,334	0,03681021	myofilin variant B	VB
AGAP008046		0,793	-0,335	0,03414776	Protein kinase C and casein kinase substrate in neurons 2 protein	UP
AGAP004689		0,797	-0,327	0,0282315	Plexin domain-containing protein	IP
AGAP011459		0,804	-0,315	0,03909671	Cyclic nucleotide-binding domain, Ion transport domain, RmlC-like jelly roll fold	IP
AGAP006645		0,809	-0,307	0,03374923	Leucine-rich repeat, typical subtype	IP
AGAP007474		0,823	-0,281	0,02721754	Talin-1	VB
AGAP009668		0,85	-0,234	0,03161692	lysosomal-associated membrane protein 1/2	VB

2646

2647  
2648  
2649  
2650  
2651  
2652  
2653  
2654  
2655  
2656  
2657  
2658  
2659

**Table S4. Full list of genes, which expression was downregulated upon APL1C depletion in RNAseq experiment.** Genes are enlisted according to the descending expression fold change between “dsAPL1C” and “dsGFP” (control) treatments. The columns include gene ID referred as the AGAP numbers; gene names, N/A means not annotated gene; expression foldchange and log2foldchange; adjusted p-value (p-adj); gene description including encoded protein function or predicted protein domain based on the databases: VB-VectorBase Community Annotation, IP-Interpro, UP-UniPro. Red dotted line indicates the cutoff tested by qPCR. Genes in purple treatments were tested by qPCR. For the rest of the genes within the cutoff, the specific primers were not available. APL1C was excluded from qPCR analysis. The “Gene name” column includes the abbreviation used in qPCR analysis. Gene name in black correspond to the annotated gene name. Gene name in red correspond to abbreviation of gene function, whereas the genes without annotated function are called by the four last digits of their AGAP number and are depicted in green.

## ANNEX

Gene ID	Gene name	FoldChange	log2FoldChange	p-adj	Gene description	Gene description source
AGAP006728	COEAE7G	10,483	3,39	4,20E-10	carboxylesterase	VB
AGAP029790	9790	10,297	3,364	0,0324689	Domain of unknown function DUF4806	IP
AGAP008246	GBD	2,577	1,366	0,016879105	gamma-butyrobetaine dioxygenase	VB
AGAP029411	9411	2,419	1,274	0,024111017	N/A	N/A
AGAP009381	CRBP	2,36	1,239	0,0375236	Cellular retinaldehyde-binding protein	UP
AGAP012477	GBD_2	2,32	1,214	0,013171475	gamma-butyrobetaine dioxygenase	VB
AGAP011676	UP2M	2,264	1,179	7,71E-05	uncoupling protein 2, mitochondrial	VB
AGAP003350	GTP	2,246	1,167	0,019596342	phosphoenolpyruvate carboxykinase (GTP)	VB
AGAP029623	9623	2,147	1,103	0,010492105	Sulfotransferase domain P-loop containing nucleoside triphosphate hydrolase	IP
AGAP007300	AP	2,122	1,085	0,019602457	alkaline phosphatase	VB
AGAP007959	7959	2,113	1,079	0,006046664	N/A	N/A
AGAP013642	U1	2,041	1,029	0,030548251	U1 spliceosomal RNA [Source:RFAM;Acc:RF00003]	
AGAP001423	BPBP	2,026	1,018	0,013739866	Bifunctional purine biosynthesis protein	VB
AGAP007122	TU1	2,018	1,013	0,000530401	tubulin, alpha 1	VB
AGAP012894		1,997	0,998	0,015519838	Histone H2B	UP
AGAP001352	1352	1,973	0,98	0,023186247	Haemolymph juvenile hormone binding Takeout superfamily	IP
AGAP005296	U11_1	1,971	0,979	0,01964634	U11/U12 small nuclear ribonucleoprotein 65 kDa protein	VB
AGAP007364	7364	1,956	0,968	0,009791078	Maternal protein exuperantia	IP
AGAP006241	INX2	1,954	0,967	9,44E-05	Innexin inx2	VB
AGAP000162	CBS	1,931	0,949	0,015459457	Cystathionine beta-synthase	UP
AGAP029317		1,912	0,935	0,039588096	Exoribonuclease, phosphorolytic domain 1 Ribosomal protein S5 domain 2-type fold PNPase/RNase PH domain superfamily	IP
AGAP000772	772	1,881	0,912	0,019328566	N/A	N/A
AGAP009751	ANCE2	1,868	0,901	0,006046664	angiotensin-converting enzyme 2	VB
AGAP003086		1,861	0,896	0,001831642	TMEM192 family	IP
AGAP004038	HPX8	1,858	0,893	0,003199427	heme peroxidase 8	VB
AGAP010540	540	1,852	0,889	0,010020081	N/A	N/A
AGAP005829	MAS_1	1,844	0,883	0,008897593	microtubule-associated protein, RP/EB family	VB
AGAP000180	PC	1,84	0,88	0,0112639	phosphoribosylaminoimidazole carboxylase	VB
AGAP004036	HPX7	1,838	0,878	0,02739326	heme peroxidase 7	VB
AGAP028655	8655	1,829	0,871	0,006887628	N/A	N/A
AGAP003591	LRR1	1,826	0,869	0,009441568	LRR-repeat protein 1	VB
AGAP007991		1,822	0,866	0,034317264	Triacylglycerol lipase family Lipase/vitellogenin Alpha/Beta hydrolase fold Lipase, N-terminal	IP
AGAP002387		1,82	0,864	0,004983856	acid phosphatase	VB
AGAP003087		1,816	0,861	0,009953441	TMEM192 family	IP
AGAP011145		1,811	0,856	0,033798283	ATP-dependent RNA helicase DDX56/DBP9	VB
AGAP000427		1,802	0,85	0,005953354	vitellogenin receptor	VB
AGAP008015		1,795	0,844	0,005266262	lamin	VB
AGAP009606		1,791	0,841	0,01042075	Domain of unknown function DUF3421 DM9 repeat	IP
AGAP011858		1,791	0,841	0,011510398	zinc finger HIT domain-containing protein 3	VB
AGAP001333	st	1,783	0,834	0,011462104	protein scarlet	VB
AGAP002581		1,782	0,833	0,019328566	MCM domain Nucleic acid-binding, OB-fold Mini-chromosome maintenance protein MCM, AAA-lid domain	IP
AGAP007549		1,774	0,827	0,030783776	Six-bladed beta-propeller, TolB-like Major royal jelly protein/protein yellow	IP
AGAP002532		1,765	0,82	0,004238876	Cyclin, C-terminal domain Cyclin, N-terminal	IP
AGAP005844		1,76	0,816	0,006233202	Bifunctional purine biosynthesis protein	VB
AGAP007505		1,76	0,816	0,014849805	vitellogenic carboxypeptidase-like protein	VB
AGAP011845		1,751	0,808	0,022152616	Histone H2A	UP
AGAP011713		1,734	0,794	0,016059772	ribonuclease P protein subunit POP4	VB
AGAP006404		1,733	0,793	0,018003641	N/A	N/A
AGAP012040		1,724	0,786	0,032254989	N/A	N/A

