npg

ORIGINAL ARTICLE

Global spread and genetic variants of the two *CYP9M10* haplotype forms associated with insecticide resistance in *Culex quinquefasciatus* Say

K Itokawa[1], O Komagata[1], S Kasai[1], H Kawada[2], C Mwatele[3], GO Dida[3], SM Njenga[3], C Mwandawiro[3] and T Tomita[1]

Insecticide resistance develops as a genetic factor (allele) conferring lower susceptibility to insecticides proliferates within a target insect population under strong positive selection. Intriguingly, a resistance allele pre-existing in a population often bears a series of further adaptive allelic variants through new mutations. This phenomenon occasionally results in replacement of the predominating resistance allele by fitter new derivatives, and consequently, development of greater resistance at the population level. The overexpression of the cytochrome P450 gene *CYP9M10* is associated with pyrethroid resistance in the southern house mosquito *Culex quinquefasciatus*. Previously, we have found two genealogically related overexpressing *CYP9M10* haplotypes, which differ in gene copy number (duplicated and non-duplicated). The duplicated haplotype was derived from the non-duplicated overproducer probably recently. In the present study, we investigated allelic series of *CYP9M10* involved in three *C. quinquefasciatus* laboratory colonies recently collected from three different localities. Duplicated and non-duplicated overproducing haplotypes coexisted in African and Asian colonies indicating a global distribution of both haplotype lineages. The duplicated haplotypes both in the Asian and African colonies were associated with higher expression levels and stronger resistance than non-duplicated overproducing haplotypes. There were slight variation in expression level among the non-duplicated overproducing haplotypes. The nucleotide sequences in coding and upstream regions among members of this group also showed a little diversity. Non-duplicated overproducing haplotypes with relatively higher expression were genealogically closer to the duplicated haplotypes than the other non-duplicated overproducing haplotypes, suggesting multiple *cis*-acting mutations before duplication.

INTRODUCTION

One of the most important observations regarding research on insecticide resistance during the last several decades may be the unforeseen dynamic nature of resistance genes, such that a resistance allele formerly dominant in a population is replaced by a fitter allele on a relatively short time scale (Guillemaud *et al.*, 1998; Field and Foster, 2002; Yebakima *et al.*, 2004; Labbe *et al.*, 2009). Although a fitter resistance allele in this context is usually an allele conferring greater resistance than the preceding allele, it can also be a resistance allele with lower fitness cost in the absence of insecticide pressure. One intriguing question raised by this phenomenon concerns the origin of the new fitter resistance allele. In some cases, the fitter alleles appear to originate from a genealogically independent lineage and not from the preceding resistance allele (Guillemaud *et al.*, 1998; Field and Foster, 2002; Hartley *et al.*, 2006). In other cases, the fitter resistance allele seems to be derived via modification of a pre-existing resistance allele (Berticat *et al.*, 2001; Schmidt *et al.*, 2010). In either case, standing or newly arising allelic variations in a locus related to insecticide resistance provide a source for such dynamics among resistance genes.

The overproducing haplotype of a cytochrome P450 (P450) gene *CYP6G1* in *Drosophila melanogaster* known to confer DDT resistance (Daborn *et al.*, 2002) is well elucidated with respect to its history of the progressive evolution (Schmidt *et al.*, 2010). *CYP6G1* encodes a cytochrome P450, which is a metabolic enzyme known to be involved in insecticide detoxification (Feyereisen, 2012). The first adaptive mutation in the *CYP6G1* haplotype lineage seems to have been the insertion of a long terminal repeat retrotransposable element, *Accord*, in the 5′ upstream region (Schmidt *et al.*, 2010), causing overexpression of *CYP6G1* (Chung *et al.*, 2007). Following the *Accord* insertion, duplication occurred in this haplotype lineage and distributed itself worldwide in the current population of *D. melanogaster* (Schmidt *et al.*, 2010). The fact that no isolate of an *Accord*-inserted haplotype without the duplication has been identified (Schmidt *et al.*, 2010) suggests that this adaptive duplication has been subjected to selective sweep at the global level. Other *cis*-acting regulatory mutations also seem to have occurred after the duplication (Schmidt *et al.*, 2010), indicating that *CYP6G1* acquired higher penetrance to the resistance phenotype through multiple mutations influencing its expression.

[1]Department of Medical Entomology, National Institute of Infectious Diseases, Tokyo, Japan; [2]Department of Vector Ecology and Environment, Institute of Tropical Medicine, Nagasaki University, Nagasaki, Japan and [3]Eastern and Southern Africa Centre of International Parasite Control, Kenya Medical Research Institute, Nairobi, Kenya
Correspondence: Dr T Tomita, Department of Medical Entomology, National Institute of Infectious Diseases, Toyama 1-23-1, Shinjuku-ku, Tokyo 162-8640, Japan.
E-mail: tomita@nih.go.jp

We have suggested a similar progressive evolution in another P450 gene in the southern house mosquito *Culex quinquefasciatus*. The JPal-per (JPP) strain of *C. quinquefasciatus* was established from a population in Saudi Arabia during 1981 with successive selection by the pyrethroid insecticide permethrin (Amin and Hemingway, 1989). Kasai *et al.* (1998) found that P450-mediated detoxification was involved in the resistance of the JPP strain. Degenerate PCR was used to isolate cDNAs of P450 genes expressed in the JPP strain (Komagata *et al.*, 2010), and subsequent microarray analysis detected *CYP9M10* as the most prominently overexpressed P450 gene in the JPP stain (compared with a susceptible strain). Overexpression of *CYP9M10* has also been observed in another resistance strain HAmCq^{G08} (Liu *et al.*, 2011), which originated in the USA during 2002 (Liu *et al.*, 2004). *CYP9M10* locus was genetically linked to the resistance phenotype (Hardstone *et al.*, 2010; Itokawa *et al.*, 2010), and more recently, Wilding *et al.* (2012) have shown the ability of *CYP9M10* to metabolize permethrin.

Although the mechanism underlying the overtranscription of *CYP9M10* is not completely clear, some insights have been obtained from studies on the JPP strain. First, the factor causing the over-transcription is *cis*-acting (Itokawa *et al.*, 2010). Second, the *CYP9M10* haplotype in the JPP strain carries a tandem duplication of a 100-kb region that includes the *CYP9M10* locus and several other genes (Itokawa *et al.*, 2010). The coding and flanking regions of *CYP9M10* in the two duplicated copies have completely identical sequences, indicating that the duplication has occurred recently. Despite the large size of the duplicated unit, the junction created by the duplication is located at only ∼1.1-kb upstream of the putative transcriptional start site of *CYP9M10*. This proximity of the junction to the coding region enables distinguishing between the two *CYP9M10* duplicants (Itokawa *et al.*, 2010). Although the duplication may contribute to the overtranscription of *CYP9M10* in JPP, it alone cannot explain the ∼200-fold overtranscription in JPP compared with a susceptible laboratory strain, Ogasawara (OGS) (Komagata *et al.*, 2010). Another notable feature of the *CYP9M10* haplotype in the JPP strain is the insertion of a 0.6-kb sequence element at 0.2-kb upstream from the putative transcription initiation site (Itokawa *et al.*, 2010). This element is a member of the repetitive *CuRE1* (Culex repetitive element 1) family (Itokawa *et al.*, 2010) that belongs to the MITE (<u>m</u>iniature <u>i</u>nverted-repeat <u>t</u>ransposable <u>e</u>lement) class of transposable elements. Although *CuRE1* is currently annotated as DNA-TA-5_CQ in Repbase (www.girinst.org/repbase/) (Kojima and Jurka, 2011), we use '*CuRE1*' here for consistency with previous researches. As the same *CuRE1* insertion exists in both duplicated copies, its insertion event should predate the duplication.

