

Research



Cite this article: Lambrechts L, Quillery E, Noël V, Richardson JH, Jarman RG, Scott TW, Chevillon C. 2013 Specificity of resistance to dengue virus isolates is associated with genotypes of the mosquito antiviral gene *Dicer-2*. *Proc R Soc B* 280: 20122437. <http://dx.doi.org/10.1098/rspb.2012.2437>

Received: 14 October 2012

Accepted: 30 October 2012

Subject Areas:

genetics, immunology, evolution

Keywords:

Aedes aegypti, dengue virus, genotype-by-genotype interaction, *Dicer-2*, RNAi

Author for correspondence:

Louis Lambrechts

e-mail: louis.lambrechts@pasteur.fr

Electronic supplementary material is available at <http://dx.doi.org/10.1098/rspb.2012.2437> or via <http://rspb.royalsocietypublishing.org>.

Specificity of resistance to dengue virus isolates is associated with genotypes of the mosquito antiviral gene *Dicer-2*

Louis Lambrechts¹, Elsa Quillery², Valérie Noël², Jason H. Richardson³, Richard G. Jarman⁴, Thomas W. Scott⁵ and Christine Chevillon²

¹Insects and Infectious Diseases, Institut Pasteur, CNRS URA 3012, 25 rue du Docteur Roux, 75724 Paris Cedex 15, France

²MIVEGEC Laboratory (UMR 5290 CNRS-IRD-Université Montpellier I-Université Montpellier II; UR 224 IRD), campus IRD, 911 avenue Agropolis, BP 64501, 34394 Montpellier Cedex 5, France

³Entomology Branch, and ⁴Viral Disease Branch, Walter Reed Army Institute of Research, Silver Spring, MD 20910, USA

⁵Department of Entomology, University of California, One Shields Avenue, Davis, CA 95616, USA

In contrast to the prevailing view that invertebrate immunity relies on broad-spectrum recognition and effector mechanisms, intrinsic genetic compatibility between invertebrate hosts and their pathogens is often highly specific in nature. Solving this puzzle requires a better understanding of the molecular basis underlying observed patterns of invertebrate host–pathogen genetic specificity, broadly referred to as genotype-by-genotype interactions. Here, we identify an invertebrate immune gene in which natural polymorphism is associated with isolate-specific resistance to an RNA virus. *Dicer-2* (*dcr2*) encodes a key protein upstream of the RNA interference (RNAi) pathway, a major antiviral component of innate immunity in invertebrates. We surveyed allelic polymorphism at the *dcr2* locus in a wild-type outbred population and in three derived isofemale families of the mosquito *Aedes aegypti* that were experimentally exposed to several, genetically distinct isolates of dengue virus. We found that *dcr2* genotype was associated with resistance to dengue virus in a virus isolate-specific manner. By contrast, no such association was found for genotypes at two control loci flanking *dcr2*, making it likely that *dcr2* contains the yet-unidentified causal polymorphism(s). This result supports the idea that host–pathogen compatibility in this system depends, in part, on a genotype-by-genotype interaction between *dcr2* and the viral genome, and points to the RNAi pathway as a potentially important determinant of intrinsic insect-virus genetic specificity.

1. Introduction

Intrinsic compatibility between invertebrate hosts and their pathogens often depends on the specific pairing of genotypes [1,2]. This genetic specificity, referred to as genotype-by-genotype interactions, has been documented in a wide variety of invertebrate host–pathogen systems [3–9] and can reach extreme levels [10]. This observation is in contrast with the conventional view that invertebrate immune systems only respond differently to broad classes of pathogens, for example, Gram-positive and Gram-negative bacteria [11]. To resolve this apparent paradox, a better understanding of the molecular basis underlying observed genotype-by-genotype interactions between invertebrates hosts and pathogens is required [12,13]. Molecular determinants of host–pathogen specificity have been well described in plant-pathogen systems [14,15] and in the case of compatibility patterns between genetically diverse pathogens and major histocompatibility complex (MHC) variants of vertebrates [16,17]. With a few exceptions [18–21], the genetic basis of invertebrate host–pathogen compatibility is poorly understood. Identifying the mechanisms underlying host–pathogen genetic specificity has important implications for a broad spectrum of evolutionary, epidemiological and medical phenomena [1,13].

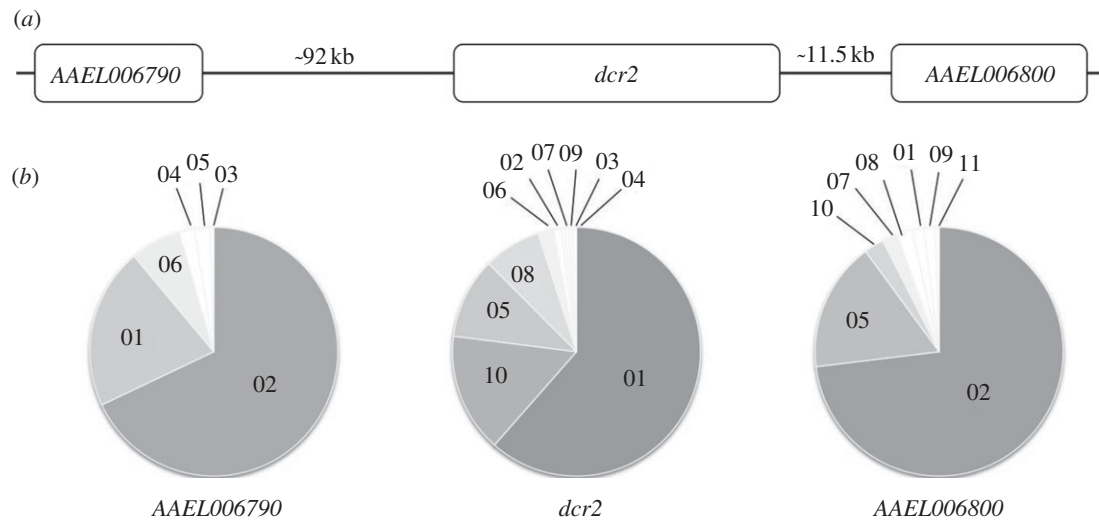


Figure 1. Allelic polymorphism in the *dcr2* genomic region of *Aedes aegypti*. (a) Schematic of the *Dicer-2* (*dcr2*) genomic region showing the position of flanking genes (*AAEL006790* and *AAEL006800*) used as control loci. Distances are not drawn to scale. (b) Pie charts of the frequency of alleles (coded with numbers, independently for each locus) identified in a sample of 100 *Ae. aegypti* females from the outbred mosquito population for the three genes shown in (a).

In the present study, we examined the role of an invertebrate immune gene in strain-specific resistance to an RNA virus. *Dicer-2* (*dcr2*) encodes a ribonuclease acting upstream of the RNA interference (RNAi) pathway [22], a major antiviral component of innate immunity in invertebrates [23–25]. Building on a previous study that revealed strong genetic specificity in the interaction between *Aedes aegypti* mosquitoes and dengue viruses (DENV) in a natural situation in Thailand [5], we sought to determine whether genetic polymorphism at the *dcr2* locus could explain, in part, the observed compatibility pattern. DENV (serotypes 1–4) are mosquito-borne RNA viruses of the genus *Flavivirus* that cause a spectrum of clinical manifestations ranging from mild febrile illness to life-threatening haemorrhagic fever, and are the most important insect-borne viral infection of humans [26]. *Aedes aegypti* is the primary mosquito vector of DENV worldwide [26].