## ANNEX

AGAP012040		1,724	0,786	0,032254989	N/A	N/A
AGAP009792		1,719	0,781	0,000530401	importin subunit alpha-2	VB
AGAP001699		1,713	0,777	0,048918391	fatty-acid amide hydrolase 2	VB
AGAP012192		1,712	0,776	0,008897593	DNA mismatch repair protein MLH1	VB
AGAP002846		1,711	0,775	0,041598991	Trunk	VB
AGAP002378		1,71	0,774	0,023876565	Adenylosuccinate lyase	UP
AGAP006727	COEAE6G	1,707	0,772	0,015519838	carboxylesterase	VB
AGAP012375		1,698	0,764	0,017608886	protein aurora borealis	VB
AGAP003974		1,695	0,761	0,008897593	tRNA (cytidine32/guanosine34-2'-O)-methyltransferase	VB
AGAP003552		1,69	0,757	0,023186247	SET and MYND domain-containing protein 5	VB
AGAP000479		1,687	0,755	0,017422566	pre-rRNA-processing protein IPI3	VB
AGAP005581		1,685	0,753	0,00105879	3-hydroxyisobutyrate dehydrogenase	UP
AGAP002478		1,684	0,752	0,008897593	cyclin B	VB
AGAP005800		1,682	0,75	0,01612065	DNA replication licensing factor MCM7	VB
AGAP013023		1,664	0,735	0,03161692	Histone RNA stem-loop-binding protein SLBP1/SLBP2 Histone RNA hairpin-binding protein, RNA-binding domain	IP
AGAP011887		1,663	0,734	0,017908403	nuclear pore complex protein Nup37	VB
AGAP005260		1,661	0,732	0,011510398	dTMP kinase	VB
AGAP008290	TRYP6	1,66	0,731	0,020786941	trypsin 6	VB
AGAP006438		1,655	0,727	0,006046664	ribosomal biogenesis protein LAS1	VB
AGAP005052		1,654	0,726	0,03653976	deoxyribodipyrimidine photo-lyase	VB
AGAP011897		1,65	0,722	0,018003641	N/A	N/A
AGAP010918		1,645	0,718	0,038914815	cleavage stimulation factor subunit 2	VB
AGAP029627		1,645	0,718	0,049973746	Copper type II, ascorbate-dependent monooxygenase, N-terminal DOMON domain PHM/PNGase F domain superfamily Tyramine beta-hydroxylase/Dopamine beta-hydroxylase	IP
AGAP012686		1,644	0,717	0,049507741	Zinc finger C2H2-type	IP
AGAP001522		1,641	0,714	0,009529416	Thyroid receptor-interacting protein 13	VB
AGAP011652		1,641	0,714	0,024988431	AAA family ATPase, CDC48 subfamily protein	VB
AGAP008738		1,638	0,712	0,048387263	Eukaryotic translation initiation factor 4E binding protein	IP
AGAP029470		1,633	0,707	0,019448835	Sterile alpha motif domain K Homology domain	IP
AGAP010789		1,624	0,7	0,023352255	EB domain	IP
AGAP029584		1,623	0,698	0,023196353	N/A	N/A
AGAP029679		1,623	0,699	0,034944849	Zinc finger, AD-type Zinc finger C2H2-type	IP
AGAP006092		1,618	0,694	0,027272801	U3 small nucleolar RNA-associated protein 4	VB
AGAP011393		1,617	0,693	0,018115369	integrator complex subunit 7	VB
AGAP000650		1,613	0,689	0,015519838	ribonucleases P/MRP protein subunit RPP40	VB
AGAP003740		1,606	0,684	0,046466887	N/A	N/A
AGAP010941		1,601	0,679	0,027784962	Zinc finger C2H2-type Zinc finger protein 511	IP
AGAP003545	Osk	1,6	0,678	0,020044331	protein oskar	VB
AGAP004825		1,598	0,676	0,022564284	PapD-like superfamily Immunoglobulin-like fold ADP-ribose pyrophosphatase, mitochondrial	IP
AGAP004008		1,597	0,676	0,029845759	Protein kinase domain Serine-threonine/tyrosine-protein kinase, catalytic domain	IP
AGAP007817		1,597	0,675	0,017908403	Protein NEDD1	UP
AGAP029259		1,591	0,67	0,0112639	N/A	N/A
AGAP007145		1,584	0,663	0,017908403	N/A	N/A
AGAP007823		1,582	0,662	0,049415485	Nucleotide-binding alpha-beta plait domain superfamily Meiosis regulator and mRNA stability factor 1 OST-HTH/LOTUS domain	IP
AGAP009048		1,582	0,662	0,049897056	protein arginine N-methyltransferase 5	VB
AGAP010609		1,582	0,662	0,049507741	DnaJ (Hsp40) homolog subfamily C	VB
AGAP012154		1,581	0,661	0,017624312	solute carrier family 15 member 1	VB
AGAP008578		1,58	0,66	0,023186247	ATP-dependent RNA helicase DDX4	VB
AGAP006938	Art7	1,578	0,658	0,033798283	Protein arginine N-methyltransferase 7	UP
AGAP007775		1,578	0,658	0,03823344	ATP-dependent DNA helicase PIF1	UP
AGAP012964		1,578	0,658	0,017608886	Sarcospan	IP
AGAP004108		1,574	0,654	0,023812903	Amalgam	VB
AGAP010791		1,569	0,65	0,008897593	polyribonucleotide nucleotidyltransferase	VB
AGAP001656		1,568	0,649	0,029622655	ribosomal RNA-processing protein 7	VB
AGAP011457		1,565	0,646	0,013171475	ribose 5-phosphate isomerase A	VB
AGAP010453		1,56	0,642	0,021290069	ribosome biogenesis protein NSA1	VB
AGAP012107		1,557	0,638	0,036201432	SUMO-conjugating enzyme UBC9	VB
AGAP012216		1,557	0,639	0,006887628	sterol O-acyltransferase	VB
AGAP010406		1,554	0,636	0,035660224	origin recognition complex subunit 3	VB
AGAP007502		1,551	0,633	0,008897593	kinesin family member 11	VB
AGAP008063		1,547	0,629	0,036201432	exosome complex component CSL4	VB
AGAP001453		1,546	0,629	0,049973746	N/A	N/A
AGAP011454		1,544	0,626	0,001831642	rotatin	VB
AGAP008879		1,539	0,622	0,011462104	Protein of unknown function DUF389	IP
AGAP003883		1,537	0,62	0,000709794	peptidyl-tRNA hydrolase, PTH2 family	VB
AGAP003578		1,535	0,618	0,033798283	aldehyde dehydrogenase (NAD+)	VB
AGAP001166		1,532	0,616	0,021172723	GINS complex subunit 4	VB