After the duplicated *CYP9M10* haplotype was found in the JPP strain, another haplotype variant of *CYP9M10* with the *CuRE1* insertion, but not duplicated, was found in the JHB-NIID-B (JNB) strain, a substrain of the JHB strain originating in South Africa during 2001 (Arensburger *et al.*, 2010; Itokawa *et al.*, 2011). Importantly, the expression level of *CYP9M10* in the JNB strain was 25-fold higher than that of the OGS strain, but 8-fold lower than that of the JPP strain and the *CYP9M10* haplotype in the JNB strain was associated with moderate resistance to permethrin (Itokawa *et al.*, 2011). A comparison of the sequences detected only three nucleotide differences between the haplotypes in the JPP and JNB strains (one within the 1.9-kb transcribed region and two within the 2.9-kb upstream flanking region) indicating that the two haplotypes share a relatively recent common ancestral haplotype. The JNB haplotype lineage was considered to have diverged from the JPP haplotype lineage before the duplication (Itokawa *et al.*, 2011). As the difference

in expression levels of the two haplotypes was significantly higher than the two-fold difference in copy number (Itokawa *et al.*, 2011), additional *cis*-acting regulatory mutation(s) as well as duplication could be involved in the difference of the expression levels between the JPP and JNB haplotypes.

Although the insertion of a transposable element such as *Accord* is potentially able to change the regulation of a nearby gene, it is not clear whether *CuRE1* itself causes the overproduction of *CYP9M10*. A reporter assay using an *Anopheles gambiae* cell line detected enhancer activity within the 1.3-kb upstream region of the *CYP9M10* JPP allele (a region common to the both duplicated copies), but exact deletion of the *CuRE1* from this sequence did not affect the reporter gene expression (Wilding *et al.*, 2012). Nonetheless, the *CuRE1* insertion is clearly associated with a decreased level of pyrethroid susceptibility within local populations of *C. quinquefasciatus* in Ghana (Wilding *et al.*, 2012), indicating that despite its inactivity in regulatory assays, the insertion may serve as a genetic marker associated with the actual resistance-conferring mutation.

As shown in the study of Schmidt *et al.* (2010), allelic series in field and laboratory populations can be useful in investigating the distribution and evolutionary history of particular haplotype lineages. As our studies for this paper initiated, three *C. quinquefasciatus* colonies collected from Kenya, Singapore and Vietnam had been inbred for <1 year (Table 1). The *CuRE1*-inserted haplotypes of both duplicated and non-duplicated forms were found in the Kenya and Singapore colonies and the *CuRE1*-inserted non-duplicated form was found in the Vietnam colony (Supplementary Table S1), indicating that both forms of the *CuRE1*-inserted *CYP9M10* haplotypes have spread through Asia and Africa. The duplicated haplotypes in the two colonies were both associated with greater resistance than the *CuRE1*-inserted non-duplicated haplotypes of each colony. The duplicated haplotypes showed the highest expression levels of *CYP9M10* mRNA, whereas the non-duplicated *CuRE1*-inserted haplotypes showed intermediate expression levels. We observed a variation in expression among the *CuRE1*-inserted non-duplicated haplotypes. Interestingly, those with relatively higher expression levels were genealogically closer to the duplicated haplotypes than the other *CuRE1*-inserted non-duplicated haplotypes.

MATERIALS AND METHODS
Mosquito colonies and strains

Three colonies of *C. quinquefasciatus* originating in Kenya (2009; KNY09), Singapore (2009; SNG09) and Vietnam (2010; VTN10) were used in this study. The PCR assay developed by Smith and Fonseca (2004) confirmed that they were all *C. quinquefasciatus* rather than other known members of the *Culex pipiens* species complex (data not shown). At the time of the experiment, each population had been maintained by inbreeding for <1 year (<12 generations) in the laboratory since its collection. The founders of the KNY09 colony were ∼100 larvae collected from several water pools or ponds within the region around Lake Victoria in the Nyanza province in western Kenya during 2009. The SNG09 colony was founded by only 4–5 larvae collected from a narrow gutter beside a street road in Ang Mo Kio, Singapore during March 2009. Both colonies exhibited extremely strong resistance to permethrin (Table 1). The VTN10 colony originated from ∼100 larvae collected from a drainage canal beside a road in Hanoi, Vietnam, in March 2010. As preliminary investigation, *CYP9M10* genotypes were investigated by genotyping PCR I and II (Itokawa *et al.*, 2010) for some specimens chosen from each colony in their second or third generation in the laboratory.

The four laboratory strains JPP, JNA, JNB and OGS had different susceptibilities to permethrin (Table 1) and different *CYP9M10* expression levels (Itokawa *et al.*, 2011). The OGS strain was fixed with a *CuRE1* non-inserted *CYP9M10* haplotype. *CYP9M10* in the OGS strain is considered nonfunctional because there is a single nucleotide deletion within the open

Table 1 Colonies and strains used in this study

Colony or strain	Locality	Year of collection	Description
KNY09	Nyanza, Kenya	2009	High resistance to permethrin (LC_{50}[a] = 6 ppm).
SNG09	Ang Mo Kio, Singapore	2009	High resistance to permethrin (LC_{50} = 50 ppm).
VTN10	Hanoi, Vietnam	2010	Moderate resistance to permethrin (LC_{50} = 0.034 ppm).
JPal-per (JPP)	Jeddah, Saudi Arabia	1981	High resistance to permethrin (LC_{50} = 10 ppm[b]). Fixed with D-Cu(+) haplotype[c].
JHB-NIIID-A (JNA)	Johannesburg, South Africa	2001	Susceptible to permethrin (LC_{50} = 0.0040 ppm[d]). Fixed with C(−) haplotype[d]. Subdivided form the JHB strain[d].
JHB-NIID-B (JNB)	Johannesburg, South Africa	2001	Moderate resistance to permethrin (LC_{50} = 0.014 ppm[d]). Fixed with C(+) haplotype[d]. Subdivided form the JHB strain[d].
Ogasawara (OGS)	Ogasawara, Japan	1968	Susceptible to permethrin (LC_{50} = 0.0040 ppm[b]). Fixed with C(−) haplotype with disruptive frame-shift mutation[c].