We genotyped *dcr2* and two flanking loci (*AAEL006790* and *AAEL006800*; figure 1a) in a wild-type *Ae. aegypti* population from Ratchaburi, Thailand and in three independent isofemale families (i.e. the progeny of individual females) derived from this population. *AAEL006800* is a gene encoding a sodium/chloride-dependent transporter located approximately 11.5 kb downstream of *dcr2*, and *AAEL006790* is a hypothetical protein-coding gene located approximately 92 kb upstream of *dcr2* (figure 1a). They were chosen in the close vicinity of *dcr2* as ‘negative controls’ to verify the locus-specificity of genotype–phenotype associations tested at the *dcr2* locus. *AAEL006800* is the closest gene downstream of *dcr2*, whereas two other short genes (*AAEL006801* and *AAEL006808*) are located between *dcr2* and *AAEL006790*. The outbred population and the isofemale families represent two complementary approaches to assess genotype–phenotype associations. The outbred population provides a representative sample of the natural genetic diversity at a locus, but has reduced power to detect genotype–phenotype associations at this locus because the genetic background, including other genes underlying the phenotype, is diverse. In other words, background variation in the outbred population results in a stringent, although less sensitive, test of genotype–phenotype associations. Conversely, isofemale families display restricted overall genetic diversity (a maximum of four parental alleles

per locus), and increased power to detect genotype–phenotype associations because siblings share a relatively homogenous genetic background. We measured the probability of viral infection of midgut epithelial cells and subsequent dissemination to secondary tissues, which are two essential steps for DENV transmission by *Ae. aegypti* bite. Both events are prerequisites for virus transmission and have been used to define a ‘midgut infection barrier’ and a ‘midgut escape barrier’ underlying *Ae. aegypti* resistance to DENV [27].

2. Material and methods

(a) Mosquitoes and viruses

A laboratory *Ae. aegypti* population was established with a large number (more than 1500) of immatures (larvae and pupae) collected during 2007 in Ratchaburi, Thailand [5]. Mosquitoes were maintained as a large, randomly mating population (more than 1800 adults per generation) under standard conditions for two generations to minimize the influence of parental effects that can be confounded with genetic effects [28]. We also derived three independent *Ae. aegypti* isofemale families (denoted as A, B and C) generated as the progeny of three inseminated females randomly chosen from the outbred population. Families A, B and C were represented by 125, 112 and 96 individual females, respectively. We examined statistical associations between genetic polymorphism and infection phenotype both in the parental, outbred population and within the isofemale families. We exposed mosquitoes to an artificial DENV infectious blood meal as described previously [5]. The parental population was exposed to two DENV-1 isolates collected in Ratchaburi, Thailand in 2007 (RTB-138 and RTB-196) represented by 20 and 40 phenotyped *Ae. aegypti* females, respectively. The three families were exposed to three other DENV-1 isolates (BKK, KPP and RTB) collected during 2007 in Thailand [5], resulting in 28–53 (mean 37) phenotyped females per family-isolate pair. Prior to their use in experimental infections, all five isolates were passaged five times in *Aedes albopictus* (C6/36) cells, with the exception of RTB-196 that was passaged two times in *Toxorhynchites splendens* mosquitoes and three times in C6/36 cells.

Two conventional indices of mosquito resistance to DENV were measured: (i) the proportion of mosquitoes that became infected and (ii) the proportion of infected mosquitoes (excluding

uninfected ones) that developed a disseminated infection (virus was detected in their head or their legs) 14 days after the infectious blood meal at 28°C [5]. These phenotypes relate to the *Ae. aegypti* 'midgut infection barrier' and 'midgut escape barrier' to DENV transmission that were previously described [27]. The complete genome sequences of BKK, KPP and RTB virus isolates (GenBank accession nos HM469966, HM469967 and HM469968, respectively), showed 2.4 per cent overall nucleotide divergence and 1.8 per cent mean pairwise nucleotide divergence. The complete genome sequences of isolates RTB-138 and RTB-196 are not available yet, but the sequence of more than 96 per cent of their genomes showed similar levels of nucleotide divergence. Because the infectious blood meal offered to mosquitoes was standardized (including the infectious dose), isolate was used as a proxy for virus genetic identity.

Phylogenetic relationships between the five DENV-1 isolates in the study and a representative set of DENV-1 isolates previously collected in southeast Asia were examined by focusing on the coding part of the viral genome (see alignment in the electronic supplementary material, file S1). Absence of recombination events was verified with the RDP3 software [29]. Phylogenetic relationships were inferred with MEGA5 [30] using the maximum-likelihood method and the general reversible model (GRT + I + I) based on a Gamma discrete distribution and a 3.63 per cent proportion of invariant sites. The bootstrap consensus tree was based on 500 replicates and the few nucleotide positions containing gaps were ignored.

(b) Mosquito genotyping

The *dcr2* gene (ID AAEL006794 in VectorBase: <http://aegypti.vectorbase.org/>) was genotyped by amplifying a 572-bp region located at positions 1559–2106 in the transcript sequence (AAEL006794-RA) consisting of (i) 205 bp in exon 7, transcript positions 1599–1803, located in a putative helicase domain; (ii) a 64-bp intron; and (iii) 303 bp in exon 8, transcript positions 1804–2106, in a putative dsRNA binding domain. Specific forward and reverse PCR primers were as follows: FOR 5'-CGTGTAGAGGAGGGGATTG, REV 5'-CAACTTGATATTGC GCATGG. PCR products were verified on an agarose gel and sequenced by the Sanger technique. The intron contained two poly-T sequences of 10+ nucleotides that appeared to present frequent insertions/deletions in this *Ae. aegypti* population, which resulted in unreadable sequencing chromatograms for heterozygotes and prevented further analysis. Exon 7, therefore, was sequenced using the forward primer and exon 8 using the reverse primer. The 205 nucleotides of exon 7 were 100 per cent conserved in both the outbred population and the isofemale families, so that *dcr2* genotype was determined based on exon 8 polymorphism. Seven biallelic single nucleotide polymorphisms (SNPs) were identified (1806C > T, 1839C > T, 1840C > A, 1851C > T, 1884G > A, 2004G > C and 2029A > G, where numbers refer to the transcript position and letters indicate the most and least frequent alleles, respectively). All are synonymous substitutions with the exception of 1840C > A and 2029A > G, which result in a leucine to isoleucine (Leu > Ile) and isoleucine to valine (Ile > Val) amino acid changes in the protein, respectively. The AAEL006800 control locus was genotyped based on six biallelic SNPs identified by amplifying a 417-bp region of exon 6 with the following PCR primers: FOR 5'-GCGTCGTGCCGGTCTG AGTC-3' and REV 5'-GGTTTCCTGCCCCAACGG-3'. The AAEL006790 control locus was genotyped based on seven biallelic SNPs identified by amplifying a 266-bp region including Intron 2–3 (144 bp) with the following PCR primers: FOR 5'-GGTTACC GTCGCTGAAAGAA-3' and REV 5'-ATGCTCCACAGTACCGA TCA-3'. All sequencing chromatograms were aligned in *Geneious Pro* v. 5.6 [31] and verified manually. Alleles were reconstructed using the PHASE algorithm implemented in *DnaSP* v. 5.0 [32],

with separate runs for each isofemale family and for the parental population. We primarily used allelic information instead of individual SNPs because SNPs are most often biallelic and, therefore, might not faithfully represent the true allelic diversity of the locus. The number of alleles identified is contingent on the number of SNPs considered per locus, which in turn depends on the size of the PCR product. We arbitrarily chose a number of SNPs that provided sufficient resolution to discriminate at least four alleles (i.e. the maximum number of alleles expected to be found in the isofemale families).