## ANNEX

AGAP010722		1,529	0,612	0,028086615	RNA recognition motif domain Tudor domain Nucleotide-binding alpha-beta plait domain superfamily SNase-like, OB-fold superfamily RNA-binding domain superfamily	IP
AGAP003090		1,528	0,612	0,013462018	ubiquitin-conjugating enzyme E2 O	VB
AGAP013324		1,528	0,612	0,036267067	putative G-protein coupled receptor GPCR	VB
AGAP028639		1,526	0,609	0,014757327	N/A	N/A
AGAP004239		1,518	0,602	0,034317264	polo-like kinase 1	VB
AGAP001529		1,515	0,6	0,001102944	H/ACA ribonucleoprotein complex subunit	UP
AGAP007417		1,515	0,599	0,029157396	E3 ubiquitin-protein ligase SHPRH	VB
AGAP007635		1,511	0,595	0,014849805	60S ribosomal protein L7	VB
AGAP000719	Ahcy13	1,51	0,595	0,028478668	Adenosylhomocysteinase	UP
AGAP011534		1,51	0,594	0,023015671	Ankyrin repeat Ankyrin repeat and LEM domain-containing protein 2	IP
AGAP005149		1,498	0,583	0,006127055	H/ACA ribonucleoprotein complex subunit 3	VB
AGAP002967		1,497	0,582	0,016802001	bloom syndrome protein	VB
AGAP003890		1,492	0,577	0,024880322	U3 small nucleolar ribonucleoprotein protein IMP4	VB
AGAP010538		1,492	0,578	0,0381339	M-phase phosphoprotein	VB
AGAP003594		1,491	0,576	0,036857325	Zinc finger, AD-type Zinc finger C2H2-type Transcription factor Grauzone	IP
AGAP009825		1,491	0,576	0,018500044	TRM13/UPF0224 family, U11-48K-like CHHC zinc finger domain Zinc finger C2H2 superfamily	IP
AGAP013441		1,487	0,572	0,023186247	SS18 family	IP
AGAP004018		1,486	0,571	0,023186247	PDZ domain Harmonin	IP
AGAP008487		1,486	0,571	0,026506785	Sphingomyelin phosphodiesterase	UP
AGAP000168	GPRFZ3	1,482	0,568	0,030490471	Putative frizzled receptor 3	UP
AGAP006233		1,481	0,566	0,023015671	Phospholipase D-like domain	IP
AGAP002088		1,48	0,566	0,018596807	U3 small nucleolar RNA-associated protein 15	VB
AGAP005174		1,473	0,559	0,02827811	nucleoporin SEH1	VB
AGAP008771		1,472	0,558	0,036187152	DNA-directed RNA polymerases I and III subunit RPAC1	VB
AGAP002394		1,471	0,557	0,036810206	DNA-directed RNA polymerase III subunit RPC3	VB
AGAP002706		1,471	0,556	0,023812903	small nuclear ribonucleoprotein G	VB
AGAP028653		1,47	0,555	0,028478668	Pleckstrin homology domain	IP
AGAP002117		1,469	0,555	0,034944849	Cell death regulator Aven	IP
AGAP005629		1,469	0,555	0,03823344	N/A	N/A
AGAP009024		1,469	0,554	0,035419211	RNA polymerase I-specific transcription initiation factor RRN3	VB
AGAP027988		1,467	0,553	0,034910795	Zinc finger, AD-type Zinc finger C2H2-type	IP
AGAP008163		1,464	0,55	0,029845759	U4/U6 small nuclear ribonucleoprotein SNU13	VB
AGAP009488		1,464	0,55	0,034910795	histone transcription regulator	VB
AGAP001595	RpS29	1,46	0,546	0,010729313	40S ribosomal protein S29	VB
AGAP029794		1,459	0,545	0,031358359	Serine proteases, trypsin domain Peptidase S1A, chymotrypsin family	IP
AGAP002409		1,457	0,543	0,006127055	tRNA pseudouridine38-40 synthase	VB
AGAP002121		1,453	0,539	0,016879105	DNA-directed RNA polymerase II subunit RPB7	VB
AGAP005320		1,45	0,536	0,039878744	U3 small nucleolar RNA-interacting protein 2	VB
AGAP009336		1,45	0,536	0,022152616	RNA recognition motif domain Nucleotide-binding alpha-beta plait domain superfamily RNA-binding domain superfamily	IP
AGAP002535		1,449	0,535	0,023186247	serine/threonine-protein phosphatase PGAM5	VB
AGAP007413		1,445	0,531	0,033798283	BRCT domain Microcephalin-like	IP
AGAP000785		1,444	0,53	0,034872633	Synaptic vesicle protein	VB
AGAP002594		1,443	0,529	0,032821552	apolipoprotein D	VB
AGAP000666		1,442	0,528	0,03823344	ubiquitin carboxyl-terminal hydrolase 16/45	VB
AGAP029449		1,441	0,527	0,014751272	Helicase, C-terminal DEAD/DEAH box helicase domain Helicase superfamily 1/2, ATP-binding domain P-loop containing nucleoside triphosphate hydrolase	IP
AGAP009512		1,44	0,526	0,019885369	E3 ubiquitin-protein ligase UBR7	VB
AGAP000535		1,438	0,524	0,030035206	Acytransferase 3 Nose resistant-to-fluoxetine protein, N-terminal	IP
AGAP001380		1,434	0,52	0,023933025	translation initiation factor 6	VB
AGAP000681		1,429	0,515	0,0436892	tRNA:m4X modification enzyme	VB
AGAP000496		1,422	0,508	0,019328566	geminin	VB
AGAP006941		1,42	0,506	0,037189609	ATP-dependent RNA helicase DDX51/DBP6	VB
AGAP009442		1,419	0,505	0,049415485	cyclin-dependent kinase regulatory subunit CKS1	VB
AGAP009431	RpS21	1,415	0,501	0,00055617	40S ribosomal protein S21	UP
AGAP003205		1,411	0,497	0,036174833	Major facilitator superfamily MFS transporter superfamily	IP

