



HAL
open science

Anopheles gambiae APL1 Is a Family of Variable LRR Proteins Required for Rel1-Mediated Protection from the Malaria Parasite, Plasmodium berghei

Michelle M Riehle, Jiannong Xu, Brian P Lazzaro, Susan M Rottschaefer, Boubacar Coulibaly, Madjou Sacko, Oumou Niare, Isabelle Morlais, Sekou F Traore, Kenneth Vernick

► **To cite this version:**

Michelle M Riehle, Jiannong Xu, Brian P Lazzaro, Susan M Rottschaefer, Boubacar Coulibaly, et al.. Anopheles gambiae APL1 Is a Family of Variable LRR Proteins Required for Rel1-Mediated Protection from the Malaria Parasite, Plasmodium berghei. PLoS ONE, 2008, 3 (11), pp.e3672. 10.1371/journal.pone.0003672 . hal-03040503

HAL Id: hal-03040503




<https://hal.science/hal-03040503>

Submitted on 4 Dec 2020

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Anopheles gambiae APL1 Is a Family of Variable LRR Proteins Required for Rel1-Mediated Protection from the Malaria Parasite, *Plasmodium berghei*

Michelle M. Riehle¹, Jiannong Xu¹, Brian P. Lazzaro², Susan M. Rottschaefer², Boubacar Coulibaly³, Madjou Sacko³, Oumou Niare³, Isabelle Morlais⁴, Sekou F. Traore³, Kenneth D. Vernick^{1,5*}

1 Department of Microbiology, University of Minnesota, Saint Paul, Minnesota, United States of America, **2** Department of Entomology, Cornell University, Ithaca, New York, United States of America, **3** Malaria Research and Training Center, University of Bamako, Bamako, Mali, **4** Laboratoire de Recherche sur le Paludisme, Institut de recherche pour le développement IRD-OCEAC, Yaoundé, Cameroun, **5** Unit of Insect Vector Genetics and Genomics, Department of Parasitology and Mycology, CNRS Unit URA3012: Hosts, Vectors and Infectious Agents, Institut Pasteur, Paris, France

Abstract

Background: We previously identified by genetic mapping an *Anopheles gambiae* chromosome region with strong influence over the outcome of malaria parasite infection in nature. Candidate gene studies in the genetic interval, including functional tests using the rodent malaria parasite *Plasmodium berghei*, identified a novel leucine-rich repeat gene, *APL1*, with functional activity against *P. berghei*.

Principal Findings: Manual reannotation now reveals *APL1* to be a family of at least 3 independently transcribed genes, *APL1A*, *APL1B*, and *APL1C*. Functional dissection indicates that among the three known *APL1* family members, *APL1C* alone is responsible for host defense against *P. berghei*. *APL1C* functions within the Rel1-Cactus immune signaling pathway, which regulates *APL1C* transcript and protein abundance. Gene silencing of *APL1C* completely abolishes Rel1-mediated host protection against *P. berghei*, and thus the presence of *APL1C* is required for this protection. Further highlighting the influence of this chromosome region, allelic haplotypes at the *APL1* locus are genetically associated with and have high explanatory power for the success or failure of *P. berghei* parasite infection.

Conclusions: *APL1C* functions as a required transducer of Rel1-dependent immune signal(s) to efficiently protect mosquitoes from *P. berghei* infection, and allelic genetic haplotypes of the *APL1* locus display distinct levels of susceptibility and resistance to *P. berghei*.

Citation: Riehle MM, Xu J, Lazzaro BP, Rottschaefer SM, Coulibaly B, et al. (2008) *Anopheles gambiae* APL1 Is a Family of Variable LRR Proteins Required for Rel1-Mediated Protection from the Malaria Parasite, *Plasmodium berghei*. PLoS ONE 3(11): e3672. doi:10.1371/journal.pone.0003672

Editor: Matthew W. Hahn, Indiana University, United States of America

Received: July 20, 2008; **Accepted:** October 20, 2008; **Published:** November 7, 2008

Copyright: © 2008 Riehle et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Funding: The reported work was supported by NIH AI042361 and AI073685 to KDV. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing Interests: The authors have declared that no competing interests exist.

* E-mail: kvernick@pasteur.fr

✉ Current address: Department of Biology, New Mexico State University, Las Cruces, New Mexico, United States of America

¶ These authors contributed equally to this work.

Introduction

Malaria is a global health problem resulting in over 1 million deaths annually, with disproportionate mortality in African children under the age of five [1]. Malaria also imposes a large economic burden on developing countries. Current efforts to control this disease are multifaceted and include use of insecticides and insect barriers, drug therapy, and strengthening healthcare and research infrastructures [2]. More consistent and widespread implementation of existing tools would be beneficial, although technical problems such as selection for chemico-resistance in vectors and parasites emphasize the need for a new generation of malaria control tools [3].

One such new approach could be limiting the genetic propensity of vector mosquitoes to serve as competent hosts for parasite development, thus decreasing or abolishing their ability to

transmit the causative agent. This approach is in its infancy and much remains to be done before we can evaluate specific genetic resistance mechanisms and the feasibility of manipulating them in nature.

We designed a phenotype-based method to genetically screen the wild *A. gambiae* population for genomic regions important in defense against *P. falciparum* [4]. Using this approach, we identified a genetic locus on chromosome 2L that consistently explains >80% of the variation in infection outcome (i.e., surviving oocyst numbers) in mosquitoes exposed to an infective bloodmeal, and thus captures most of the natural genetic variation for *P. falciparum* resistance or susceptibility [5]. The genetic interval, currently ~10 Mb, was termed the *Plasmodium*-Resistance Island (PRI). We then employed the rodent malaria laboratory model of *P. berghei* [6–12] to functionally screen candidate genes in the PRI. This work identified *APL1*, a novel leucine-rich repeat (LRR)

containing protein [5]. When *APL1* transcript abundance was reduced by RNAi gene knockdowns, the number of *P. berghei* oocysts was increased up to 20-fold, showing it to be a potent factor for host defense against *P. berghei* infection [5].

Here, we reannotate the original *APL1* gene as a gene family of 3 related members, *APLIA*, *B*, and *C*. Gene-specific RNAi assays show that all of the malaria-protective activity we previously reported for *A. gambiae APL1* can now be attributed exclusively to *APLIC*. We functionally dissect the position of *APLIC* in mosquito immune signaling networks, placing *APLIC* as a required node in Rel1-mediated host defense against *P. berghei* infection. Finally, we identify haplotypes in the *APL1* locus that are genetically associated with the degree of phenotypic susceptibility to *P. berghei* infection.