[a]50% lethal concentration.
[b]Kasai et al. (1998).
[c]Itokawa et al. (2010).
[d]Itokawa et al. (2011).

reading frame, resulting in frame shift (Itokawa et al., 2010). The Johannesburg (JHB) strain that was used for the recent C. quinquefasciatus genome project (Arensburger et al., 2010) has segregated both CuRE1 non-inserted and CuRE1-inserted CYP9M10 haplotypes (Itokawa et al., 2011). The JNA and JNB strains are sister strains that have been divided from the JHB strain (Itokawa et al., 2011) as being fixed with CuRE1 non-inserted and CuRE1-inserted CYP9M10 haplotypes, respectively. More detailed information on these strains is provided in the references cited in Table 1.

Conventions for specifying CYP9M10 haplotypes

In this paper, CYP9M10 haplotypes are classified using three classification cues 'form', 'allele' and 'origin.' The 'forms' are defined by the presence or absence of the CuRE1 insertion in the upstream region and duplication. Three forms of CYP9M10 haplotypes were considered: CuRE1 non-inserted (Cu(−)), CuRE1-inserted but non-duplicated (Cu(+)) and CuRE1-inserted and duplicated (D-Cu(+)). The 'alleles' were distinguished by the 1.6-kb coding sequence (CDS) of CYP9M10, that is, from start to stop codon excluding the 57-bp single intron (Itokawa et al., 2010). The D-Cu(+) haplotype could have two equal or different alleles as duplicated copies. The 'origin' indicates the colony or strain in which the haplotype was initially included. Although it is possible that identical-by-state haplotypes of the same origin are not identical-by-descent from the initial founders of each colony, here we do not further distinguish identical-by-state haplotypes of the same origin. Thus, the identity of each haplotype is specified using the combination of form, allele and origin: for example, 'D-Cu(+)[SNG09][*1–*3]' indicates a haplotype of the D-Cu(+) form with *1 and *3 alleles (as duplicated copies) included in the SNG09 colony.

Bioassays and genotyping of the KNY09-F1, SNG09-F1 and VTN10-F1 cohorts

The first-instar larvae of the OGS strain were exposed to $0.025\,mg\,ml^{-1}$ of tetracycline following the study of Portaro and Barr (1975) to eradicate the bacterial endosymbiont Wolbachia, which potentially causes cytoplasmic incompatibility in inter-strain crosses. Approximately 100 virgin females from each KNY09, SNG09, and VTN10 colony were mated with 20–30 newly hatched males from the OGS strain to generate the F1 cohorts KNY09-F1, SNG09-F1 and VTN10-F1, respectively. The larvae of each cohort were reared under our standard laboratory condition (25 °C, 16L:8D). Fourth-instar larvae were placed in a plastic cup containing distilled water (20–25 larvae per 50 ml). We then added 0.5 ml permethrin dissolved in ethanol at the designated concentration to start the bioassay. Forty to fifty larvae (two cups) were tested at each concentration. The end point of the bioassay, surviving or dead, was the ability of the larva to move up to the surface when stirred up after 24 h. After the bioassay, dead and surviving larvae from the same concentration group were preserved separately in acetone. Apparent dead larvae were collected before 24 h because leaving dead larvae in water for a long period can hinder the
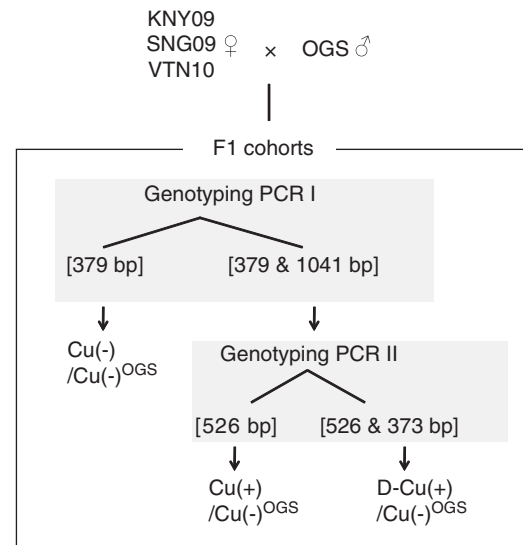

Figure 1 The scheme for genotyping individuals in KNY09-F1, SNG09-F1 and VTN10-F1. The size of expected products in genotyping PCR I and II are shown. The haplotypes inherited from the OGS males are indicated as Cu(−)[OGS].

extraction of DNA of sufficient quality for the downstream assays. The stored samples were dried in acetone and treated with a REDExtract-N-Amp Tissue PCR Kit (Sigma-Aldrich, St Louis, MO, USA) to be used as PCR template.

Genotyping PCR I and II (Itokawa et al., 2010) are genotyping methods for CYP9M10 in terms of the CuRE1 insertion and the duplication (here referred to as 'form' in this paper). Genotyping PCR I amplified the inserted CuRE1 element with its flanking regions. A longer (1041-bp) and a shorter (379-bp) fragments are amplified from CuRE1-inserted and non-inserted haplotypes, respectively (Supplementary Figure S1). Genotyping PCR II amplifies two fragments of different sizes from each of the two CYP9M10 copies in the duplicated haplotype (a longer 526-bp fragment from the upstream and a shorter 373-bp fragment from the downstream copy), thereby detecting the duplicated haplotype (Supplementary Figure S1). The genotypes of individual larvae in cohorts KNY09-F1, SNG09-F1 and VTN10-F1 were determined using these genotyping PCRs (Figure 1a). First, genotyping PCR I was applied to all samples. Samples yielding only the shorter fragment in genotyping PCR I were determined as the Cu(−)/Cu(−) genotype and the other samples, which yielded both longer and shorter fragments, were carried to the next genotyping PCR II. In genotyping PCR II, samples that yielded only the longer fragment were determined as the Cu(+)/Cu(−) genotype and the other samples, which

yielded both the longer and shorter fragments, were determined as the D-Cu(+)/Cu(−) genotype.

gDNA and RNA extraction, and quantitative PCR

Fourth-instar larvae were homogenized individually in 250 μl of ISOGEN (Nippongene, Tokyo, Japan), before adding 62.5 μl chloroform and centrifuging at 12 000 × g for 15 min at 4 °C. The upper aqueous phase was used to extract total RNA and synthesize cDNA, as described by Itokawa et al. (2010). Genomic DNA (gDNA) was precipitated from the remaining inter- and phenol–chloroform phases by addition of 75 μl of ethanol and centrifugation at 2000 × g for 5 min at 4 °C. The precipitate was washed twice with 70% ethanol and dissolved in 100 μl of TE buffer (pH 8.0). gDNA was used for genotyping by the previously mentioned methods.

cDNA was used to measure the relative expression level of CYP9M10 by the $2^{−\Delta\Delta Ct}$ method (Livak and Schmittgen, 2001), with ribosomal protein S3 gene (RPS3) as an internal control gene. The primers, reagents and PCR conditions were those described in the study by Komagata et al. (2010). Data in our previous study (Komagata et al., 2010) has showed that the amplification efficacies for CYP9M10 and RPS3 with under the condition are both nearly 1.00 (data not shown). One cDNA sample from the OGS strain was used as an inter-plate calibrator and was measured on every plate. The relative expression level, $2^{−\Delta\Delta Ct}$, was calculated based on the calibrator sample on the same plate. Each sample was measured twice and the mean value was used.