## ANNEX

AGAP011383		1,41	0,495	0,037458371	rRNA-processing protein EBP2	VB
AGAP000549		1,409	0,494	0,049507741	2-oxoisovalerate dehydrogenase E2 component (dihydrolipoyl transacylase)	VB
AGAP002961		1,409	0,495	0,036810206	ribosomal RNA-processing protein 12	VB
AGAP009830		1,406	0,491	0,01612065	Chitin binding domain	IP
AGAP008292	TRYP4	1,403	0,489	0,036201432	trypsin 4	VB
AGAP001224		1,402	0,488	0,025215999	cleavage and polyadenylation specificity factor subunit 3	VB
AGAP005917		1,399	0,484	0,036201432	ATP-dependent DNA helicase Q4	VB
AGAP013400		1,399	0,485	0,014432066	Cytosolic fatty-acid binding Lipocalin/cytosolic fatty-acid binding domain Calycin Intracellular lipid binding protein	IP
AGAP029374		1,398	0,484	0,039588096	Zinc finger, AD-type Domain of unknown function DUF4806	IP
AGAP006234		1,397	0,483	0,049507741	protein SHQ1	VB
AGAP011966		1,397	0,482	0,036187152	transcription termination factor 2	VB
AGAP005922		1,392	0,477	0,044831691	DNA-directed RNA polymerase III subunit RPC5	VB
AGAP028053		1,389	0,474	0,047484681	N/A	N/A
AGAP007668	eIF3g	1,387	0,472	0,006887628	Eukaryotic translation initiation factor 3 subunit G	UP
AGAP002151		1,386	0,471	0,033160605	Regulator of chromosome condensation, RCC1 Regulator of chromosome condensation 1/beta-lactamase-inhibitor protein II	IP
AGAP008118		1,385	0,47	0,023352255	serine/threonine-protein kinase Chk1	VB
AGAP007020	TPX5	1,378	0,462	0,049615058	thioredoxin peroxidase 5	VB
AGAP000374		1,377	0,461	0,034944849	anaphase-promoting complex subunit 1	VB
AGAP003029		1,375	0,459	0,036201432	Forkhead-associated (FHA) domain Zinc finger, PHD-type SMAD/FHA domain superfamily PHD finger protein 12, MRG binding domain superfamily	IP
AGAP002337	eIF3d	1,366	0,45	0,027842842	Eukaryotic translation initiation factor 3 subunit D	UP
AGAP000416		1,364	0,448	0,030053919	hydroxyacylglutathione hydrolase	VB
AGAP002106		1,363	0,447	0,0324689	xeroderma pigmentosum group C-complementing protein	VB
AGAP002113		1,363	0,446	0,023409625	cytochrome b5	VB
AGAP002985		1,363	0,447	0,017608886	Structural maintenance of chromosomes protein 6	UP
AGAP002653		1,357	0,44	0,045285977	phenylalanyl-tRNA synthetase alpha chain	VB
AGAP003264		1,355	0,438	0,036810206	histone acetyltransferase 1	VB
AGAP007026	Ccs	1,355	0,439	0,043199303	copper chaperone for superoxide dismutase	VB
AGAP011309		1,355	0,438	0,044633666	nuclear GTP-binding protein	VB
AGAP010978		1,346	0,429	0,037106786	nuclear pore complex protein Nup62	VB
AGAP006997	Aats-ala-m	1,345	0,428	0,036201432	alanyl-tRNA synthetase, mitochondrial	VB
AGAP002404		1,343	0,425	0,049615058	cullin 5	VB
AGAP001141		1,339	0,421	0,044831691	Zinc finger C2H2-type	IP
AGAP009862		1,335	0,417	0,048227633	WD40 repeat WD40/YVTN repeat-like-containing domain superfamily	IP
AGAP010586	AIMP1	1,335	0,417	0,029255508	aminoacyl tRNA synthase complex-interacting multifunctional protein 1	VB
AGAP011677		1,335	0,417	0,020085987	Elongation factor EFG, domain V-like Transcription factor, GTP-binding domain Translation elongation factor EFTu-like, domain 2	IP
AGAP008000		1,334	0,416	0,044640528	sepiapterin reductase	VB
AGAP002968		1,324	0,405	0,048634299	thiamine pyrophosphokinase	VB
AGAP005115		1,323	0,404	0,044831691	N/A	N/A
AGAP003063		1,319	0,399	0,044831691	Rwd domain-containing protein	VB
AGAP009091		1,317	0,397	0,034625882	aromatic-L-amino-acid decarboxylase	VB
AGAP004563	ANCE9	1,313	0,393	0,049507741	angiotensin-converting enzyme 9	VB
AGAP006947		1,311	0,391	0,044831691	Cytochrome B561-related	IP
AGAP003486		1,302	0,38	0,034625882	N/A	N/A
AGAP008090		1,302	0,38	0,036201432	nuclear pore complex protein Nup205	VB
AGAP008635		1,295	0,373	0,032228725	protein SDA1	VB
AGAP005990		1,285	0,362	0,046758013	glycolipid transfer protein domain-containing protein 1	VB
AGAP008795		1,285	0,362	0,029622655	Zinc finger, AD-type Zinc finger C2H2-type	IP
AGAP009920	RpL37a	1,283	0,359	0,035903402	60S ribosomal protein L37a	VB
AGAP000826		1,282	0,358	0,044052437	cap-specific mRNA (nucleoside-2'-O-)-methyltransferase 1	VB
AGAP011173	RpL11	1,278	0,354	0,039015507	60S ribosomal protein L11	VB
AGAP009572	RpS15a-2	1,276	0,351	0,033798283	40S ribosomal protein S15a	VB
AGAP002539		1,269	0,343	0,035419211	integrator complex subunit 3	VB
AGAP007740	RpLP1	1,264	0,338	0,027842842	60S ribosomal protein LP1	VB
AGAP012219	Med1	1,253	0,326	0,03589132	mediator of RNA polymerase II transcription subunit 1	VB
AGAP002525		1,223	0,29	0,049507741	eukaryotic translation initiation factor 2A	VB