Results

Reannotation of the ENSEMBL prediction for *APL1*

Examination of *APL1* at the time of our original description [5] suggested that its annotation as a single gene (EN-SANGG00000012041 in ENSEMBL version 44 and earlier) was incorrect. The previous ENSEMBL prediction for *APL1* lacked start and stop codons, predicting a partial protein consisting of little more than a string of LRR domains. Resequencing of

genomic DNA and archived clones from the original *A. gambiae* sequencing project [13], as well as transcript mapping, revealed that the previous *APL1* gene represented the erroneous annotation of a gene family comprised of at least 3 tandem LRR-containing genes, here named *APLIA*, *APLIB*, and *APLIC* (Figure 1A). Each of the 3 genes has a short 5' exon followed by a small intron and a longer second exon, and each has a block of LRR motifs flanked by an N-terminal signal peptide and C-terminal coiled coil domains (Figure 1B). The three individual *APL1* genes display sequence similarity that likely results from gene duplication and functional diversification (diagonals in Figure S1), and we therefore class them together as the *APL1* family.

There were notable structural differences between the resequencing results and the ENSEMBL genome assembly (discussed in order below): i) the presence of structurally polymorphic haplotypes, and ii) a polymorphic and/or active transposable element. First, all three *APL1* family genes display major structural haplotypes (Figure 1B). The differences between allelic forms are most striking for *APLIA* and *APLIC*. Variants can differ in the locations of their predicted stop codons, resulting in predicted proteins of distinct lengths, and also by the presence of multiple polymorphic insertion-deletion (indel) sites within the protein coding sequence (CDS). The indels are precisely in-frame with the surrounding protein. Thus, the indels do not introduce missense

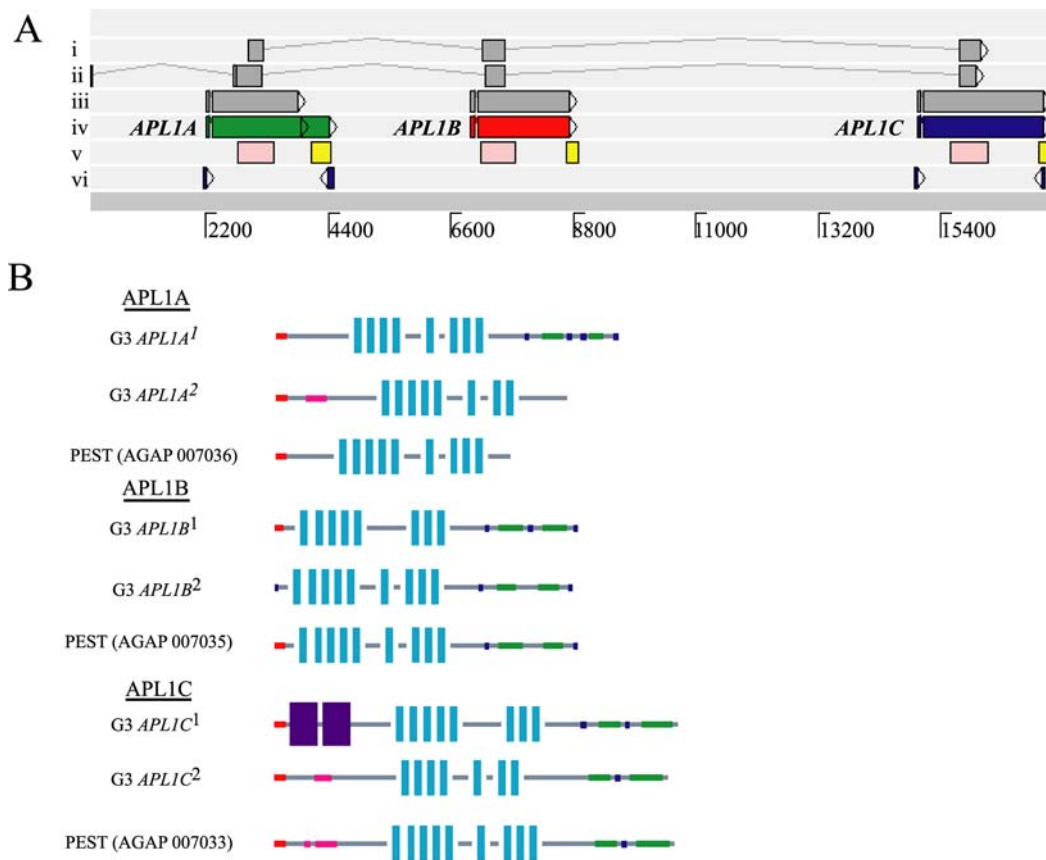


Figure 1. A. Reannotation of the *APL1* region. i) Ensembl release version 36, ii) Ensembl release version 41, iii) Ensembl release version 45, iv) Empirical annotation of *APLIA*, *B* and *C* in this article and Vectorbase manual annotation database, v) Fragments used for RNA interference assays; common dsRNA fragment knocking down *APLIA*, *B* and *C* (pink), unique dsRNA fragments used to delimit transcripts. **B. *APL1* family protein functional motifs.** Predicted peptide domains are indicated as follows: red, signal peptide; green, coiled-coil domain; light blue vertical bars, leucine rich repeats; blue, regions of intrinsic disorder, pink, segments of low complexity; and purple, repeat regions. Haplotypic versions of the *APL1* proteins as discussed in the text differ in functional predictions, with PEST strain predictions being most similar to the *APLIA*², *B*², and *C*² haplotypic forms. doi:10.1371/journal.pone.0003672.g001

mutations but rather encode small peptide cassettes that are present or absent, respectively, in the predicted finished protein. In a sample of wild and colony mosquitoes, indel alleles appear consistently linked to specific surrounding nucleotide variants (discussed below), thus establishing the indels as reliable markers for stable haplotypes that encode predicted proteins of distinct sizes and structure. The haplotypes are designated by the gene name followed by a superscript number (Figure 1B). The superscript 2 haplotype for each gene is most similar to the variant found in the PEST strain used for the *A. gambiae* genome sequence.

Another structural difference revealed by resequencing occurs upstream of the *APLIA* gene, where we found that a tract of Ns in the public genome assembly is actually (in PEST strain plasmid clone 19600445759751) a TA-III-Ag miniature inverted transposable element (MITE, [14]). The ambiguous bases in the ENSEMBL genome sequence may result from the difficulty in sequencing the repetitive region, or perhaps from the polymorphic state of the MITE in the sequence template. The G3 strain lacks the MITE in this genomic location.