(c) Population genetics

The occurrence of linkage disequilibrium (LD) was tested for each pair of loci in the outbred population using exact probability tests for genotypic disequilibrium (batches = 5000; iterations = 20 000; dememorization = 1000) implemented in *GENEPOP* v. 4.0 [33]. Recombination parameters were estimated in *PHASE* v. 2.1.1 [34] using the general model for recombination rates of multi-allelic loci other than microsatellites (i.e. that does not assume a stepwise mutation model). All default priors were unchanged with the exception of a prior background recombination parameter of 2×10^{-5} . Different prior values (range 4×10^{-6} – 8×10^{-3}) for the background recombination parameter did not qualitatively change the results. The final run was iterated 10 times (all loci). Results shown are those with the best average goodness-of-fit from 10 independent runs of the algorithm. The possibility that laboratory colonization resulted in a demographic bottleneck was examined in the outbred population using two complementary analyses. First, because allele number decreases faster than heterozygosity when populations shrink in size, a signature of bottlenecks observed in subsequent generations is $H_e > H_{eq}$, where H_e is the expected heterozygosity in the population of interest under Hardy–Weinberg equilibrium and H_{eq} is the expected heterozygosity in a population with the same sample size and allele number at mutation-drift equilibrium [35]. This analysis was performed using *BOTTLENECK* v. 1.2 [36]. Second, we calculated Tajima's *D* and Fu and Li's *F** statistics for each of the three genes in *DnaSP* v. 5.0 [32]. Negative values of *D* and *F** indicate an excess of rare alleles relative to expectation under demographical stability and quasi-neutral evolution, and are interpreted as a recent population bottleneck (or a selective sweep in a demographically stable population).

(d) Genotype–phenotype associations

For the outbred population, only the two most common genotypes of each locus were included in the genotype–phenotype association analyses because the sample sizes for other genotypes were too small for a meaningful analysis. The proportion of infected mosquitoes and the proportion of infected mosquitoes that developed a disseminated infection (excluding uninfected ones) were analysed using a nominal logistic regression that included the effects of genotype, isolate, infectious dose and their interactions up to the second order. The isolate \times infectious dose interaction was omitted from the model because one of the two isolates had only been used at one virus concentration and, therefore, this interaction could not be tested. For the isofemale families, the analysis was conducted in two steps. First, the proportion of infected females and the proportion of infected females with a disseminated infection were analysed with a nominal logistic regression that included the effects of family, isolate, genotype and their interactions up to the second order. Genotype was nested within family because each family contained a different set of genotypes. Second, because the main effect of the family proved insignificant in the initial analysis, the same analysis was performed without including the family, so that all genotypes

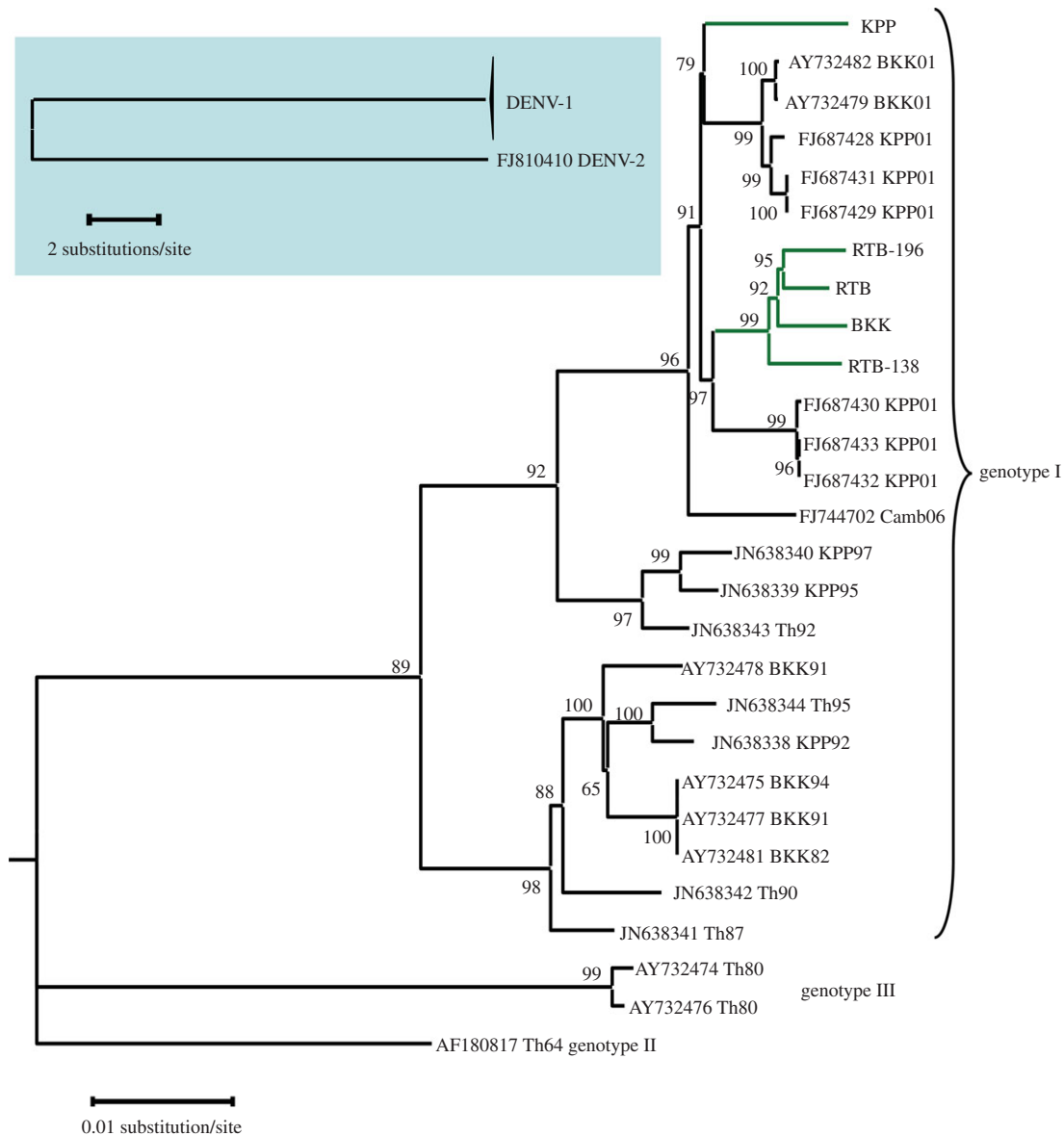


Figure 2. Phylogenetic relationships among DENV-1 isolates used in this study. Background genomic sequences retrieved from GenBank are labelled with their accession number followed by an indication of their geographical origin (Camb, Th, BKK and KPP stand for Cambodia, Thailand, Bangkok and Kamphaeng Phet, respectively), and by two digits referring to the sampling year. Major clades are labelled according to previously described DENV-1 ‘genotypes’. The five DENV-1 isolates used in this study are indicated by green terminal branches. They were isolated in 2007 from three different Thai localities: Bangkok (BKK), Kamphaeng Phet (KPP) and Ratchaburi (RTB, RTB-138, RTB-196). The inset in the upper left corner shows the position of the DENV-2 isolate that was used to root the tree. Bootstrap support values are shown next to the relevant nodes.

were pooled regardless of family. Differences were considered statistically significant at $p < 0.05$. All statistical analyses were performed with *JMP* v. 10.0 (SAS Institute Inc., NC, USA).