## ANNEX

2664

2665 **Table S5. Full list of genes, which expression is upregulated upon APL1C depletion in**  
 2666 **RNAseq experiment.** Genes are enlisted according to the ascending expression fold change  
 2667 between “dsAPL1C” and “dsGFP” (control) treatments. The columns include gene ID referred  
 2668 as the AGAP numbers; gene names, where N/A means not annotated gene; expression  
 2669 foldchange and log2foldchange; adjusted p-value (p-adj); gene description including encoded  
 2670 protein function or predicted protein domain based on the databases: VB- VectorBase  
 2671 Community Annotation, IP-Interpro, UP-UniPro. Genes in purple treatments were tested by  
 2672 qPCR. For the rest of the genes within the cutoff, the specific primers were not available. The  
 2673 “Gene name” column include the abbreviation used in qPCR analysis. Gene name in black  
 2674 correspond to the annotated gene name. Gene name in red correspond to abbreviation of  
 2675 gene function, whereas the genes without annotated function are called by the four last digits  
 2676 of their AGAP number and are depicted in green.

Gene_orientation_restriction	Primer sequence
APL1C_5'_EcoRI	TTGAATTCGAGCATTGAGTACCACAATGTGCTGGTTACACGCCG
APL1C_3'_V5Xba	CCTCTAGACTTTGTAACGCGACGCGTATC
LRIM1_5'_EcoRI	TTGAATTCAACGCGAAACGAAAGATGATGTCG
LRIM1_3'_V5Xba	CCTCTAGATCCCAGCTGGCTCGCTAAATTCTGC
RFPAPL1C_5'_EcoRI	TTGAATTCGAGCATTGAGTACCACAATGTGCTGGTTACACGCCG
RFPAPL1C_3'_SacII	AACCGCGGTGCAACTACCACATGTG