APLIC is required for the control of *P. berghei* infection

We previously demonstrated that the *APLI* family had a pronounced effect on *P. berghei* oocyst intensity in an RNAi gene expression knockdown assay [5]. In that case, the double-stranded RNA (dsRNA) fragment injected into mosquitoes was fortuitously common to a portion of all three *APLI* family genes (Figure 1A, track v, homology regions of *APLI*-common dsRNA indicated by pink bars; also Figure S1B), because at the time *APLI* was annotated as a single gene. Based on our current reannotation of *APLI*, we wondered whether the source of our previously reported *APLI* knockdown phenotype was a single *APLI* family member, or alternatively a combined effect of all 3 members. To test this, we conducted knockdown experiments using new dsRNA constructs specific for each of the *APLI* family members (locations of dsRNAs shown in Figure 1A, track v, yellow bars). The results indicate that among the *APLI* family, *APLIC* alone is responsible for the control of *P. berghei* oocyst intensity (Figure 2). Oocyst loads in either *APLIA* or *APLIB* knockdown mosquitoes were statistically indistinguishable from those of the GFP controls, while *APLIC* knockdown mosquitoes carried oocyst loads ~20 times greater than GFP controls ($p < 0.05$). The effect of the *APLI*-common dsRNA fragment that silences the three genes was not different from the *APLIC* specific dsRNA, indicating that the effect of *APLIC* is both necessary and sufficient to explain *APLI* protective function against infection with *P. berghei*.

APLIC activity is required for anti-*P. berghei* protection mediated by the Toll/Rel1 pathway

The Toll signaling pathway controls cellular and humoral innate immune signaling, including activation of anti-fungal and anti-Gram positive bacteria defense in *Drosophila* [15,16] The genes for the core components of the pathway are also found in the *A. gambiae* and *Aedes aegypti* genomes [17,18]. Mosquito Rel1, ortholog of *Drosophila* Dorsal and functional analog of *Drosophila* Dif, is an ultimate transcription factor of the Toll pathway [11,19,20]. Under naive conditions, Rel1 is retained in the cytoplasm due to binding of the inhibitor, Cactus. Activation of the Toll receptor by its ligand spatzle in response to pathogens results in the disassociation of Cactus from Rel1 [21]. Released Rel1 translocates to the nucleus where it transactivates target genes. Recently, Frolet et al. [11] have shown that Rel1 regulates the transcription of *TEP1* and *LRIM1*, two anti-*Plasmodium* genes in *A. gambiae*, and that depletion of the Rel1 inhibitor, Cactus, by

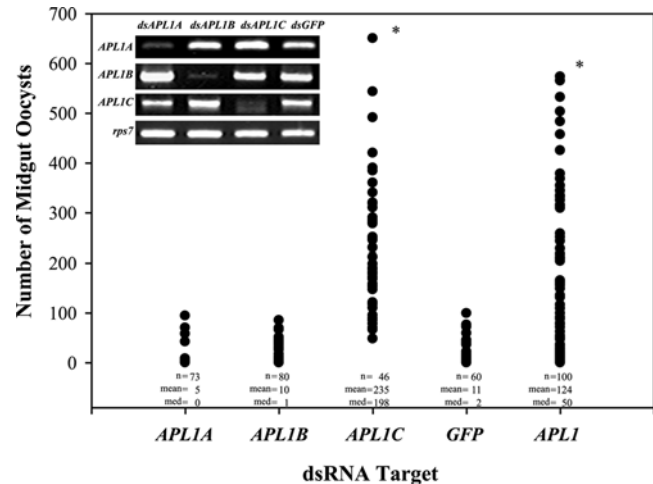


Figure 2. Among the *APLI* family, only *APLIC* specifically protects against *P. berghei* infection. The three *APLI* family genes were individually assayed to determine their relative contributions to host defense against *P. berghei* infection. Horizontal axis indicates the gene target of RNAi knockdown, where *APLIA*, *B* and *C* are gene-specific knockdowns, *GFP* is an irrelevant control dsRNA, and *APLI* is a shared dsRNA that simultaneously knocks down all 3 *APLI* family genes. Vertical axis indicates the number of midgut oocysts 7–8 d following IBM, with sample size (n), mean (mean), and median (med). Infection levels in *APLIA* and *APLIB* silenced mosquitoes were not different from *GFP* controls. However, treatment with either *dsAPLIB* or *dsAPLI* (targeting all 3 genes) permitted significantly greater oocyst development than the other treatments (asterisk, $p < 0.05$ by Dunn's Multiple Comparison after Kruskal Wallis one way ANOVA on ranks). Infections of *APLIC* and *APLI* silenced mosquitoes were not different from each other, indicating that the function of *APLIC* alone is sufficient to explain all of the increased permissiveness caused by complete *APLI* family knockdown. The result for each knockdown target represents pooled data from at least 2 independent replicate experiments. One of the common *APLI* knockdowns was done alongside an *APLIC* kd, and when assayed in the same replicate, there was no significant difference between *APLIC* kd and common *APLI* kd ($p = 0.481$). The inset gel photo shows representative examples of gene knockdown efficiency for the 3 members of the *APLI* gene family. Labels above gel images indicate the dsRNA that was used for the knockdown. Labels to the left of images indicate the transcript detected by RT-PCR on RNA purified from the treated mosquitoes. *rps7* was used as a control for cDNA input in PCR.
doi:10.1371/journal.pone.0003672.g002

RNAi strongly promoted mosquito host defense against *P. berghei*. In the same study, it was claimed that *APLI* was not transcriptionally regulated by Rel1. However, the PCR assay used for the *APLI* expression in their study was actually specific for the AGAP007037 gene (upstream of *APLIA*), which was incorrectly annotated as an exon of *APLI*.