Sequencing alleles in the KNY09-F1, SNG09-F1 and VTN10-F1 cohorts inherited from their maternal colonies

Each individual of the cohorts KNY09-F1, SNG09-F1 and VTN10-F1 (see Results) inherited a CYP9M10 haplotype from each maternal population and another from the OGS strain. To isolate and analyze their maternal haplotypes, we first tried to design a PCR primer that may amplify only maternal haplotypes and not the OGS haplotype. By comparing the sequences of four known CYP9M10 coding and franking regions among the OGS, JPP, JNA and JNB strains, we designed a reverse primer P32R56, 5′-GGGACATAATGCA-TAATGTGCAGTA-3′, targeting a few base pairs downstream from the putative transcriptional termination site. The P32R56 primer was designed to include mismatches only to the OGS allele among the four haplotypes to prevent amplification from the OGS haplotype. A common forward primer, 5′-CACCTACATATTTAAGAACGCCG-3′, was designed to anneal 19-bp upstream from the putative transcriptional start site. PCR was conducted using these primers and KOD-FX DNA polymerase (Toyobo, Osaka, Japan) with the following amplification cycle: 94 °C for 1 min followed by 30 cycles of 94 °C for 15 s, 58 °C for 15 s, and 68 °C for 1 min and 30 s. A fragment encompassing the overall coding region of CYP9M10 was amplified by this PCR from the gDNA of the JPP, JNA and JNB strains, whereas no product was amplified from the gDNA of the OGS strain (data not shown). We applied this PCR to the 72 gDNA samples extracted from the individuals used in the expression analysis. The gDNA of all individuals of KNY09-F1 and SNG09-F1 produced products with the expected size. However, among the 24 individuals of VTN09-F1, the PCR amplified products from only 15 individuals. The nine individuals of VTN09-F1 from which we could not obtain a PCR product by the above method, probably owing to primer mismatches, were analyzed by a different method described later in this section. The resulting products were sequenced directly. In the D-Cu(+) haplotypes of the SNG09-F1 cohort, we found heterogeneous overlapping signals in the sequence chromatogram. This was considered due to difference in the sequences of the two duplicated copies. Then, the two alleles (duplicated copies) of CYP9M10 in this haplotype were amplified separately by the two forward primers specific for each duplicant (targeting upstream from the junction) used in Itokawa et al. (2010) and a reverse primer P32R56 that did not anneal to the OGS haplotype, as described above. PCR was conducted with KOD-FX DNA polymerase using the following amplification cycles for both PCRs: 94 °C for 1 min followed by 30 cycles of 94 °C for 15 s, 58 °C for 15 s, and 68 °C for 3 min and 30 s. The amplified products were sequenced directly. As described above, the gDNA of nine VTN09-F1 individuals showed no product of PCR using the P32R56 primer. Based on the quantitative PCR analysis, the CYP9M10 expression levels

in those samples seemed higher than the mean among the OGS strain individuals (see Results). Thus, we expected that the alleles inherited from VTN09-F1 were abundant compared with the paternal OGS allele in the cDNA of all nine samples given that these alleles may have higher allele-specific expression than the OGS allele. We then used the cDNA, rather than gDNA, of these samples as a template and subjected them to PCR, simply amplifying the CYP9M10 cDNA end-to-end using the specific primers described in the study by Itokawa et al. (2010). As expected, the sequence of the maternal allele could be distinguished from the paternal OGS allele in the chromatogram after direct sequencing because the signal from the OGS allele was visibly weaker or barely visible compared with the maternally inherited allele. There is only one intron in CYP9M10 with a length of 57 bp (Itokawa et al., 2010). As the sequences obtained from cDNA lacked this intron, the subsequent phylogenetic analysis for all the alleles was performed using only open reading frame, that is, the region from start to stop codon excluding the intron.

A reverse primer P32R58 5′-TCCGCCTCGATTGGAACCACA-3′ was designed for PCR to amplify the upstream region of each maternal haplotype, without amplifying from the OGS haplotype. The P32R58 primer targeted the region including the + 404 nucleotide, which is deleted only in the OGS allele (the uniqueness of this deletion was confirmed from the comparison of sequences obtained from the experiments above). The forward primer okaP32F34 5′-TGACATTCTTGTTGGCGTTG-3′ was used with P32R58. The primers amplify a ∼2.2-kb region upstream from the putative transcriptional initiation site (the length is in reference to the JPP haplotype). As okaP32F34 did not work for individuals with the Cu(−) haplotype in F1-VTN10 cohort, another forward primer P32UPSF50 5′-TGAGTACGCAATTTGAGCTGTG AGC-3′ targeting 113-bp upstream from the okaP32F34's annealing site was used alternatively for these samples. PCR was conducted with KOD-FX DNA polymerase using the following amplification cycles: 94 °C for 1 min followed by 30 cycles of 94 °C for 15 s, 58 °C for 15 s, and 68 °C for 5 min. The amplified fragments were directly sequenced.

Allele-specific quantitative PCR

The CYP9M10 allelic ratio ([maternal allele]/[OGS allele]) within the cDNA sample of F1 individuals was measured by allele-specific quantitative PCR (Germer et al., 2000), as described in the study by Itokawa et al. (2010), but a different primer set were used here. By comparing the sequences of all CYP9M10 alleles involved in the F1-cohort, we designed two allele-specific forward primers, P32F60 5′-GCGGAAATCGATCAAGTCAAGGAAC-3′ specific to all alleles other than the OGS allele and P32F60-OGS 5′-GCAGAAA TCGATCACGTCAAAGAGC-3′ (boldfaced letters indicate SNPs) specific only to the OGS allele in the coding region. The primer 5′-CGGTTGGTTAGGCCG AGGGG-3′ was used as a common reverse primer. A fine quantitative standard curve was obtained by mixing two CYP9M10 PCR fragments amplified from JPP and OGS in appropriate ratios as described in the study by Itokawa et al. (2010). The amplification program and reagents for the quantitative PCR followed the study of Itokawa et al. (2010).

Statistical and phylogenetic analysis

The statistical analyses in this study were performed in R (R Development Core Team, 2010). Interval estimates for mortality in the bioassay were performed with the binom.test() function using the method of Clopper and Pearson (1934). Two-sided Fisher's exact tests, Wilcoxon's rank-sum tests and Kruskal–Wallis rank-sum test were performed with the fisher.test(), wilcox.test() and kruskal.test() functions, respectively. Alignment of DNA sequences was performed using the MUSCLE algorithm (Edgar, 2004) in MEGA5 (Tamura et al., 2011) and manually modified in BioEdit (Hall, 1999). Construction of a maximum-parsimony tree and bootstrap tests with 1000 replications were performed in MEGA5. Estimation of the minimum number of recombination events by Hudson and Kaplan's method (1985) was performed with DnaSPv5 (Librado and Rozas, 2009).