Full datasets, including phenotypes, genotypes and individual SNP calls are provided for both the outbred population (see the electronic supplementary material, file S2) and the isofemale families (see the electronic supplementary material, file S3).

3. Results

Phylogenetic analysis of the coding region of viral genomes detected no major recombination event among the five DENV-1 isolates used in the present study and other viruses that recently circulated in Thailand. The five DENV-1 isolates of the study belong to two sister groups nested within a large phylogenetic lineage referred to as genotype I (figure 2). All five isolates are closely related to viruses that circulated during 2001 in Bangkok and Kamphaeng Phet (see BKK01

and KPP01 in figure 2). Four of the five isolates cluster together in one sister group, whereas the KPP isolate diverged slightly earlier and belongs to the other sister group. Pairwise differences between isolates were nucleotide substitutions scattered throughout the viral genome (see the electronic supplementary material, file S1).

Single nucleotide polymorphisms (SNPs) surveyed at the *dcr2* locus in a sample of 100 individuals of the outbred *Ae. aegypti* population defined 17 genotypes resulting from 10 different *dcr2* alleles (provisionally named 01 to 10) with frequencies ranging from 0.5 to 61.5 per cent (figure 1b). Flanking genes displayed a similar pattern of polymorphism with 11–12 genotypes based on six to eight alleles per locus (figure 1b). Consistent with their close physical proximity, genotypic LD was significant between each pair of loci (*dcr2*-AAEL006790: $p = 0.0017$; *dcr2*-AAEL006800: $p = 0.0023$; AAEL006790-AAEL006800: $p = 0.0433$). Fifteen of the 42 possible pairs of individual SNPs resulted in

significant LD between *dcr2* and *AAEL006790*, whereas 8 of the 42 possible pairs of SNPs showed significant LD between *dcr2* and *AAEL006800*. The median value of recombination rate estimated for the *dcr2* genomic region was 4×10^{-6} per generation (95% CI 1.0×10^{-8} – 3.7×10^{-4}). Assuming an effective population size between 500 and 1000 individuals, this corresponds to 0.1–0.2% recombination per Mb, in agreement with a previous genome-wide estimate of 0.15 per cent recombination per Mb [37,38]. There was no indication of a population bottleneck that may have occurred during laboratory colonization of the population. Heterozygosity calculated from the allele frequencies (H_e) at the three loci did not deviate significantly from expected heterozygosity at mutation-drift equilibrium (H_{eq}), with a tendency for a heterozygosity deficit instead of the expected heterozygosity excess following a population bottleneck (*dcr2*: $H_e = 0.59$, $H_{eq} = 0.69$, $p = 0.15$; *AAEL006790*: $H_e = 0.49$, $H_{eq} = 0.53$, $p = 0.35$; *AAEL006800*: $H_e = 0.44$, $H_{eq} = 0.62$, $p = 0.13$). Similarly, Tajima's D and Fu and Li's F^* statistics did not detect any departure from neutral evolution at mutation-drift equilibrium, with a tendency for a deficit of rare alleles instead of the expected excess of rare alleles following a population bottleneck (*dcr2*: $D = 0.515$, $p > 0.1$, $F^* = 0.361$, $p > 0.1$; *AAEL006790*: $D = 1.04$, $p > 0.1$, $F^* = 1.26$, $p > 0.1$; *AAEL006800*: $D = -0.09$, $p > 0.1$, $F^* = 0.795$, $p > 0.1$).

In the three isofemale families, parental haplotypes could be inferred from the relative proportions of genotype combinations at the three loci (see the electronic supplementary material, figure S1). In families A and B ($n = 125$ and $n = 112$, respectively), individuals were partitioned into four clusters of genotypes with approximately equal frequencies that were consistent with a single possible combination of parental haplotypes (see the electronic supplementary material, figure S1). Only one recombinant (0.8%) was observed in family A. In family C, 100 per cent of individuals ($n = 96$) had the same genotype at all three loci. Taken together, patterns of polymorphism confirmed that a single mating pair had founded each isofemale family.

Females in the outbred population were experimentally exposed to two genetically distinct (1.2% nucleotide divergence) DENV-1 isolates that were designated as RTB-138 and RTB-196. Out of the 17 different *dcr2* genotypes identified in the population, only the two most common *dcr2* genotypes (01–01 and 01–10) had large enough sample sizes ($n > 10$) to be analysed statistically. We analysed both the proportion of virus-exposed mosquitoes that became infected and the proportion of infected mosquitoes that developed a disseminated viral infection 14 days after the infectious blood meal as a function of the *dcr2* genotype (01–01 or 01–10), the virus isolate (RTB-138 or RTB-196) and the infectious dose (RTB-196 was used at two different concentrations in the artificial infectious blood meal). The only factor that significantly influenced the probability of infection was the *dcr2* genotype \times isolate interaction (likelihood-ratio $\chi^2 = 5.75$, d.f. = 1, $p = 0.0165$), indicating that (i) there was a significant genotype–phenotype association and (ii) this association was virus isolate-specific. Including the two flanking genes in the model did not change the significance of the *dcr2* genotype \times isolate interaction (LR $\chi^2 = 5.04$, d.f. = 1, $p = 0.0247$). This interaction remained the only significant effect when *dcr2* genotypes were defined by the most informative individual SNP alone (i.e. with the highest minor allele frequency) at the *dcr2* locus (LR

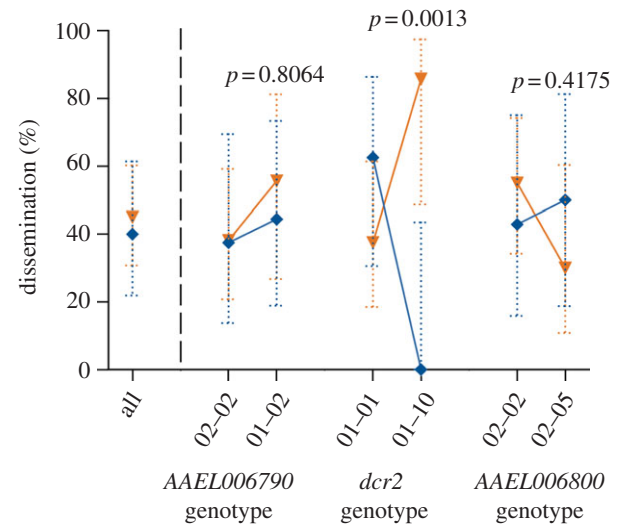


Figure 3. Isolate- and locus-specific association between viral dissemination and *dcr2* genotype in the outbred mosquito population. Interaction plot showing the percentage of mosquitoes in the outbred population that developed a disseminated viral infection by two different DENV isolates (blue, RTB-138; orange, RTB-196) as a function of the two most common genotypes at the *dcr2* locus and at two control flanking loci (right-hand side of the dashed vertical line), compared with the percentage of dissemination in the entire population (left-hand side of the dashed vertical line). Dotted, vertical bars are 95% CI of the percentages. P -values above the graph indicate the statistical significance of the genotype \times isolate interaction term in a multifactorial logistic regression accounting for the effects of infectious dose, isolate and genotype. Note that for simplicity, uninfected individuals are included in this analysis so that dissemination prevalence is a composite phenotype reflecting both midgut infection and viral dissemination. Separate analyses showed a similar effect of the genotype \times isolate interaction on both virus infection and dissemination probabilities.