A specific assay for the expression of *APLIC*, the only anti-*P. berghei* gene among the 3 *APLI* genes, reveals that *APLIC* is in fact regulated by Rel1, as determined by its expression in Rel1 and Cactus knockdown mosquitoes (Figure 3). *APLIC* transcription was reduced in *dsRel1* treated mosquitoes, and increased in *dsCactus* treated mosquitoes (Figure 3A). Furthermore, *APLIC* protein abundance was elevated following *P. berghei* infection, and this elevation was enhanced by *dsCactus* (Figure 3B). Boosting Rel1 signaling by depletion of Cactus enables mosquitoes to eliminate almost all invaded malaria parasites within 48 h post-infection [11]. If *APLIC* is a host-protective gene regulated by Rel1/Cactus signaling, then the anti-*P. berghei* effect of *dsCactus* should be reduced in the absence of *APLIC* expression. To test this hypothesis, we examined the effect of a *dsCactus*/*dsAPLIC* double

knockdown on the outcome of *P. berghei* infection. As expected, control mosquitoes treated only with *dsCactus* eliminated all invaded parasites by either lysis or melanization responses. Most *dsCactus* mosquitoes (24/36) harbored no parasites on day 7 post-IBM, while the remaining third of the mosquitoes (12/36) displayed melanized parasites (range 1–19 melanized parasites/mosquito) but no normal oocysts. Interestingly, this elevated

parasite resistance of *dsCactus* treated mosquitoes was completely reversed when *APL1C* was silenced simultaneously with *Cactus*, because the oocyst load in mosquitoes treated with *dsCactus*/*dsAPL1C* was equivalent to that in *dsAPL1C*/*dsGFP* treated mosquitoes (Figure 3C). These results indicate that *APL1C* is a required mediator of the anti-*P. berghei* effect controlled by Rel1 signaling, because constitutive activation of Rel1 signaling by

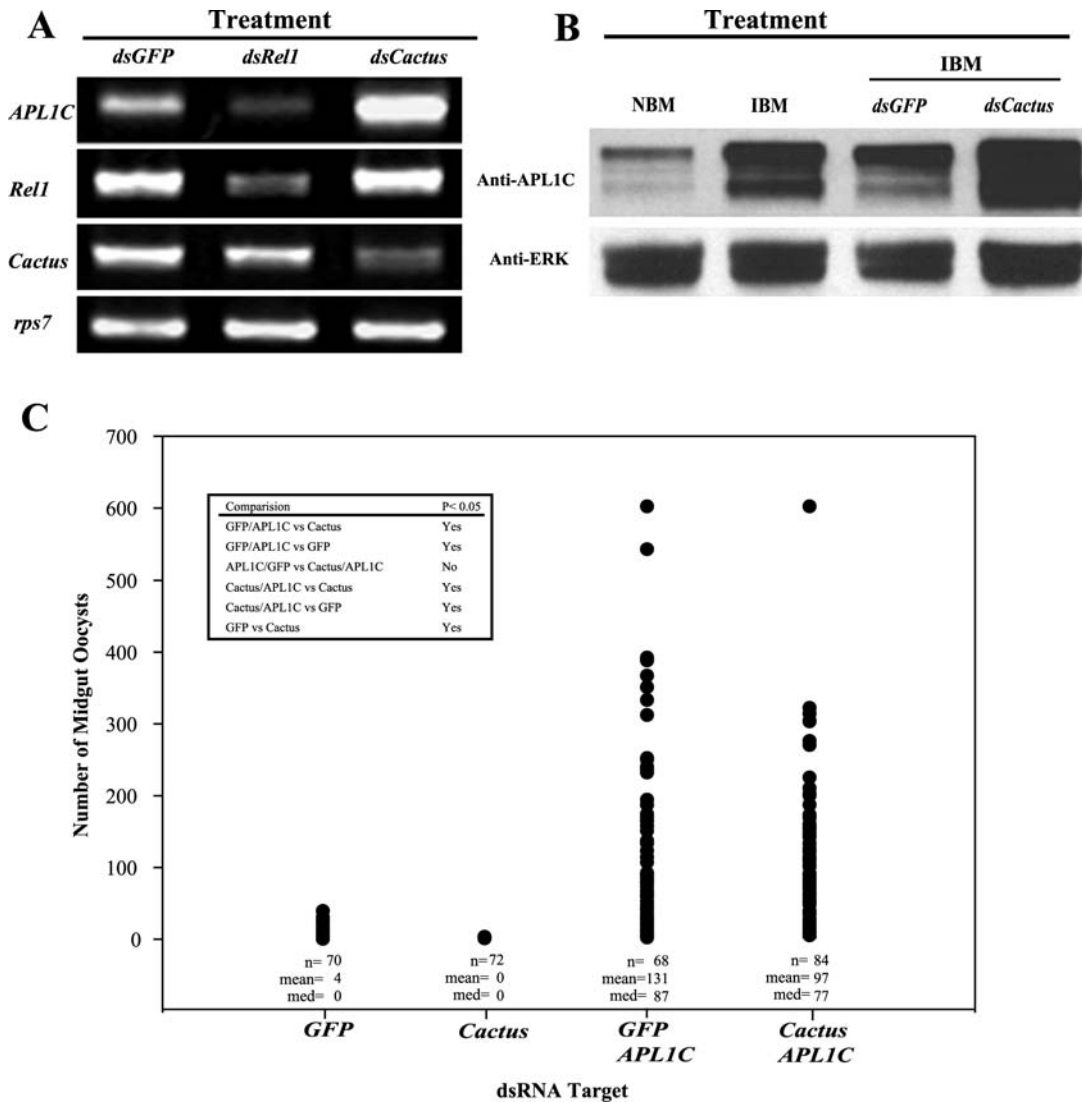


Figure 3. A. *APL1C* mRNA is regulated by the Rel1/Cactus immune signaling pathway. Semi-quantitative RT-PCR was used to measure the effect of *Rel1* and *Cactus* knockdown on *APL1C* transcript abundance. Labels above gel indicate the dsRNA that was used for the knockdown, labels to the left indicate the transcript whose abundance was measured. *APL1C* level was decreased by *Rel1* silencing, and increased by *Cactus* silencing, consistent with a model whereby *APL1C* RNA is regulated positively by Rel1 and negatively by Cactus. Specificity is shown by *dsRel1* and *dsCactus* silencing of cognate RNA levels. **B. *APL1C* protein abundance is regulated by *P. berghei* infection and Cactus.** Western blot analysis of *APL1C* protein abundance 24 h after *P. berghei* infection in naïve, *dsCactus* knockdown, and *dsGFP* control mosquitoes. Total protein abundance of ERK detected by anti-ERK antibodies was used as a protein loading control. **C. *APL1C* is required for Rel1-mediated host-defense against *P. berghei*.** The functional effect of *Cactus* and *APL1C* activity on the outcome of *P. berghei* infection was tested using gene knockdowns. The RNAi knockdown target is shown on the horizontal axis. *GFP* was used as a control dsRNA. The vertical axis shows the number of midgut oocysts 7–8 d following a *P. berghei* IBM. Sample sizes (n), means (mean), and medians (med) are given for knockdowns, which were each pooled data from 2 independent experiments. The oocyst loads differed significantly ($p < 0.001$) among samples by Kruskal Wallis one way ANOVA on Ranks. Pairwise comparisons using Dunn's method of multiple comparisons are given in the inset box. *dsCactus*-treated mosquitoes were completely protected from infection, with no successful oocyst development. In distinction, the double knockdown of *APL1C* simultaneously with *Cactus* produced mosquitoes with ~25-fold more oocysts than the *dsGFP*-treated controls, a result that was no different than the *dsAPL1C*-treated mosquitoes. Thus, in the absence of *APL1C*, Rel1 activation has no effect on parasite development, indicating the requirement of the presence of *APL1C* for Rel1-mediated host defense against *P. berghei*. Note that *dsGFP* was included in the *APL1C* knockdown so that total amount of input dsRNA per mosquito was the same as in the *Cactus-APL1C* double knockdown, eliminating dsRNA concentration as a variable. doi:10.1371/journal.pone.0003672.g003