RESULTS

Singapore, Kenya and Vietnam C. quinquefasciatus colonies

Three C. quinquefasciatus laboratory-inbred colonies that had been collected recently from Kenya (KNY09), Singapore (SNG09) and

Table 2 Number of dead and survived larvae in cohorts KNY09-F1, SNG09-F1 and VTN10-F1 in the bioassay

F1 cohort	Permethrin (ppm)	Response	CYP9M10 genotype			P-value[a]
			D-Cu(+)/Cu(−)^{OGS}	Cu(+)/Cu(−)^{OGS}	Cu(−)/Cu(−)^{OGS}	
KNY09-F1	0.015	Dead	0	0		—
		Survived	28	11		
		Mortality (95% CI)[b]	0 (0–0.12)	0 (0–0.28)		
	0.04	Dead	0	10		4.5×10^{-7}
		Survived	26	3		
		Mortality (95% CI)	0 (0–0.13)	0.77 (0.46–0.95)		
	0.15	Dead	18	14		0.067
		Survived	6	0		
		Mortality (95% CI)	0.75 (0.53–0.90)	1 (0.77–1)		
SNG09-F1	0.015	Dead	0	0		—
		Survived	10	30		
		Mortality (95% CI)	0 (0–0.31)	0 (0–0.12)		
	0.04	Dead	0	17		2.0×10^{-3}
		Survived	10	13		
		Mortality (95% CI)	0 (0–0.31)	0.57 (0.37–0.75)		
	0.15	Dead	9	27		0.024
		Survived	3	0		
		Mortality (95% CI)	0.75 (0.43–0.95)	1 (0.87–1)		
VTN10-F1	0.015	Dead		0	5	0.30
		Survived		12	23	
		Mortality (95% CI)		0 (0–0.26)	0.18 (0.061–0.37)	
	0.04	Dead		7	28	1.2×10^{-3}
		Survived		5	0	
		Mortality (95% CI)		0.58 (0.28–0.85)	1 (0.88–1)	
	0.15	Dead		11	29	—
		Survived		0	0	
		Mortality (95% CI)		1 (0.72–1)	1 (0.88–1)	

[a]Two-sided Fisher's exact test.
[b]95% confidence interval calculated by the method of Clopper and Pearson (1934).

Vietnam (VTN10) had been reared for <12 generations at the time of the experiment. Genotyping PCR I and II (Itokawa *et al.*, 2010) diagnosing the genotype of *CYP9M10* with respect to form, indicated that all colonies segregated multiple *CYP9M10* haplotype forms, that is, each colony contained at least two of the three forms of the *CYP9M10* haplotype (Supplementary Table S1). Such segregation of the *CYP9M10* genotypes is useful for testing the genotype (of *CYP9M10* forms)–phenotype (pyrethroid susceptibility) association. However, because we had no co-dominant assay for the D-Cu(+) and Cu(+) haplotypes, the two genotypes, D-Cu(+)/Cu(+) and D-Cu(+)/D-Cu(+), were indistinguishable (Supplementary Figure S1). This constraint was problematic particularly in the KNY09 and SNG09 colonies in which those genotypes segregated. Then, we crossed a batch of females from each colony with males from the susceptible laboratory strain OGS. As the OGS strain is homogeneous for the Cu(−) *CYP9M10* haplotype (Itokawa *et al.*, 2010), all possible genotypes from these crossings could be correctly diagnosed using the two PCR genotyping methods (Figure 1). Of note, because there was a nucleotide deletion in the coding region of the *CYP9M10* allele in OGS that resulted in a disruptive frame shift (Itokawa *et al.*, 2010), only the maternal *CYP9M10* haplotype was functional in each progeny.

Association between haplotypic form and pyrethroid susceptibility

The progeny of each crossing are referred to as a 'F1 cohort.' Members of each F1 family were reared as a single batch, and the fourth-instar larvae were selected using three concentrations of permethrin (0.015, 0.04 and 0.15 ppm), which were expected from the results of our previous study (Itokawa *et al.*, 2010) to incur medium lethality. Dead and surviving larvae were each genotyped for *CYP9M10* after the assay. The F1 cohorts KNY09-F1 and SNG09-F1 were generated by crossing KNY09♀ × OGS♂ and SNG09♀ × OGS♂, respectively. The individuals in these cohorts inherited the Cu(+) and D-Cu(+) haplotypes from each maternal colony. The individuals in VTN10-F1, which was derived from a VTN10♀ × OGS♂ crossing, inherited the Cu(−) and Cu(+) haplotype from their maternal (VTN10) colony. In the KNY09-F1 and SNG09-F1 cohorts, individuals that inherited the D-Cu(+) haplotypes from each maternal colony could tolerate a higher concentration of permethrin than those that inherited the Cu(+) haplotype (Table 2). On the other hand, individuals that inherited the Cu(+) haplotype from VTN10 mothers tolerated higher concentration of permethrin than those that inherited the Cu(−) haplotype in the VTN10-F1 cohort (Table 2).
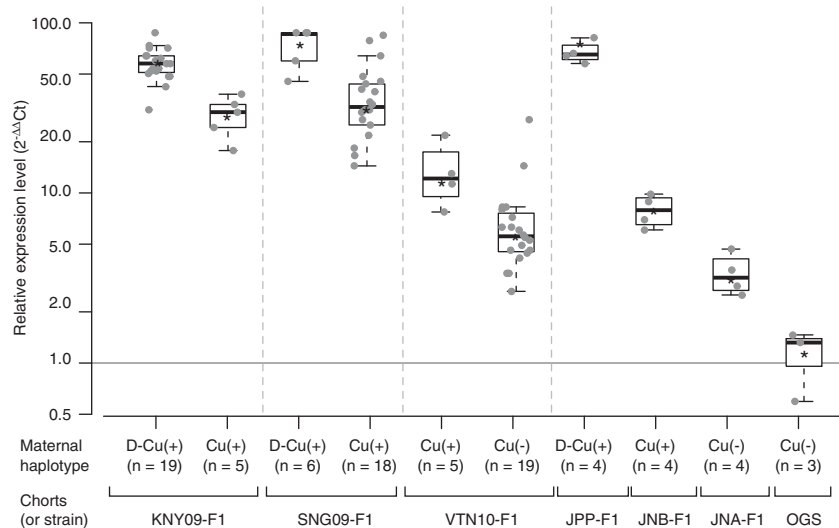
Figure 2 Box plot for relative expression levels ($2^{-\Delta\Delta Ct}$) of *CYP9M10* in the F1 cohorts. Expression levels are shown relative to one individual sample from the OGS strain as an inter-plate calibrator. The data are broken down according to the cohorts and the forms of maternal *CYP9M10* haplotypes. Each gray dot indicates data for individual larva. The vertical axis was shown in logarithmic scale. Asterisks indicate geographic means in each group.

Expression level of *CYP9M10*

Twenty four fourth-instar larvae were collected from each of the three F1 cohorts used in the bioassay, and RNA and DNA were simultaneously extracted from each individual larva. The extracted RNA was used to synthesize cDNA and measure the expression level of *CYP9M10* in each individual, while the gDNA was used to genotype the same individual. We also generated another three F1 cohorts by crossing females from the previously investigated laboratory strains (Table 1) with males from the OGS strain JPP-F1 (JPP♀ × OGS♂), JNA-F1 (JNA♀ × OGS♂) and JNB-F1 (JNB♀ × OGS♂), to make a comparison. Figure 2 shows the relative *CYP9M10* expression levels of individuals normalized by *RPS3* and broken down according to cohort and the form of maternal haplotype inherited. In both KNY09-F1 and SNG09-F1, individuals that inherited D-Cu(+) showed significantly higher expression than those that inherited the Cu(+) haplotype ($P<0.01$ in Wilcoxon rank-sum test for both cohorts; the differences in means within each cohort were 2.0- and 2.2-fold, respectively). In the VTN10-F1 cohort, the mean expression level in individuals that inherited the Cu(+) haplotype was higher than in those that inherited the Cu(−) haplotype, with the difference marginally significant ($P<0.05$ by Wilcoxon rank-sum test and 2.1-fold difference in means).