2677

2678 **Table S6. Primers sequences used for the plasmids construction.**

2679

# ANNEX

Annotation ID <sup>a</sup>	Gene name	Primer f 5'-3' <sup>b</sup>	Primer r 5'-3' <sup>b</sup>	Source <sup>c</sup>	Purpose
N/A	GFP	<u>GAATTTGTAATACGACTCACTATAGGCGATGGT</u> GAGCAAGGGCGAG	<u>GAATTTGTAATACGACTCACTATAGGCGCTTACTT</u> GTACAGCTCGTC	<a href="https://doi.org/10.1371/journal.ppat.1005306">https://doi.org/10.1371/journal.ppat.1005306</a>	dsRNA synthesis
AGAP007036	APL1A	<u>TAATACGACTCACTATAGGACTACCACCAGCCGAAAGATG</u>	<u>TAATACGACTCACTATAGGACTTGGTCTTGTATAGTACAATGG</u>	<a href="https://doi.org/10.1371/journal.ppat.1005306">https://doi.org/10.1371/journal.ppat.1005306</a>	dsRNA synthesis
AGAP007033	APL1C	<u>TAATACGACTCACTATAGGAGGCCAAGAAGACCGCAATCC</u>	<u>TAATACGACTCACTATAGGATCACAGTGATTTACGGGTGTGC</u>	<a href="https://doi.org/10.1371/journal.ppat.1005306">https://doi.org/10.1371/journal.ppat.1005306</a>	dsRNA synthesis
AGAP008072	NOX5	<u>TAATACGACTCACTATAGGCTCGTACGAATGGTCGAGTGA</u>	<u>TAATACGACTCACTATAGGCGCAAACTGGTCGCCTTGT</u>	DOI: 10.1126/sciimmunol.aal1505	dsRNA synthesis
AGAP009033	HPX2	<u>TAATACGACTCACTATAGGACGACGACGGTGTGTACAAG</u>	<u>TAATACGACTCACTATAGGATACTCGGCCGAATCGAAC</u>	DOI: 10.1126/sciimmunol.aal1505	dsRNA synthesis
AGAP010815	TEP1	<u>TAATACGACTCACTATAGGTTTGTGGGCTTAAAGCGCTG</u>	<u>TAATACGACTCACTATAGGACACGTAACCGCTCGGTAAG</u>	<a href="https://doi.org/10.1371/journal.ppat.0020052">https://doi.org/10.1371/journal.ppat.0020052</a>	dsRNA synthesis
AGAP010816	TEP3	<u>TAATACGACTCACTATAGGGCACCTCGACTGAGAAAGGTTTG</u>	<u>TAATACGACTCACTATAGGCTGATTATTTATATAGTTTAC</u>	<a href="https://doi.org/10.1371/journal.ppat.1005306">https://doi.org/10.1371/journal.ppat.1005306</a>	dsRNA synthesis
AGAP010592	S7	<u>CACCGCCGTGTACGATGCCA</u>	<u>ATGGTGGTCTGCTGGTCTT</u>	<a href="https://doi.org/10.1371/journal.ppat.1005306">https://doi.org/10.1371/journal.ppat.1005306</a>	qPCR
N/A	16S_V4q	<u>GTGCCAGCMGCCGCGGTAA</u>	<u>GGACTACHVGGGTWTCTAAT</u>	<a href="https://doi.org/10.1371/journal.ppat.1005306">https://doi.org/10.1371/journal.ppat.1005306</a>	qPCR
AGAP007036	APL1A	<u>GACTGCAAGCCGAGATCGATACC</u>	<u>CATCCATCTGGTCTTGAAGCTTA</u>	<a href="https://doi.org/10.1371/journal.ppat.1005306">https://doi.org/10.1371/journal.ppat.1005306</a>	qPCR
AGAP007033	APL1C	<u>AAGCAGGCTGAGTTGAGACAGG</u>	<u>GCCCAAGTAACATGACACAC</u>	<a href="https://doi.org/10.1371/journal.ppat.1005306">https://doi.org/10.1371/journal.ppat.1005306</a>	qPCR
AGAP012386	eater	<u>TTACCCGCTGTGCGAGGATGCAAGC</u>	<u>GTCACGTGCATAGTAGCGTCCCGTAGC</u>	DOI: 10.1073/pnas.1900147116	qPCR
AGAP009762	nimrodB2	<u>CAATCTGTCTCAAATGGCTGCTTCCACG</u>	<u>GCTGCAAACTTCGGTCCAGTGCATT</u>	DOI: 10.1073/pnas.1900147116	qPCR
AGAP009033	HPX2	<u>CGCTTGTACAAACAGATGA</u>	<u>CGACCAGATGGCGAAGTAT</u>	DOI: 10.1126/sciimmunol.aal1505	qPCR
AGAP008072	NOX5	<u>TCATGCTACGCTACTGGAAG</u>	<u>CGAGAAAGTCCACCTTGG</u>	DOI: 10.1126/sciimmunol.aal1505	qPCR
AGAP010815	TEP1	<u>AAAGCTGTTGGCTCAGGG</u>	<u>TTTCCACACACCAAAACGAA</u>	DOI: 10.1371/journal.ppat.1002023	qPCR
AGAP010816	TEP3	<u>ACCCGACGGCTACGTGATGG</u>	<u>CAACCTTTCTCAGTCGAGGT</u>	<a href="https://doi.org/10.1371/journal.ppat.1005306">https://doi.org/10.1371/journal.ppat.1005306</a>	qPCR
AGAP005503	SDR11	<u>CAGTGTGGTAGGCCATACGG</u>	<u>TTACGCTCACTAGGGACCG</u>	TM	qPCR
AGAP010058	PAS1	<u>CTCGGACACAACCTAGTTGG</u>	<u>AGCTGGAAGATCGCACTGGC</u>	TM	qPCR
AGAP009217	CLIPB12	<u>ATAATTCTCAGATGTGTGCG</u>	<u>CCATCCCAAAAGCATGTGCC</u>	TM	qPCR
AGAP001669	CPR6	<u>AAGACTGTTTGTATGTTGACC</u>	<u>AGATCCATCTCGTTTTGGCC</u>	TM	qPCR