Cactus depletion, which is normally entirely protective for mosquitoes, had no effect in the absence of *APLIC*. Treatment of mosquitoes with dsRNA for two other mosquito immune factors, TEPI and LRIM1, only partially reversed the dsCactus phenotype in double knockdowns, allowing development of some normal *P. berghei* parasites [11]. However, because Cactus depletion confers no host-protective phenotype without the presence of *APLIC*, it appears that the effect of *APLIC* is functionally dominant in this anti-*P. berghei* pathway. We hypothesize that *APLIC* may be a functionally upstream signaling node responsible for the coordinated Rel1-dependent control of multiple host protective factors.

Haplotypes at the *APLI* locus are genetically associated with protection from *P. berghei* infection

Resequencing and reannotation of the *APLI* locus identified major structural haplotypes that include alleles of *APLIA*, *APLIB*, and *APLIC* (Figure 1B). We hypothesized that the genetic and predicted protein variation in this locus could control differences for host defense against *P. berghei*. If true, one could test for the phenotypic effect of different haplotypes by detecting association of haplotype-specific markers with phenotypic outcome after *P. berghei* infection.

To design a genetic assay for haplotype, we resequenced part of the *APLIA* and *APLIC* genes in a sample of colony and wild mosquitoes. Patterns of specific SNPs distinguish the *APLI* haplotypes and are stable among mosquitoes of distinct geographic origins (Figure 4A). Thus, we designed a simple PCR fragment length assay to detect the genotype of indels in *APLIA* (Figure 4B, C). Typing of multiple mosquitoes using the PCR assay indicated that the indels are linked to and diagnostic for the *APLI* haplotypes (not shown). In summary, the *APLIA*¹ haplotype bears the deleted variant of an indel located in exon 2, while *APLIA*² bears the inserted variant for this indel. *APLIA*² bears the deleted variant of an indel located 5' to the predicted translation start site, while *APLIA*¹ bears the inserted variant of this indel (Figure 4B). The exon 2 indel is an in-frame CDS deletion, which consequently alters the length of the predicted protein product but does not introduce missense/nonsense mutations (not shown).

We challenged the G3 strain *A. gambiae* with a *P. berghei*-infected bloodmeal, and used the indel genotyping assay to query for an association between *APLI* genotype and infection outcome (Figure 4D). Mosquitoes that were homozygous for *APLIA*² had significantly lower oocyst loads (average = 5.3 ± 2.5) than *APLIA*¹/*APLIA*² heterozygotes (average = 31.7 ± 6; Mann Whitney Rank Sum Test, p < 0.001). Furthermore, the prevalence of infection in homozygous *APLIA*² mosquitoes was lower (37.5%; 9/24) than that in *APLIA*¹/*APLIA*² heterozygotes (84.5%; 41/48; Fisher's Exact Test, p < 0.001).

Thus, the haplotype-tagging indels are linked to genetic variation that controls significant difference in numbers of surviving *P. berghei* oocysts. We emphasize that the indels serve as genetic markers to detect the phenotypic effect of linked variant sequences, and linkage does not imply that the marker itself underlies the observed phenotypic variation. Like any other genetic marker in common use, for example almost all microsatellites and SNPs in any given genome, the allelic variation of the marker itself is presumed to be neutral for the trait under examination. Thus, further genetic studies will be necessary to resolve the underlying cause at the *APLI* locus of this phenotypic difference.

Based on the observation that *APLIA*² homozygotes are less susceptible and *APLIA*¹/*APLIA*² heterozygotes are more susceptible to *P. berghei* infection, one might predict that *APLIA*¹ homozygotes would in turn be exceedingly susceptible for infection. However, the

G3 colony of *A. gambiae* displays a low frequency of *APLIA*¹ homozygotes (8%) and thus the sample sizes of these individuals (n = 2) obtained after a moderate number of infections was too small to include in the statistical analyses. The *APLIA*¹ homozygotes must certainly be generated each generation by the mating of heterozygotes, but due to their low frequency their fate and potential infection phenotype remains an open question. Typing mosquito developmental stages using the haplotype diagnostic showed the same haplotype frequencies in eggs and larvae as in adults (data not shown), suggesting that loss of the *APLIA*¹ homozygotes is not developmental but rather prezygotic.

Discussion

We demonstrate by RNAi gene silencing experiments that *APLIC* activity is necessary to provide a high level of host-protective activity against *P. berghei* infection. By genetic association, we show that variant haplotypes at the *APLI* locus confer distinct levels of parasite susceptibility or resistance. Further functional and cell biological studies will be necessary to understand the role of *APLIC* in host defense against *P. berghei* and other parasite species, most importantly the human malaria parasite, *P. falciparum*. Further genetic studies will be required to dissect the functional basis of the distinct infection phenotypes linked to the *APLI* haplotypes.

Methods

Sequence Generation and Comparison

For PCR amplification and subsequent sequencing of the *APLI* region, DNA was isolated from a single female G3 strain mosquito using Qiagen DNeasy (Qiagen, CA, USA). Diploid sequence was obtained via standard Sanger sequencing at the BioMedical Genomics Center at the University of Minnesota and assembled using Seqman (DnaStar, WI, USA). Due to the diploid nature of the starting DNA, no phase information is known and ambiguous base calls are given for reliably called heterozygous positions.