Allelic and haplotypic variations in *CYP9M10*

The results shown in Figure 2 indicated that the *CYP9M10* mRNA transcription levels in F1-cohort individuals corresponds to the form of the maternal haplotype. Individuals that inherited the D-Cu(+) haplotype in the KNY09-F1 and SNG09-F1cohorts had an expression level similar to that of JPP-F1. Thus, we considered that D-Cu(+) haplotypes in the KNY09 and SNG09 colonies showed the same haplotype-specific expression level as the JPP D-Cu(+) haplotype. However, the situation for the Cu(+) and Cu(−) haplotypes was not as clear. The KNY09-F1 and SNG09-F1 individuals that inherited the Cu(+) haplotype as maternal haplotype apparently had higher expression levels than those that inherited the Cu(+) haplotype in the VTN10-F1 and JNB-F1 cohorts (Figure 2). Furthermore, the expression levels among the SNG09-F1 individuals that inherited the Cu(+) haplotype as maternal haplotype varied greatly (Figure 2).

Table 3 Alleles associated with each form of haplotype in each strain and colony

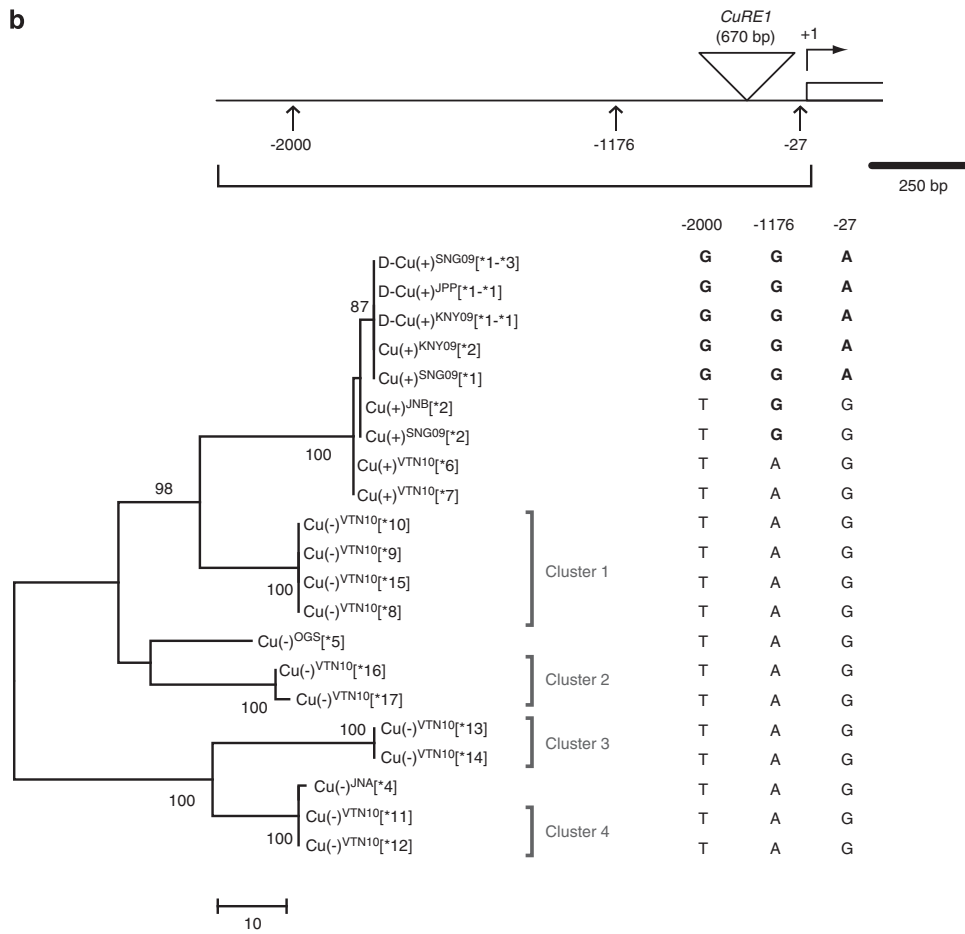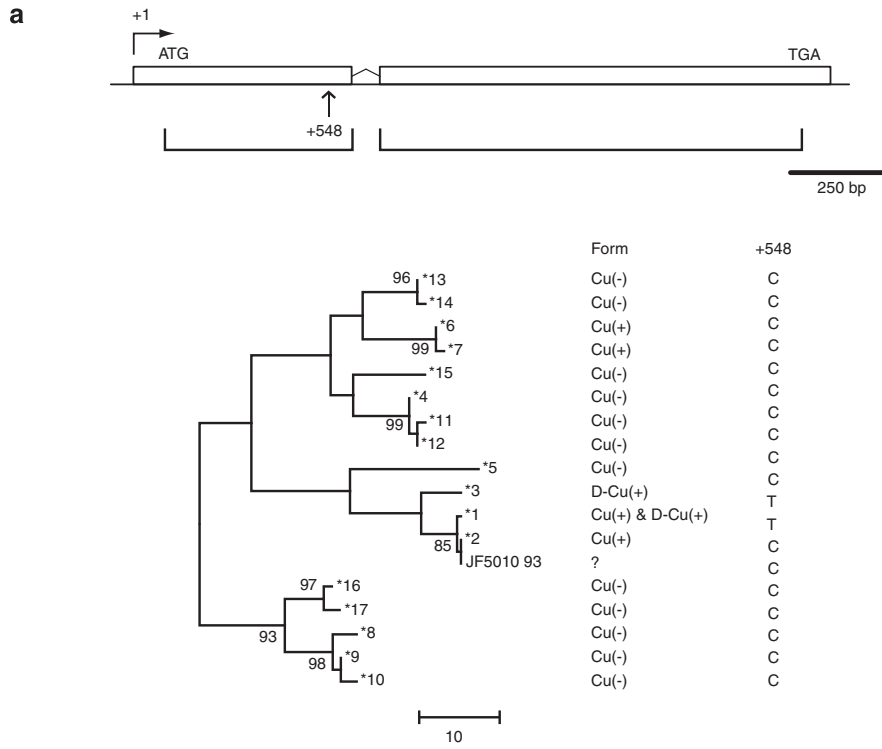| Strain	Form	Allele	Colony	Form	Allele
JPP	D-Cu(+)	[*1–*1][a]	KNY09	D-Cu(+)	[*1–*1]
JNB	Cu(+)	[*2][b]		Cu(+)	[*2]
JNA	Cu(−)	[*4][c]	SNG09	D-Cu(+)	[*1–*3]
OGS	Cu(−)	[*5][d]		Cu(+)	[*1] & [*2]
			VTN10	Cu(+)	[*6] & [*7]
				Cu(−)	[*8]–[*17]

[a]Itokawa et al., 2010 (GenBank ID: AB551111 and AB551112).
[b]Itokawa et al., 2011 (AB607193).
[c]Itokawa et al., 2011 (AB607192).
[d]Itokawa et al., 2010 (AB551113).