AGAP008512	8512	<u>TTGGAACGGCTTTTGGACCG</u>	<u>CTTGTGGTGAACATTTGGC</u>	TM	qPCR
AGAP000113	113	<u>CAACTTCGCGAAACCTCGCC</u>	<u>CTGTTGCGAGAAAGCTGCC</u>	TM	qPCR
AGAP011509	COE	<u>GTGCCATACAGGGGGATGAT</u>	<u>ATCAACTCCACTGTCCAGCG</u>	TM	qPCR
AGAP013201	3201	<u>ATGATGTTCAATTGCGAGCC</u>	<u>AGATCTATCCCTGGAGAACCG</u>	TM	qPCR
AGAP006400	APH2	<u>CATGTCCAGACAAACCGTGA</u>	<u>TGCCGAGGTTGATTTCAACA</u>	TM	qPCR
AGAP007319	7319	<u>TTATGAAAGTGAAGCTGACGC</u>	<u>CCTCCGCAACCTGGTTATGCC</u>	TM	qPCR
AGAP010854	SCF45	<u>ACCTGTAGGACTTAGCACGCC</u>	<u>AAGGTGCTAAAGTAGCTGC</u>	TM	qPCR
AGAP011564	GGT	<u>ATTACCGGATCATACCGG</u>	<u>CGTTACGACCCCATCGCTG</u>	TM	qPCR
AGAP012648	2648	<u>GGTCCATGCCTCAAGAAGTC</u>	<u>AAGATTGCAAAAGGTGCGAG</u>	TM	qPCR
AGAP029110	9110	<u>CGAGCTACTAACGCTTGGCC</u>	<u>GGGAATTACAGGATACGCGG</u>	TM	qPCR
AGAP003334	CPLX2	<u>GTCCACGACGCCCGGTCGTG</u>	<u>CCACCGGGACCGGGTGCACGGA</u>	TM	qPCR
AGAP001748	Ch5	<u>GGCTACTTGAAGCTGGCGG</u>	<u>TCGTGATCTCAATTACCGG</u>	TM	qPCR
AGAP001470	1470	<u>GAAAGTAAAGCTCCGACACCT</u>	<u>CTGGCCGGCTGCTCGAGCT</u>	DOI: 10.1073/pnas.1900147116	qPCR
AGAP006728	COEAE7G	<u>CGGTGTAAAGCTGGCACCGC</u>	<u>TCGACCGGTGCTGGTACCG</u>	TM	qPCR
AGAP029790	9790	<u>GCAAGTAAAGCTGAAGTAG</u>	<u>TCATCGCTGATCGAAACG</u>	TM	qPCR
AGAP008246	GBD_46	<u>GGTCGGGACTGTGAACAAGAA</u>	<u>ACGAGGGTTTTCGACGAAACT</u>	TM	qPCR
AGAP029411	9411	<u>GCGATTTCTAGCCTTCTGTC</u>	<u>CGTGTGATGAGCGTAGG</u>	TM	qPCR
AGAP009381	CRBP	<u>GTGTGCGAATCCAAGTCAGC</u>	<u>CCACTCTACCCTGCAACCTG</u>	TM	qPCR
AGAP012477	GBD_77	<u>CCATGTCACTGGTTGCTG</u>	<u>ATAGAGCCGATTTGTGTGG</u>	TM	qPCR
AGAP011676	UP2M	<u>TTGCTTGAAGCTCTCACCTG</u>	<u>AAGACGAAACTCCCGCGA</u>	TM	qPCR
AGAP003350	GTP	<u>TCAGTTCAAGCAACCAATCCA</u>	<u>GCCTTAGGCACGAGAGTGT</u>	TM	qPCR
AGAP029623	9623	<u>CGATGAAGCGATCGAACTCT</u>	<u>GCAAGTTGCTGCTCGATGAA</u>	TM	qPCR
AGAP007300	AP	<u>AGTTTCCGACGTTGGTCTC</u>	<u>CACCGAGCGTTGCATAGTTG</u>	TM	qPCR
AGAP007959	7959	<u>TTGCCCTGAATATCGTGCTT</u>	<u>ACCCGCGCTGTTGCTATTAT</u>	TM	qPCR
AGAP001423	BPBP	<u>ATTATTGCGCGGAAAGTGTG</u>	<u>TCCTTCTTTCGCGACCACT</u>	TM	qPCR
AGAP007122	TU1	<u>GCTTCTGCTGTTCCACTCG</u>	<u>GGGTAGACGGCGAACTGTAG</u>	TM	qPCR
AGAP001352	1352	<u>GTTTCTTCTGCTCAACCCG</u>	<u>AGAGCCGGATTGGGACAAAC</u>	TM	qPCR
AGAP005296	U11_5296	<u>GAAAGGACAGCGCTTCTGTTA</u>	<u>CGCTTGGTGCATGCAATTTT</u>	TM	qPCR
AGAP007364	7364	<u>CGAGTCGGTTAAGTCTCTCG</u>	<u>TCGTCCAAACTCACTTCGCT</u>	TM	qPCR
AGAP006241	INX2	<u>GGTCCAGAGTGTGTCGAAGG</u>	<u>CACACGCTTCCCTCATCGTA</u>	TM	qPCR
AGAP000162	CBS	<u>GCAACGACGTTCCGGTATTCG</u>	<u>TCCAAACTGCCAACAGGGT</u>	TM	qPCR
AGAP000772	772	<u>TGCTAGGGGTACAGTGAAGA</u>	<u>GCCCCAGCACAATGCAAAAT</u>	TM	qPCR
AGAP009751	ANCE2	<u>ACTCCGGTGTGGAGCCGCGCG</u>	<u>CACGTAAGTCCGCGCAGCG</u>	TM	qPCR
AGAP004038	HPX8	<u>ACGGGCCCGGACCAATCCG</u>	<u>CCGCTAGTGACCCCTCCAGG</u>	TM	qPCR
AGAP010540	540	<u>GATTGGCCGGAAGGAAGTGA</u>	<u>CACGTGTTTGAGACGATCC</u>	TM	qPCR
AGAP005829	MAS_5829	<u>GGAAGGATACGATCCGCGAG</u>	<u>GTTCTGATCACCATCCGCT</u>	TM	qPCR
AGAP000180	PC	<u>AGCAGGAGCAGATTTGTGTG</u>	<u>CTTCATGTCGATCAGCGCAC</u>	TM	qPCR
AGAP004036	HPX7	<u>GTAGCGTGTACAGGAGGCGG</u>	<u>TACCGATGGTGGTGTGCG</u>	TM	qPCR
AGAP028655	8655	<u>TACAAGTACAGCTCCGTGCG</u>	<u>GGGGGCACTTCTACCTAGA</u>	TM	qPCR
AGAP003591	LRR1	<u>GTTCTGACACCCGAGCACG</u>	<u>TTCCCCGSGTAGACACCGG</u>	TM	qPCR
AGAP007036	APL1A_GSP	<u>CATTTGCTGCACGTGGAACGTAACCAGC</u>	N/A	TM	RACE
AGAP007033	APL1C_GSP	<u>GGCAAGCTTTAAGTTGGCGAAACGCA</u>	N/A	DOI: 10.1371/journal.pone.0003672	RACE
AGAP007036	APL1A_GSP_ne	<u>TAGGCTGAAAACCTTTACCTCGACCAC</u>	N/A	TM	RACE
AGAP007033	APL1C_GSP	<u>GTTGGATAAAGAGGGCAAAGCAGGC</u>	N/A	TM	RACE