Three PEST clones (Clones 19600445682690, 19600445759751, 19600445654898) spanning this *APLI* region were obtained from the Malaria Research and Reference Reagent Resource Center, MR4 (www.mr4.org). Clones overlapping the *APLI* gene predictions were sequenced and assembled as described above (map in Figure S1C). The three assembled sequences (G3, ENSEMBL, and PEST clones) were aligned and a comparison plot drawn using ACT [22]. The score cut-off was set to 100, the % ID cutoff to 50% and the minimum size of matches was set to 100 bp or greater.

Manual Annotation of the *APLI* Gene Family

Both molecular biology and *in silico* approaches were used to manually annotate the three member genes of the *APLI* family. All wet biology work was done on cDNA synthesized from total RNA isolated from a pool of 30 G3 strain female mosquitoes. Gene model predictions from a whole genome *in silico* prediction of the *A. gambiae* genome [23] were considered alongside data from 5' and 3' RACE reactions (First Choice RLM-RACE Kit, Ambion, CA, USA) run with gene specific primers for *APLIA* and *APLIC* including *apl1A_5'_end_R* GATGGCTGTCCCTCCGTTGGTACAGGC, *apl1C_5'_end_R* CCGTAATTTGGCTGACTTCTGTAGATT, *apl1A_3'_end_F* CAGCAGCAGCTCCTAGCAAGACTGCA, and *apl1C_3'_end_F* GGCAAGCGTTTAAAGTTGCGC-GAAACGCA and 5' forward and 3' reverse adaptor primers. The *APLIB* transcript was determined by designing primers to include all possible upstream start and downstream stop codons and screening for PCR amplification. Once the extent of transcripts was

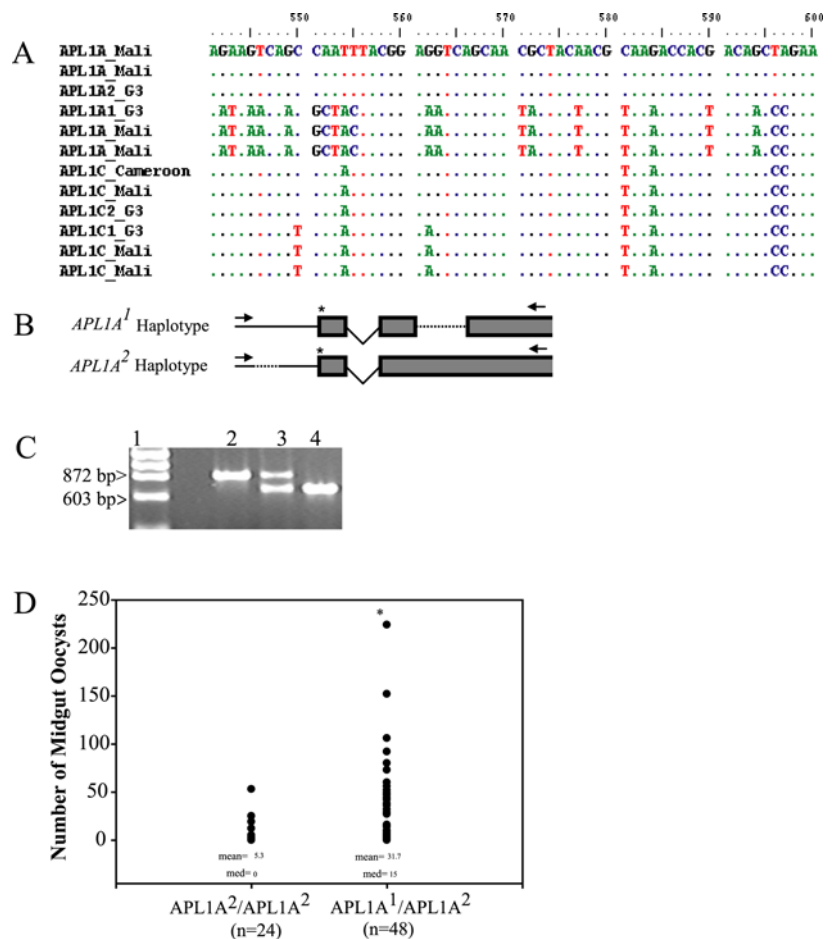


Figure 4. A. *APL1* alleles are comprised of blocks of linked SNPs. Nucleotide alignment shows *APL1A* and *APL1C* genetic variation in a sample of *A. gambiae* wild and colony mosquitoes. Haplotypes are shared over the resequenced region between individual mosquitoes, indicating that the variants are stable haplotypes that exist in *A. gambiae* at appreciable frequency. **B. Genotyping assay for segregating indels in the *APL1* locus.** A PCR assay was designed to test the state of 2 indels in *APL1A* that are markers for allelic haplotypes at the *APL1* locus. Arrows show forward and reverse primers used in the diagnostic assay, dashed lines indicate indels (located within exon 2 of *APL1A*¹, and above the translation start in *APL1A*²), asterisk indicates the start codon of *APL1A*. **C. Indel genotyping assay.** A representative gel of PCR products from the indel diagnostic assay, lane 1, Lambda HindIII/PhiX HaeIII marker, lane 2, *APL1A*² homozygote (854 bp), lane 3, a heterozygote with codominant bands, lane 4, *APL1A*¹ homozygote (663 bp). **D. *APL1* locus haplotypes control distinct levels of protection from *P. berghei* infection.** Homozygous *APL1A*² mosquitoes were significantly less permissive for parasite development than *APL1A*¹/*APL1A*² heterozygotes (asterisk, $p < 0.001$, Mann Whitney Rank Sum Test; sample size (n), mean (mean), and median (med) given for each haplotype). Homozygous *APL1A*¹ individuals are present at low frequency in the sampled colony and thus were not included in the analysis (see Results for further details on *APL1A*¹ haplotype frequency). doi:10.1371/journal.pone.0003672.g004

determined using this combinatorial approach, complete sequences were amplified by PCR and Sanger sequenced resulting in complete sequence of the coding regions and the ability to predict the underlying polypeptide. The resulting new gene models were submitted to Vectorbase (www.vectorbase.org) as manual annotation entries (data deposition *APL1*, manual). Three independently transcribed genes with start and stop codons are now included beginning with ENSEMBL release version 45: AGAP007036, *APL1A*; AGAP007035, *APL1B*; and AGAP007033, *APL1C*.