$\chi^2 = 5.75$, d.f. = 1, $p = 0.0165$). Among infected mosquitoes (excluding uninfected), the only factor that significantly influenced the probability of dissemination was again the *dcr2* genotype \times isolate interaction (LR $\chi^2 = 4.57$, d.f. = 1, $p = 0.0325$). The result was robust when using the most informative individual SNP alone (LR $\chi^2 = 4.57$, d.f. = 1, $p = 0.0325$). Because the interaction effect resulted in a similar pattern for infection and dissemination, for the sake of simplicity, the cumulative probability of infection + dissemination is shown in figure 3. Although the infection + dissemination probability of both isolates was similar (approx. 50%) for mosquitoes with genotype 01–01, it was strikingly different for mosquitoes with genotype 01–10 (figure 3). Among mosquitoes with genotype 01–10, none had a disseminated infection with isolate RTB-138, whereas almost 90 per cent had a disseminated infection with isolate RTB-196.

Because of the significant LD, distribution overlap between the two most common *dcr2* genotypes and the two most common genotypes at the control loci was 76.3 per cent for *AAEL006790* and 88.1 per cent for *AAEL006800*. In contrast to the *dcr2* locus, however, no genotype \times isolate interaction associated with infection probability was detected at either control locus when considering their two most common genotypes (*AAEL006790*: LR $\chi^2 = 0.26$, d.f. = 1, $p = 0.6090$; *AAEL006800*: LR $\chi^2 = 0.05$, d.f. = 1, $p = 0.8186$) despite an increase in statistical power due to larger sample sizes ($n = 61$, $n = 74$ and $n = 79$ for *dcr2*, *AAEL006790* and *AAEL006800*, respectively). The genotype \times isolate

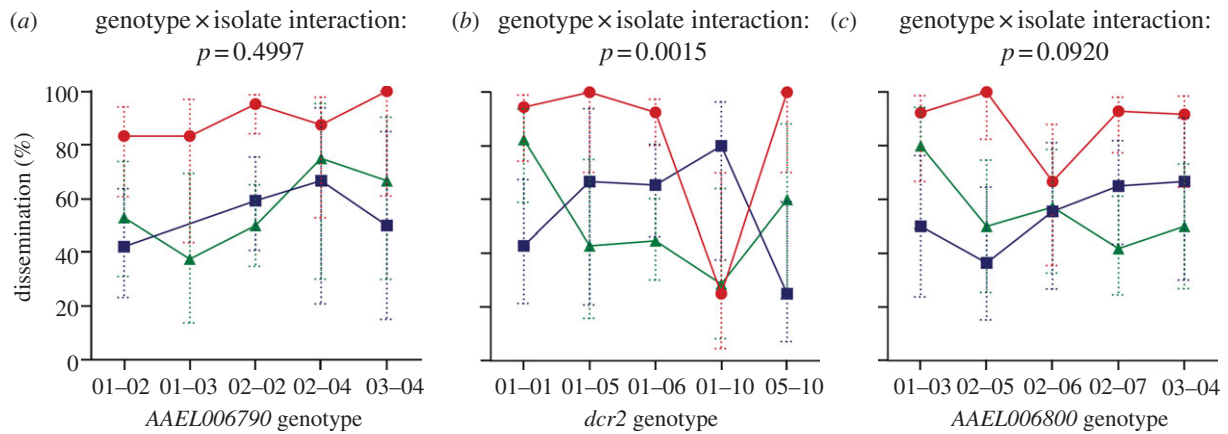


Figure 4. Isolate- and locus-specific association between viral dissemination and *dcr2* genotype in the mosquito families. Interaction plot showing the percentage of infected mosquitoes in the three isofemale families that developed a disseminated viral infection from three different DENV isolates (green, BKK; blue, KPP; red, RTB) as a function of the genotype at the *dcr2* locus (b) and at the two control loci (a,c). Dotted, vertical bars are 95% CI of the percentages. *P*-values above the graphs indicate the statistical significance of the genotype \times isolate interaction term in a two-way logistic regression accounting for the effects of isolate and genotype. Genotypes from the three families are pooled because the main effect of family was insignificant and thus removed from the analysis. Note that for purposes of clarity, one data point (KPP isolate, genotype 01–03) was omitted in (a) because it consisted of only two individuals.

interaction remained insignificant at each control locus when the genotypes were defined by the most informative individual SNP alone (AAEL006790: LR $\chi^2 = 0.26$, d.f. = 1, $p = 0.6090$; AAEL006800: LR $\chi^2 = 0.05$, d.f. = 1, $p = 0.8186$). Likewise, when the analysis of dissemination probability (excluding uninfected females) was performed for the two most common genotypes of each control locus, no significant genotype \times isolate interaction was detected (AAEL006790: LR $\chi^2 < 0.01$, d.f. = 1, $p > 0.9999$; AAEL006800: LR $\chi^2 = 2.32$, d.f. = 1, $p = 0.1280$). The genotype \times isolate interaction remained insignificant when the analysis of dissemination at control loci was performed with the most informative individual SNP alone (AAEL006790: LR $\chi^2 < 0.01$, d.f. = 1, $p > 0.9999$; AAEL006800: LR $\chi^2 = 2.32$, d.f. = 1, $p = 0.1280$). In combination, our analysis demonstrated that among the three neighbour genes examined, the genotype \times isolate interaction was specific to *dcr2* (figure 3).

Females in the isofemale families derived from the outbred *Ae. aegypti* population were exposed to three different, genetically distinct (2.4% overall nucleotide divergence) DENV-1 isolates designated RTB, BKK and KPP. Analysis of the proportion of infected females including the effects of family, isolate, *dcr2* genotype (nested within the families because each family contained a different set of *dcr2* genotypes; electronic supplementary material, figure S1) and their interactions confirmed the significant effect of the family \times isolate interaction (LR $\chi^2 = 13.4$, d.f. = 4, $p = 0.0096$) previously reported [5], but detected no influence of the *dcr2* genotype (LR $\chi^2 = 2.29$, d.f. = 4, $p = 0.6825$) or its interaction with the isolate (LR $\chi^2 = 6.20$, d.f. = 8, $p = 0.6302$). An initial genotype–phenotype association analysis of dissemination (excluding uninfected females) that included the effects of family, isolate and *dcr2* genotype (nested within family) revealed a strong main effect of the isolate (LR $\chi^2 = 29.5$, d.f. = 2, $p < 0.0001$), but no overall influence of the family (LR $\chi^2 = 0.69$, d.f. = 2, $p = 0.7077$). In this analysis, the genotype \times isolate interaction (nested within the family) was statistically significant (LR $\chi^2 = 23.8$, d.f. = 8, $p = 0.0025$), indicating that within each separate family, there was an isolate-specific genotype–phenotype