2680

2681

<sup>a</sup> Applicable to mosquito genes; ID based on the Reference Strain Anopheles gambiae PEST.

2682

<sup>b</sup> T7 sequence is underlined.

2683

2684

<sup>c</sup> In case of published primers sequence, link to the publication source is indicated. TM refers to the primers designed in this manuscript.

2685

**Table S7. List of the primers used for dsRNA synthesis, qPCR and RACE.**

2686 **ABSTRACT**

2687

2688 **Abstract:** *Anopheles* mosquito is a vector of *Plasmodium* parasite, the causative agent  
 2689 of malaria. The *Anopheles* leucine-rich immune molecule (LRIM) family members are  
 2690 leucine-rich repeat (LRR) proteins, which certain were described as key antagonists of  
 2691 *Plasmodium* parasites in *Anopheles* mosquito midgut. APL1C (*Anopheles*  
 2692 *Plasmodium-responsive factor*) is a representative of LRIM members which  
 2693 specifically protects against rodent malaria parasites by stabilizing the complement-  
 2694 like protein TEP1. By combining cell biology with functional genomic approaches, this  
 2695 study shows that mosquito bloodmeal induce the presence of an extracellular layer of  
 2696 APL1C protein surrounding the midgut beneath of the basal lamina. Consistently with  
 2697 the formation of this layer, APL1C binds to the ookinetes that emerged on the basal  
 2698 side of the midgut. This presence occurs independently from TEP1 function, requires  
 2699 the contribution of the phagocytic cells and nitration pathway. In addition, APL1C  
 2700 defence function is not restricted to the ookinete in the midgut but it also acts against  
 2701 the latest *Plasmodium* stage, the sporozoites. APL1C inhibits salivary glands infection  
 2702 prevalence, and consistently, it also binds to the surface of the sporozoites in the  
 2703 hemocoel. However, unlike to the midgut stages, anti-sporozoites APL1C-dependent  
 2704 mechanism involves different partners. Moreover, RNAseq study revealed APL1C  
 2705 gene targets, including genes with immune-like function. These results generate novel  
 2706 biological insight for the function of APL1C, and probably other LRIM and LRR family  
 2707 members, as a pathogen recognition receptor inducing immune response against  
 2708 pathogens that come in contact with mosquito hemolymph compartment.

2709

2710 **Key words:** *Anopheles*, leucine-rich repeat, pathogen recognition, immune  
 2711 signalling, complement-like pathway, *Plasmodium*.

2712

2713 **Résumé :** L'anophèle est le moustique vecteur du parasite *Plasmodium*, responsable  
 2714 du paludisme. Chez ce moustique, des protéines de la famille LRIM (*Anopheles*  
 2715 *leucine-rich immune molecule*) ont été décrites comme antagonistes du  
 2716 développement de *Plasmodium*. L'une d'entre elles, APL1C (*Anopheles Plasmodium-*  
 2717 *responsive factor*) protège spécifiquement le moustique contre les parasites  
 2718 *Plasmodium* de rongeurs en collaboration avec la protéine TEP1 du complément. En  
 2719 combinant des approches de biologie cellulaire et de génomique fonctionnelle, mon  
 2720 travail de thèse montre que le repas sanguin des moustiques induit le recrutement  
 2721 d'APL1C au niveau du pôle basal de l'épithélium digestif. Ce positionnement d'APL1C  
 2722 lui permet de se lier aux stades ookinètes du parasite émergeant du côté basal de  
 2723 l'épithélium, et ce, indépendamment de la fonction de TEP1. Néanmoins, cette action  
 2724 d'APL1C requiert la contribution des phagocytes et de la Nitration. Par ailleurs, mon  
 2725 travail montre que l'action d'APL1C ne se restreint pas à l'ookinète car elle agit aussi  
 2726 contre le dernier stade de développement de *Plasmodium*, les sporozoïtes. Avec la

## ABSTRACT

2727 capacité de se lier aux sporozoïtes, APL1C contrôle la prévalence d'infection des  
2728 glandes salivaires du moustique, mais avec des partenaires différents de ceux  
2729 agissant sur les ookinètes. Finalement, une étude transcriptomique m'a permis  
2730 d'identifier des facteurs agissant en aval d'APL1C. L'ensemble de ces résultats génère  
2731 de nouvelles connaissances relatives à la fonction des membres de la famille LRIM en  
2732 tant que récepteurs de reconnaissance d'agents pathogènes capable de déclencher  
2733 une réponse immunitaire dans différents compartiments tissulaires du moustique.

2734 **Mots clés** : *Anopheles*, motifs riches en leucines, reconnaissance microbienne,  
2735 signalisation immunitaire, système assimilé au complément, *Plasmodium*.

2736

2737