Protein Structural Architecture of the APL1 Family

Predicted peptide sequences of the haplotypic forms of the APL1 family were generated from the resequencing data described above and were aligned using ClustalX [24]. SMART [25] was used to obtain predictions for the protein domains using HMMER searches and also searching for outlier homologues, Pfam domains, signal peptides, internal repeats, and regions of intrinsic protein disorder. The protein domains depicted in Figure 1B are adapted

from SMART output that only displays domains more significant than established cutoffs. When two or more features occupy the same region, the following order of preference was used SMART>PFAM>PROSPERO repeats>Signal peptide>Transmembrane>Coiled coil>Unstructured regions>Low complexity.

APL1C Antibodies

An *APL1C* specific gene fragment was PCR amplified with the following primers *APL1C_F* 5' CAAGCGTTTAAGTTGCGC-GAAACGCAG and *APL1C_R* 5' CTACTTTGTAACGC-GACGCGTATCTGG. The fragment was expressed from the pet-46 Ek/LIC vector (Novagen, CA, USA), and was used for production of rabbit antisera. IgG was purified from serum using Protein A IgG Purification Kit from Pierce Biotechnology (Rockford, IL, USA), and made to a concentration of 1 mg/ml. Mosquito proteins were separated by 4–12% SDS-PAGE gels of the NuPAGE Novex Bis-Tris system (Invitrogen, Carlsbad, California, USA), transferred to the Nitrocellulose membrane.

Immunoblotting with the antibody against APL1C (1:1000 dilution) was done with the Protein Detector LumiGLO Western Blotting kit (KPL, Gaithersburg, Maryland, USA). The protein loads on the blot were checked by staining with anti-ERK2 total antibody (item K-23 #sc-153, Santa Cruz Biotechnology, Santa Cruz, CA, USA) at a concentration suggested by the supplier.

RNA Interference Assays

Gene specific fragments of *APL1 A, B* and *C*, *Rel*, *Cactus* and *GFP* were produced by PCR using oligos tagged at 5' end with T7 promoter sequence. The primers used were: GFP_F 5' AGTGGA-GAGGGTGAAGGTGA, GFP_R 5' CACGTGTCTTG-TAGTTCCCG, APL1A_F 5' CTACCACCTGCCGAAAGATG, APL1A_R 5' TCTGGTCTTGATAGTACAATGG, APL1B_F 5' TGAGAAACAAATAAGTTCAAAGTCC, APL1B_R 5' ACTCGCAAAGCTCAGCAAACAC, APL1C_F 5' CCAAGAA-GAACCGCAATCC and APL1C_R 5' TCACAGTGATTT-CAGGGTGTGC, REL1_F 5'GGCCCTAGTCAGCCG-CAGCCG and REL1_R 5'GGGGGGTTGGAATGGATGCTT; CACTUS_F 5' GGTGGTGCCTGCGATTGCTGG and CACTUS_R 5' GGCTTTTCGTTCAAGTTCTGTGC (all primers contained a T7 promoter sequence, 5' TAATACGACTCACTA-TAGG for use in synthesizing dsRNA; the GFP fragment was used as a dsRNA control). The PCR products were used as templates for dsRNA synthesis using the MEGAscript T7 Kit (Ambion, TX, USA). Four d after the dsRNA treatment, knockdown of the target gene was verified. dsRNA treated mosquitoes were fed on mice infected with PbGFPCON (8–12% parasitaemia with mature gametocytes), a transformed strain of *P. berghei* constitutively expressing GFP [26] at 8–12% parasitaemia with mature gametocytes. Seven to 8 d following infective blood meal, midgut oocysts were counted using a fluorescence microscope. To compare oocyst numbers across treatments non-parametric statistical tests were used, including the Mann Whitney Rank Sum Test and the Kruskal Wallis ANOVA on ranks. For each RNAi experiment at least 30 mosquitoes survived and were counted for oocyst load. Two to 4 independent replicate infections were performed and data was pooled prior to statistical analysis.

Indel Genotype-Phenotype Association Study

During the sequencing of the *APL1* gene family in a single G3 colony mosquito 2 segregating variants (haplotypes) were discovered (see Results). Based on available DNA sequence from homozygotes of the two segregating haplotypes, a PCR based diagnostic assay was designed to detect 2 segregating indels within *APL1A*. The first indel (46 bp in length) was located 109 bp upstream of the predicted translation start site and the second indel (237 bp in length) was located in the second exon. Amplification across this variable region results in differences in PCR product length of 191 bp in the *APL1A* gene, and was used to genotype mosquitoes for allelic haplotypes of the *APL1* locus. Homozygous *APL1A*¹ individuals displayed a single PCR of 663 bp, homozygous *APL1A*² individuals displayed a single PCR band of 854 bp, and heterozygotes displayed both bands. The

primers used in the haplotype diagnostic assay were 5' GCT GGA TCC CAA CTA GTG CTG TT and 5' AGT AAA GCA GCG GGC AGT TTG C. PCR conditions were 94°C for 1 minute, followed by 30 cycles of 94°C for 30 sec, 58°C for 30 sec, and 72°C for 45 sec and a final extension of 72°C for 7 minutes. To query for an association between haplotype and infection phenotype, G3 mosquitoes were fed on mice infected with the PbGFPCON strain of *P. berghei* (18) as described above. Seven to eight days post blood feeding midguts were dissected, oocysts counted, mosquito genomic DNA extracted from carcasses, and *APL1A* haplotype was determined by the diagnostic assay for each individual. Three replicate infections were performed and data was analyzed with a non-parametric Wilcoxon Mann Whitney test.

Supporting Information

Figure S1 A. Reannotation of the *APL1* region (reproduced from main text for clarity with Figure S1). i) Ensembl release version 36, ii) Ensembl release version 41, iii) Ensembl release version 45, iv) Empirical annotation of *APL1A, B* and *C* in this article and Vectorbase manual annotation database, v) Fragments used for RNA interference assays; common dsRNA fragment knocking down *APL1A, B* and *C* (pink), unique dsRNA fragments at the 3' end of each gene used for gene-specific knockdowns (yellow), vi) 5' and 3' RACE fragments used to delimit transcripts. B. Genomic similarity. ACT Sequence Comparison plot of i) a single G3 female, ii) the genomic sequence of *A. gambiae* from ENSEMBL and iii) sequence from three PEST clones spanning the region. Score cut-off was set at a minimum of 100, per cent ID cut-off was set to a minimum 50% and minimum size of matches was set to 100 bp or greater. Greater sequence identity is indicated by darker shade of red. Extensive sequence similarity is evident across the *APL1* locus region, with many regions >95% and most >90%. The assembled PEST strain sequence at ENSEMBL shows greater similarity to independent PEST clones than it does to G3. The LRR regions of the *APL1* genes show greatest intergene similarity (diagonals). The largest region of sequence dissimilarity occurs upstream of the 5' end of the *APL1A* gene. Here the ENSEMBL sequence is a string of Ns (see white box in track ii), the PEST clone has a miniature inverted transposable element (MITE) of the TA-III-Ag family based on terminal inverted repeat sequence and the G3 sequence has no MITE. C. Overlap of the PEST strain clones with the *APL1* gene family. a) The *APL1* gene family as presented throughout this paper, *APL1A* (green), *APL1B* (red), and *APL1C* (blue) (b) The three PEST clones obtained from MR4.