association. Because the main family effect was not a significant predictor of the phenotype, it was removed from the subsequent analysis. Thus, all *dcr2* genotypes were pooled regardless of family. In this analysis, again the probability of virus dissemination was significantly influenced by the *dcr2* genotype \times isolate interaction (LR $\chi^2 = 25.0$, d.f. = 8, $p = 0.0015$). Including the two flanking genes in the model did not change the significance of the *dcr2* genotype \times isolate interaction (LR $\chi^2 = 25.4$, d.f. = 8, $p = 0.0013$). Analysis of the most informative individual SNP at the *dcr2* locus confirmed a significant genotype \times isolate interaction (LR $\chi^2 = 15.0$, d.f. = 4, $p = 0.005$). By contrast, no significant genotype \times isolate interaction was detected when the same analysis was performed at either one of the control loci (AAEL006790: LR $\chi^2 = 7.35$, d.f. = 8, $p = 0.4997$; AAEL006800: LR $\chi^2 = 13.6$, d.f. = 8, $p = 0.0920$) despite an identical statistical power due to equal sample size and genotype number at all three loci. The genotype \times isolate interaction remained insignificant at the control loci when genotypes were defined by the most informative individual SNP (AAEL006790: LR $\chi^2 = 1.17$, d.f. = 2, $p = 0.557$; AAEL006800: LR $\chi^2 = 4.15$, d.f. = 2, $p = 0.126$). Analyses performed with the isofemale families confirmed the conclusion obtained from the parental outbred population; i.e. that isolate-specific resistance is associated with polymorphism at the *dcr2* locus, but not with other polymorphisms present in the same chromosomal region (figure 4).

The ranking order of isolates did not strongly differ across genotypes for the two control loci, with isolate RTB resulting in the highest percentage of viral dissemination (approx. 90%) and isolates KPP and BKK achieving a similar level (approx. 50%). By contrast, there was a marked difference in the ranking order of isolates across *dcr2* genotypes (figure 4). This strong interaction appeared to be primarily driven by genotype 01–10, which conferred strong resistance against isolate RTB (approx. 25% dissemination versus >90% for other *dcr2* genotypes). Genotype 01–10 was not a universal resistance genotype, however, because approximately 80 per cent of mosquitoes with this genotype had a disseminated infection with isolate KPP. Thus, the isolate- and

locus-specific association detected between the *dcr2* genotype and the viral dissemination phenotype in the outbred population was also found in the isofemale families exposed to three additional DENV-1 isolates.

4. Discussion

Although the occurrence of specific genetic interactions between invertebrate hosts and their pathogens is ubiquitous in nature, the underlying molecular basis has rarely been defined. In the present study, we document a statistical association consistent with a genotype-by-genotype interaction between the mosquito antiviral gene *dcr2* and DENV. The *dcr2* locus encodes the ribonuclease Dicer-2 that acts upstream of the RNAi pathway, an important antiviral defence mechanism of plants and invertebrates [39]. Dicer-2 recognizes and cleaves long double-stranded RNA (dsRNA) molecules resulting from secondary structures or replication intermediates. It processes viral dsRNA into 21–23 bp small interfering RNAs (siRNAs), which are incorporated into an RNA-induced silencing complex (RISC), leading to sequence-specific degradation of the target viral RNA [22]. Transient silencing of key genes of the RNAi pathway, including *dcr2*, results in a significant increase in DENV-2 titres in *Ae. aegypti* mosquitoes [24]. Constitutive impairment of the RNAi pathway also results in increased dissemination of Sindbis virus from the midgut of *Ae. aegypti* [40]. In *Drosophila*, Dicer-2 displays an exceptionally elevated rate of amino acid substitution, suggestive of rapid evolution through natural selection by RNA viruses [41,42]. It is worth noting that the low DENV prevalence in wild *Ae. aegypti* populations [43] and a relatively modest fitness cost of infection [44] make it unlikely that *dcr2* polymorphism is primarily shaped by DENV-mediated selection.

Our analyses were based on the same tests of genotype–phenotype associations at one candidate gene, *dcr2* and two control genes flanking the *dcr2* locus. Although statistical power was identical or greater for the control genes than for the candidate gene, *dcr2* was the only locus for which the interaction between mosquito genotype and virus isolate significantly influenced the outcome of infection. Our results point to the RNAi pathway as a possibly important determinant of intrinsic insect-virus genetic specificity. This is one of the few candidate immune genes underlying strain-specific resistance identified in a natural invertebrate population. This finding highlights the difficulty associated with the genetic mapping of a locus involved in strain-specific resistance. Because the effect of such a locus is most apparent by comparing resistance patterns across several pathogen strains, it will often go undetected by conventional mapping strategies that use a single pathogen strain [45]. Mapping loci involved in specific host–pathogen interactions requires that multiple combinations of host and pathogen genotypes are considered simultaneously [13].

Although alternatives are unlikely, our study design does not allow us to conclusively assign a causal role to the *dcr2* locus because the statistical association could result from undetected LD between *dcr2* and the causal polymorphism(s). Below, we discuss three potential scenarios. (i) Large chromosomal inversions may link physically distant genetic loci. Existence of inversion polymorphisms has been hypothesized in *Ae. aegypti* because some genomic regions have a

reduced recombination rate [46]. Direct evidence for segregating chromosomal inversions, however, has only been detected in the forest form *Ae. aegypti formosus*, which is found in West Africa [47]. To the best of our knowledge, segregating chromosomal inversions have not been detected among wild populations of the domestic form *Ae. aegypti aegypti* that we used in our study. If *dcr2* is located within a chromosomal inversion, flanking genes would remain valid controls. In the unlikely event that the inversion breakpoint was located precisely at the *dcr2* locus, at least one of the two flanking genes would remain a valid control locus. In contrast to this hypothesis, both flanking loci showed similar levels of LD with *dcr2* in the outbred population and we failed to detect the genotype–phenotype association found at the *dcr2* locus. (ii) A demographic bottleneck during laboratory colonization of the mosquito population may have artificially increased LD. Maintaining a large population size at each generation since the initial field collection, however, should have limited genetic drift during the colonization process. Accordingly, population genetic parameters in the outbred population did not display any signature of a recent demographic bottleneck. (iii) Long-range LD may exist between *dcr2* and a physically unlinked locus due to epistatic selection [48]. In this case, the *dcr2* locus would be in stronger LD with one or more distant gene(s) than with physically close loci because of functional relationships between the genes. Even under this scenario, however, our conclusion that polymorphism at the *dcr2* locus influences the phenotype would remain valid.

We do not expect that the *dcr2* locus is solely responsible for the variation in *Ae. aegypti* resistance to DENV, which is known to be a polygenic trait [27]. In the present study, a variance component analysis estimated that the specific *dcr2* genotype \times virus isolate interaction explained 17.8 per cent of the total variance in the probability of viral dissemination among the isofemale families. The existence of other genetic factors influencing mosquito resistance to DENV may help to explain why in the isofemale families *dcr2* genotype was only associated with isolate-specific probability of viral dissemination from the midgut to the rest of the mosquito body, but not with midgut infection probability. In the outbred population, *dcr2* genotype influenced both infection of midgut epithelial cells and subsequent dissemination to secondary tissues. We speculate that the isofemale families may have captured a subset of polymorphisms at other genetic loci involved in resistance to DENV that were not fully representative of the outbred population.