Similarly, there was also a relatively large variance in expression among VTN10-F1 individuals that inherited the Cu(−) haplotype as maternal haplotype (Figure 2). These results suggested the presence of 'within-form' variation in expression level.

We accordingly amplified and sequenced the regions including the whole coding region of the maternal *CYP9M10* alleles for all 72 individuals used in the expression analysis from their gDNA or cDNA samples to resolve the variation of the alleles at a higher resolution. By combining this data with previously reported sequences of the four laboratory strains, 17 different *CYP9M10* alleles (*1–*17) were defined in total (Table 3). The alleles in the previously reported haplotypes in the JPP (duplicated), JNB, JNA and OGS strains (Itokawa et al., 2010, 2011; Komagata et al., 2010) are assigned as *1, *2, *4 and *5, respectively (Table 3). Only one allele (*2) was associated with the Cu(+) haplotypes in the KNY09-F1 cohort. However, two distinct alleles (*1 and *2) were found to be associated with Cu(+) haplotypes in the SNG09-F1 cohort (Table 3). Among the D-Cu(+) haplotypes isolated from the KNY09-F1 individuals, all shared the same allele (*1) as the two duplicated copies. In contrast, each duplicated loci in the D-Cu(+) haplotype in the SNG09-F1 individuals harbored two distinct alleles (*1 and *3) where the allele *1 resided in the upstream locus whereas allele *3 was in the downstream locus (Table 3). Twelve distinct alleles (*6–*17) were

a

Form	+548
Cu(-)	C
Cu(-)	C
Cu(+)	C
Cu(+)	C
Cu(-)	C
Cu(-)	C
Cu(-)	C
Cu(-)	C
Cu(-)	C
D-Cu(+)	T
Cu(+) & D-Cu(+)	T
Cu(+)	C
?	C
Cu(-)	C
Cu(-)	C
Cu(-)	C
Cu(-)	C
Cu(-)	C

b

	-2000	-1176	-27
D-Cu(+)[SNG09][*1-*3]	**G**	**G**	**A**
D-Cu(+)[JPP][*1-*1]	**G**	**G**	**A**
D-Cu(+)[KNY09][*1-*1]	**G**	**G**	**A**
Cu(+)[KNY09][*2]	**G**	**G**	**A**
Cu(+)[SNG09][*1]	**G**	**G**	**A**
Cu(+)[JNB][*2]	T	**G**	G
Cu(+)[SNG09][*2]	T	**G**	G
Cu(+)[VTN10][*6]	T	A	G
Cu(+)[VTN10][*7]	T	A	G
Cu(-)[VTN10][*10]	T	A	G
Cu(-)[VTN10][*9]	T	A	G
Cu(-)[VTN10][*15]	T	A	G
Cu(-)[VTN10][*8]	T	A	G
Cu(-)[OGS][*5]	T	A	G
Cu(-)[VTN10][*16]	T	A	G
Cu(-)[VTN10][*17]	T	A	G
Cu(-)[VTN10][*13]	T	A	G
Cu(-)[VTN10][*14]	T	A	G
Cu(-)[JNA][*4]	T	A	G
Cu(-)[VTN10][*11]	T	A	G
Cu(-)[VTN10][*12]	T	A	G

found from the 24 individuals in VTN10-F1. Two alleles (*6 and *7) were associated with the Cu(+) haplotypes whereas the other 10 alleles (*8–*17) were found in the Cu(−) haplotypes in this cohort.

The sequences of the coding regions (from the start codon to the stop codon excluding the intron) in all alleles found in this study were aligned and phylogenetically analyzed (Figure 3a). This analysis also included the cDNA sequence from the HAmCq^{G08} strain (JF501093 in GenBank) recently deposited by Liu et al. (2011). The minimum number of recombination events, R_m, was 12 as estimated by Hudson and Kaplan's method (1985), and there were also other specific evidences of past recombination events that will be detailed below. The phylogenetic tree shown in Figure 3a (and also that in Figure 3b), therefore, is used mainly as a visual representation of sequence similarity among the alleles, and the tree topology does not necessarily represent a strict genealogical relationship among the alleles. Alleles *1 and *2, which were both associated with the CuRE1 insertion, differed by only one nucleotide substitution at the +548 position (numbering is with respect to the *1 allele in the JPP strain, considering the putative transcriptional initiation site as +1), that is, thymine in allele *1 but cytosine in allele *2. In our previous paper, we postulated that allele *1 was directly derived from allele *2 by a C548T nucleotide substitution that occurred just before the duplication (Itokawa et al., 2011). Indeed, all alleles, except alleles *1 and *3, carried a cytosine at this nucleotide locus, indicating C548 was ancestral (Figure 3a and Supplementary Table S2). Allele *3, found in the downstream copy of the D-Cu(+) haplotype in the SNG09-F1 cohort also carried a thymine at this locus, but its sequence differed from that of allele *1 by nine nucleotide substitutions. All nine SNPs were located only within a 10% fraction of the 3′ end, and only three of nine were unique in this allele (Supplementary Table S2), indicating that this allele had resulted from a past recombination in allele *1 after duplication, rather than accumulation of de novo substitutions after the duplication or an independent duplication event. Of note, the cDNA sequence of CYP9M10 in the HAmCq^{G08} strain (JF501093) was completely identical to that in allele *2 (Figure 3a). The twelve alleles in the VTN10-F1 cohort (*6–*17) exhibited considerable sequence diversity (Figure 3a). The Cu(+) haplotypes in this cohort were associated with two alleles (*6 and *7) differing from each other by only one nucleotide substitution. The sequences of these alleles were apparently distant from those of the group consisting of alleles associated with the CuRE1 insertion (alleles *1 and *2) (Figure 3a).

We extended our analysis to the upstream region of CYP9M10 including the site of the CuRE1 insertion. In our previous study, we found two polymorphic sites in the upstream region (at −27 and −2000: positions are with reference to the upstream copy of the JPP haplotype relative to the putative transcriptional initiation site +1) between the Cu(+)^{JNB}[*2] and D-Cu(+)^{JPP}[*1–*1] haplotypes. The region corresponding to the 2.2-kb upstream region in the upstream copy of JPP CYP9M10 haplotype of the four Cu(+) and two D-Cu(+) haplotypes were sequenced to resolve the diversity among these haplotypes belonging to those forms more deeply. The