Found at: doi:10.1371/journal.pone.0003672.s001 (1.62 MB TIF)

Author Contributions

Conceived and designed the experiments: MMR JX SFT KV. Performed the experiments: MMR JX BPL SMR BC MS ON IM. Analyzed the data: MMR JX KV. Contributed reagents/materials/analysis tools: MMR JX BPL. Wrote the paper: MMR JX KV.

References

- Hay SI, Guerra CA, Tatem AJ, Noor AM, Snow RW (2004) The global distribution and population at risk of malaria: past, present, and future. *Lancet Infect Dis* 4: 327–336.
- Breman JG, Alilio MS, Mills A (2004) Conquering the intolerable burden of malaria: what's new, what's needed: a summary. *Am J Trop Med Hyg* 71: 1–15.
- Vernick KD, Waters AP (2004) Genomics and malaria control. *N Engl J Med* 351: 1901–1904.
- Niare O, Markianos K, Volz J, Oduol F, Toure A, et al. (2002) Genetic loci affecting resistance to human malaria parasites in a west african mosquito vector population. *Science* 298: 213–216.
- Riehle MM, Markianos K, Niare O, Xu J, Li J, et al. (2006) Natural malaria infection in *Anopheles gambiae* is regulated by a single genomic control region. *Science* 312: 577–579.
- Osta MA, Christophides GK, Kafatos FC (2004) Effects of mosquito genes on *Plasmodium* development. *Science* 303: 2030–2032.
- Volz J, Muller HM, Zdanowicz A, Kafatos FC, Osta MA (2006) A genetic module regulates the melanization response of *Anopheles* to *Plasmodium*. *Cell Microbiol* 8: 1392–1405.
- Volz J, Osta MA, Kafatos FC, Muller HM (2005) The roles of two clip domain serine proteases in innate immune responses of the malaria vector *Anopheles gambiae*. *J Biol Chem* 280: 40161–40168.

9. Abraham EG, Pinto SB, Ghosh A, Vanlandingham DL, Budd A, et al. (2005) An immune-responsive serpin, SRPN6, mediates mosquito defense against malaria parasites. *Proc Natl Acad Sci U S A* 102: 16327–16332.
10. Arrighi RB, Lycett G, Mahairaki V, Siden-Kiamos I, Louis C (2005) Laminin and the malaria parasite's journey through the mosquito midgut. *J Exp Biol* 208: 2497–2502.
11. Frolet C, Thoma M, Blandin S, Hoffmann JA, Levashina EA (2006) Boosting NF-kappaB-dependent basal immunity of *Anopheles gambiae* aborts development of *Plasmodium berghei*. *Immunity* 25: 677–685.
12. Michel K, Budd A, Pinto S, Gibson TJ, Kafatos FC (2005) *Anopheles gambiae* SRPN2 facilitates midgut invasion by the malaria parasite *Plasmodium berghei*. *EMBO Rep* 6: 891–897.
13. Holt RA, Subramanian GM, Halpern A, Sutton GG, Charlab R, et al. (2002) The genome sequence of the malaria mosquito *Anopheles gambiae*. *Science* 298: 129–149.
14. Tu Z (2001) Eight novel families of miniature inverted repeat transposable elements in the African malaria mosquito, *Anopheles gambiae*. *PNAS* 98: 1699–1704.
15. Lemaitre B, Hoffmann J (2007) The Host Defense of *Drosophila melanogaster*. *Annu Rev Immunol* 25: 697–743.
16. Lemaitre B, Nicolas E, Michaut L, Reichhart JM, Hoffmann JA (1996) The dorsoventral regulatory gene cassette *spatzle/Toll/cactus* controls the potent antifungal response in *Drosophila* adults. *Cell* 86: 973–983.
17. Christophides GK, Zdobnov E, Barillas-Mury C, Birney E, Blandin S, et al. (2002) Immunity-related genes and gene families in *Anopheles gambiae*. *Science* 298: 159–165.
18. Nene V, Wortman JR, Lawson D, Haas B, Kodira C, et al. (2007) Genome Sequence of *Aedes aegypti*, a Major Arbovirus Vector. *Science*.
19. Shin SW, Bian G, Raikhel AS (2006) A toll receptor and a cytokine, Toll5A and Spz1C, are involved in toll antifungal immune signaling in the mosquito *Aedes aegypti*. *J Biol Chem* 281: 39388–39395.
20. Bian G, Shin SW, Cheon HM, Kokoza V, Raikhel AS (2005) Transgenic alteration of Toll immune pathway in the female mosquito *Aedes aegypti*. *Proc Natl Acad Sci U S A* 102: 13568–13573.
21. Belvin MP, Jin Y, Anderson KV (1995) Cactus protein degradation mediates *Drosophila* dorsal-ventral signaling. *Genes Dev* 9: 783–793.
22. Carver TJ, Rutherford KM, Berriman M, Rajandream M-A, Barrell BG, et al. (2005) ACT: the Artemis comparison tool. *Bioinformatics* 21: 3422–3423.
23. Li J, Riehle MM, Zhang Y, Xu J, Oduol F, et al. (2006) *Anopheles gambiae* genome reannotation through synthesis of ab initio and comparative gene prediction algorithms. *Genome Biol* 7: R24.
24. Thompson JD, Gibson TJ, Plewniak F, Jeanmougin F, Higgins DG (1997) The CLUSTAL_X windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. *Nucleic Acids Res* 25: 4876–4882.
25. Letunic I, Copley RR, Pils B, Pinkert S, Schultz J, et al. (2006) SMART 5: domains in the context of genomes and networks. *Nucleic Acids Res* 34: D257–260.
26. Franke-Fayard B, Trueman H, Ramesar J, Mendoza J, van der Keur M, et al. (2004) A *Plasmodium berghei* reference line that constitutively expresses GFP at a high level throughout the complete life cycle. *Mol Biochem Parasitol* 137: 23–33.