The underlying mechanism of virus isolate-specific resistance of different *dcr2* genotypes remains to be characterized. We hypothesize that non-synonymous polymorphisms within *dcr2* may result in differential dsRNA binding affinities of the Dicer-2 protein for particular dsRNA sequences and thus differential recognition of particular viral strains, leading to variation in the efficiency of RNAi-mediated antiviral activity. This hypothesis is supported by results of an *in vitro* study that detected significant variation among different DENV strains in their sensitivity to Dicer-2 knock-down [49]. Alternatively, different viral RNAi suppressors may differentially affect the activity of Dicer-2 allelic variants [23,50,51]. Future studies will dissect the mechanistic basis of Dicer-2 implication in insect-virus genetic specificity.

The authors thank Apolline Pichon for help with mosquito genotyping and three anonymous reviewers for helpful comments. L.L. was supported by Marie Curie Outgoing International Fellowship MOIF-CT-2006-039855 from the Sixth Framework Programme of the European Commission and ANR-09-RPDOC-007-01 from the French Agence Nationale de la Recherche. C.C. was funded by the French governmental agencies Centre National de la

Recherche Scientifique (CNRS) and Institut de Recherche pour le Développement. This work received partial support from the CNRS programme 'Maladies Infectieuses Emergentes'. The opinions or assertions contained herein are the private views of the authors and are not to be construed as reflecting the official views of the United States Army, Royal Thai Army or the United States Department of Defense.

References

- Schmid-Hempel P, Ebert D. 2003 On the evolutionary ecology of specific immune defence. *Trends Ecol. Evol.* **18**, 27–32. (doi:10.1016/S0169-5347(02)00013-7)
- Little TJ, Hultmark D, Read AF. 2005 Invertebrate immunity and the limits of mechanistic immunology. *Nat. Immunol.* **6**, 651–654. (doi:10.1038/ni1219)
- Carius HJ, Little TJ, Ebert D. 2001 Genetic variation in a host–parasite association: potential for coevolution and frequency-dependent selection. *Evolution* **55**, 1136–1145.
- de Roode JC, Altizer S. 2009 Host–parasite genetic interactions and virulence–transmission relationships in natural populations of Monarch butterflies. *Evolution* **64**, 502–514. (doi:10.1111/j.1558-5646.2009.00845.x)
- Lambrechts L, Chevillon C, Albright RG, Thaisomboonsuk B, Richardson JH, Jarman RG, Scott TW. 2009 Genetic specificity and potential for local adaptation between dengue viruses and mosquito vectors. *BMC Evol. Biol.* **9**, 160. (doi:10.1186/1471-2148-9-160)
- Lambrechts L, Halbert J, Durand P, Gouagna LC, Koella JC. 2005 Host genotype by parasite genotype interactions underlying the resistance of anopheline mosquitoes to *Plasmodium falciparum*. *Malar. J.* **4**, 3. (doi:10.1186/1475-2875-4-3)
- Schmid-Hempel P, Pühr K, Krüger N, Reber C, Schmid-Hempel R. 1999 Dynamic and genetic consequences of variation in horizontal transmission for a microparasitic infection. *Evolution* **53**, 426–434. (doi:10.2307/2640779)
- Schulenburg H, Ewbank JJ. 2004 Diversity and specificity in the interaction between *Caenorhabditis elegans* and the pathogen *Serratia marcescens*. *BMC Evol. Biol.* **4**, 49. (doi:10.1186/1471-2148-4-49)
- Webster JP, Woolhouse MEJ. 1998 Selection and strain specificity of compatibility between snail intermediate hosts and their parasitic schistosomes. *Evolution* **52**, 1627–1634. (doi:10.2307/2411336)
- Luijckx P, Ben-Ami F, Mouton L, Du Pasquier L, Ebert D. 2011 Cloning of the unculturable parasite *Pasteuria ramosa* and its *Daphnia* host reveals extreme genotype–genotype interactions. *Ecol. Lett.* **14**, 125–131. (doi:10.1111/j.1461-0248.2010.01561.x)
- Lemaitre B, Reichhart JM, Hoffmann JA. 1997 *Drosophila* host defense: differential induction of antimicrobial peptide genes after infection by various classes of microorganisms. *Proc. Natl Acad. Sci. USA* **94**, 14 614–14 619. (doi:10.1073/pnas.94.26.14614)
- Schmid-Hempel P. 2005 Natural insect host–parasite systems show immune priming and specificity: puzzles to be solved. *Bioessays* **27**, 1026–1034. (doi:10.1002/bies.20282)
- Lambrechts L. 2010 Dissecting the genetic architecture of host–pathogen specificity. *PLoS Pathog.* **6**, e1001019. (doi:10.1371/journal.ppat.1001019)
- De Wit PJ, Mehrabi R, Van den Burg HA, Stergiopoulos I. 2009 Fungal effector proteins: past, present and future. *Mol. Plant Pathol.* **10**, 735–747. (doi:10.1111/j.1364-3703.2009.00591.x)
- Ellis JG, Dodds PN, Lawrence GJ. 2007 Flax rust resistance gene specificity is based on direct resistance–avirulence protein interactions. *Annu. Rev. Phytopathol.* **45**, 289–306. (doi:10.1146/annurev.phyto.45.062806.094331)
- Gilbert SC *et al.* 1998 Association of malaria parasite population structure, HLA, and immunological antagonism. *Science* **279**, 1173–1177. (doi:10.1126/science.279.5354.1173)
- Kubinak JL, Ruff JS, Hyzer CW, Slev PR, Potts WK. 2012 Experimental viral evolution to specific host MHC genotypes reveals fitness and virulence trade-offs in alternative MHC types. *Proc. Natl Acad. Sci. USA* **109**, 3422–3427. (doi:10.1073/pnas.1112633109)
- Carpenter JA, Hadfield JD, Bangham J, Jiggins FM. 2012 Specific interactions between host and parasite genotypes do not act as a constraint on the evolution of antiviral resistance in *Drosophila*. *Evolution* **66**, 1114–1125. (doi:10.1111/j.1558-5646.2011.01501.x)
- Fleuriet A. 1980 Polymorphism of the hereditary sigma virus in natural populations of *Drosophila melanogaster*. *Genetics* **95**, 459–465.
- Harris C, Lambrechts L, Rousset F, Abate L, Nsango SE, Fontenille D, Morlais I, Cohuet A. 2010 Polymorphisms in *Anopheles gambiae* immune genes associated with natural resistance to *Plasmodium falciparum*. *PLoS Pathog.* **6**, e1001112. (doi:10.1371/journal.ppat.1001112)
- Mone Y, Gourbal B, Duval D, Du Pasquier L, Kieffer-Jaquinod S, Mitta G. 2010 A large repertoire of parasite epitopes matched by a large repertoire of host immune receptors in an invertebrate host/parasite model. *PLoS Negl. Trop. Dis.* **4**, e813. (doi:10.1371/journal.pntd.0000813)
- Bernstein E, Caudy AA, Hammond SM, Hannon GJ. 2001 Role for a bidentate ribonuclease in the initiation step of RNA interference. *Nature* **409**, 363–366. (doi:10.1038/35053110)
- Galiana-Arnoux D, Dostert C, Schneemann A, Hoffmann JA, Imler JL. 2006 Essential function *in vivo* for Dicer-2 in host defense against RNA viruses in *Drosophila*. *Nat. Immunol.* **7**, 590–597. (doi:10.1038/ni1335)
- Sanchez-Vargas I, Scott JC, Poole-Smith BK, Franz AW, Barbosa-Solomieu V, Wilusz J, Olson KE, Blair CD. 2009 Dengue virus type 2 infections of *Aedes aegypti* are modulated by the mosquito's RNA interference pathway. *PLoS Pathog.* **5**, e1000299. (doi:10.1371/journal.ppat.1000299)
- Wang XH, Aliyari R, Li WX, Li HW, Kim K, Carthew R, Atkinson P, Ding SW. 2006 RNA interference directs innate immunity against viruses in adult *Drosophila*. *Science* **312**, 452–454. (doi:10.1126/science.1125694)
- Gubler DJ. 1998 Dengue and dengue hemorrhagic fever. *Clin. Microbiol. Rev.* **11**, 480–496.
- Black WC, Bennett KE, Gorrochotegui-Escalante N, Barillas-Mury CV, Fernandez-Salas I, de Lourdes Munoz M, Farfan-Ale JA, Olson KE, Beaty BJ. 2002 Flavivirus susceptibility in *Aedes aegypti*. *Arch. Med. Res.* **33**, 379–388. (doi:10.1016/S0188-4409(02)00373-9)
- Lynch M, Walsh B. 1998 *Genetics and analysis of quantitative traits*. Sunderland, MA: Sinauer Associates.
- Martin DP, Lemey P, Lott M, Moulton V, Posada D, Lefevre P. 2010 RDP3: a flexible and fast computer program for analyzing recombination. *Bioinformatics* **26**, 2462–2463. (doi:10.1093/bioinformatics/btq467)
- Tamura K, Peterson D, Peterson N, Stecher G, Nei M, Kumar S. 2011 MEGA5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. *Mol. Biol. Evol.* **28**, 2731–2739. (doi:10.1093/molbev/msr121)
- Drummond AJ, Ashton B, Cheung M, Heled J, Kearse M, Moir R, Stones-Havas S, Thierer T, Wilson A. 2009 GENEIOUS v4.8. See <http://www.geneious.com/>.
- Librado P, Rozas J. 2009 DnaSP v5: a software for comprehensive analysis of DNA polymorphism data. *Bioinformatics* **25**, 1451–1452. (doi:10.1093/bioinformatics/btp187)
- Raymond M, Rousset F. 1995 GENETPOP (version 1.2): population genetics software for exact tests and ecumenicism. *J. Hered.* **86**, 248–249.
- Li N, Stephens M. 2003 Modeling linkage disequilibrium and identifying recombination hotspots using single-nucleotide polymorphism data. *Genetics* **165**, 2213–2233.

35. Cornuet JM, Luikart G. 1996 Description and power analysis of two tests for detecting recent population bottlenecks from allele frequency data. *Genetics* **144**, 2001–2014.
36. Piry S, Luikart G, Cornuet JM. 1999 Bottleneck: a computer program for detecting recent reduction in the effective population size using allele frequency data. *J. Hered.* **90**, 502–503. (doi:10.1093/jhered/90.4.502)
37. Nene V *et al.* 2007 Genome sequence of *Aedes aegypti*, a major arbovirus vector. *Science* **316**, 1718–1723. (doi:10.1126/science.1138878)
38. Severson DW, Meece JK, Lovin DD, Saha G, Morlais I. 2002 Linkage map organization of expressed sequence tags and sequence tagged sites in the mosquito, *Aedes aegypti*. *Insect Mol. Biol.* **11**, 371–378. (doi:10.1046/j.1365-2583.2002.00347.x)
39. van Rij RP, Berezikov E. 2009 Small RNAs and the control of transposons and viruses in *Drosophila*. *Trends Microbiol.* **17**, 163–171. (doi:10.1016/j.tim.2009.01.003)
40. Khoo CC, Piper J, Sanchez-Vargas I, Olson KE, Franz AW. 2010 The RNA interference pathway affects midgut infection- and escape barriers for Sindbis virus in *Aedes aegypti*. *BMC Microbiol.* **10**, 130. (doi:10.1186/1471-2180-10-130)
41. Obbard DJ, Gordon KH, Buck AH, Jiggins FM. 2009 The evolution of RNAi as a defence against viruses and transposable elements. *Phil. Trans. R. Soc. B* **364**, 99–115. (doi:10.1098/rstb.2008.0168)
42. Obbard DJ, Jiggins FM, Halligan DL, Little TJ. 2006 Natural selection drives extremely rapid evolution in antiviral RNAi genes. *Curr. Biol.* **16**, 580–585. (doi:10.1016/j.cub.2006.01.065)
43. Yoon IK *et al.* 2012 Fine scale spatiotemporal clustering of dengue virus transmission in children and *Aedes aegypti* in rural Thai villages. *PLoS Negl. Trop. Dis.* **6**, e1730. (doi:10.1371/journal.pntd.0001730)
44. Lambrechts L, Scott TW. 2009 Mode of transmission and the evolution of arbovirus virulence in mosquito vectors. *Proc. R. Soc. B* **276**, 1369–1378. (doi:10.1098/rspb.2008.1709)
45. Wilfert L, Schmid-Hempel P. 2008 The genetic architecture of susceptibility to parasites. *BMC Evol. Biol.* **8**, 187. (doi:10.1186/1471-2148-8-187)
46. Yan G, Christensen BM, Severson DW. 1997 Comparisons of genetic variability and genome structure among mosquito strains selected for refractoriness to a malaria parasite. *J. Hered.* **88**, 187–194. (doi:10.1093/oxfordjournals.jhered.a023087)
47. Bernhardt SA, Blair C, Sylla M, Bosio C, Black WC. 2009 Evidence of multiple chromosomal inversions in *Aedes aegypti formosus* from Senegal. *Insect Mol. Biol.* **18**, 557–569. (doi:10.1111/j.1365-2583.2009.00895.x)
48. Zapata C, Nunez C, Velasco T. 2002 Distribution of nonrandom associations between pairs of protein loci along the third chromosome of *Drosophila melanogaster*. *Genetics* **161**, 1539–1550.
49. Mukherjee S, Hanley KA. 2010 RNA interference modulates replication of dengue virus in *Drosophila melanogaster* cells. *BMC Microbiol.* **10**, 127. (doi:10.1186/1471-2180-10-127)
50. Aliyari R, Wu Q, Li HW, Wang XH, Li F, Green LD, Han CS, Li WX, Ding SW. 2008 Mechanism of induction and suppression of antiviral immunity directed by virus-derived small RNAs in *Drosophila*. *Cell Host Microbe* **4**, 387–397. (doi:10.1016/j.chom.2008.09.001)
51. van Rij RP, Saleh MC, Berry B, Foo C, Houk A, Antoniewski C, Andino R. 2006 The RNA silencing endonuclease Argonaute 2 mediates specific antiviral immunity in *Drosophila melanogaster*. *Genes Dev.* **20**, 2985–2995. (doi:10.1101/gad.1482006)