10 distinct maternal Cu(−) haplotypes in VTN10-F1 (Cu(−)^{VTN10}[*8]–[*17]) were also analyzed as well to compare the diversity of the upstream sequence between Cu(−) and CuRE1-inserted haplotypes. The Cu(+) and D-Cu(+) haplotypes showed high nucleotide identity (there were only three polymorphic sites among all haplotypes) compared with that among the Cu(−) haplotypes (Figure 3b). The upstream sequences of the Cu(+)^{VTN10}[*6] and Cu(+)^{VTN10}[*7] haplotypes also belonged to this monophyletic clade including other Cu(+) and D-Cu(+) haplotypes (Figure 3b). Thus, the discrepancy of alleles (CDSs) associated with these haplotypes from other Cu(+) and D-Cu(+) seen in Figure 3a were considered due to past recombination(s). The three polymorphic sites among Cu(+) and D-Cu(+) within this region were at −27 (T/G), −1176 (A/G) and −2000 (G/A). Segregation at these sites was seen only among Cu(+) haplotypes (Figure 3b), whereas the nucleotides at the same sites were homogeneous in all Cu(−) haplotypes (T −2000; A −1176; G −27) and were considered ancient. The D-Cu(+) haplotypes were also homogeneous in the three sites, but the nucleotides at all sites were different from those in the Cu(−) haplotypes (G −2000; G −1176; A −27) that were considered as derived. Cu(+)^{SNG09}[*1] and Cu(+)^{KNY09}[*2] carried nucleotides at the three sites considered as derived, that is, these two Cu(+) haplotypes shared a sequence fully identical to that of to the D-Cu(+) haplotypes in upstream region (Figure 3b).

Allele-specific expression levels

The relative abundance of a maternal CYP9M10 allele to the paternal OGS allele (*5) within each cDNA sample of the F1 cohorts was measured. Such allelic imbalance within a transcriptome of a heterozygous individual represents a difference in efficacy of cis-acting regulatory elements between the two alleles (Cowles et al., 2002). Thus, to compare the efficacies of cis-acting regulatory elements linked with each allele, the allele-specific expression levels of the maternal allele normalized by allele *5 (referred to as relative allele-specific expression level or rASE) was considered to provide relatively stable values among different genetic backgrounds compared with the $2^{-\Delta\Delta Ct}$ values. The rASEs of the individuals in each F1 cohort are plotted in Figure 4. The data for the two Cu(+) haplotypes in the VTN10-F1 cohort were pooled as a single haplotype category (Cu(+)^{VTN10}[*6] & [*7]) because the Cu(+)^{VTN10}[*6] haplotype was found in only one individual. The D-Cu(+) haplotypes showed the highest mean rASEs and the Cu(+) haplotypes showed a middle range of mean rASEs. In fact, there appeared to be two groups ('lower' and 'higher') among Cu(+) haplotypes differing in mean rASE; the higher group consisted of Cu(+)^{KNY09}[*2] and Cu(+)^{SNG09}[*1], and the lower group consisted of Cu(+)^{SNG09}[*2], Cu(+)^{VTN10}[*6] & [*7], and Cu(+)^{JNB}[*2] (Figure 4). While the rASEs of Cu(−) haplotypes were located at the lower extremity of the whole distribution of rASEs, there was a relatively large variance among the Cu(−) haplotypes inherited from VTN10 colony. Moreover, all Cu(−) haplotypes showed rASE far higher than 1, which

Figure 3 Maximum parsimony (MP)-trees using CDSs of (a) the 18 CYP9M10 alleles and (b) the upstream regions of 21 haplotypes. The lengths of the branches indicate the number of changes over the whole sequence. A total of (a) 1613 and (b) 1436 nucleotide sites (after all positions containing gaps or missing data were eliminated) were used. The trees are bootstrap consensus trees inferred from 1000 replicates. Bootstrap values of >80% are shown. (a) A MP-tree for the CDSs (alleles). The regions used for the analysis is shown in the top. A total of 1613 positions were used. The haplotype forms with which each allele was associated and the bases at the +548 sites in each allele are indicated in the right box. (b) A MP-tree for upstream sequences. The regions used for the analysis is shown in the top. Nucleotides at the three polymorphic sites (−2000, −1176 and −27) among the Cu(+) and D-Cu(+) haplotypes are shown in right. Haplotypes in the VTN10 colony were clustered into four clusters (clades 1–4) according to the sequence. Note that the trees do not indicate the strict genealogical relationship as there are past recombinations, that is, the trees are mainly used for visual representation for sequence similarities. Although there is no out-group sequence, the tree is drawn like rooted tree for visual conciseness.
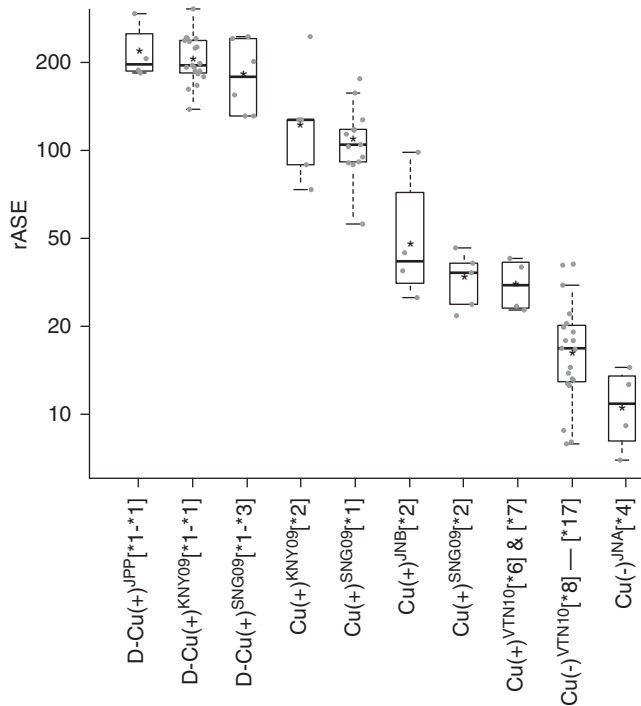
Figure 4 Box plot for allele-specific expressions of each maternal haplotype in the F1-cohorts relative to the paternal OGS allele (rASE). Each gray dot indicates individual larva. The vertical axis is shown in logarithmic scale. Asterisks indicate the mean values of each haplotype.

indicates that the rASE in Cu(−)^{OGS}[*5] does not represents a typical basal ASE, but rather a 'particularly low' ASE (Figure 4). The mean rASE of Cu(−)^{JNA}[*4] was 10 (the 95% confidence interval was 6.1–18 by one-sample t-test using log-transformed values of rASEs). This result was inconsistent with our previous report that the mean expression (not in rASE) level of JNA larvae was only 1.5-fold higher than that of OGS larvae (Itokawa et al., 2011). The reason of this discrepancy is not yet clear so far, but, we speculate that some influence from genetic background exists. In any case, a notice should be required to refer to the values of CYP9M10 relative expression levels calculated in past studies using the OGS strain as a reference strain (Itokawa et al., 2010, 2011; Komagata et al., 2010). In the VTN10-F1 cohort, rASE varied greatly also among the Cu(−) haplotypes inherited from the VTN10 colony (Figure 4). We tentatively clustered the 10 haplotypes associated with this form into four clades (Clade 1–4), using the upstream sequences (Figure 3b). Based on this classification, the independence of the rASEs from the clades was rejected by the Kruskal–Wallis rank-sum test (P<0.01) (Supplementary Figure S2). This result means that there are variations in cis-acting regulatory elements even among these Cu(−) haplotypes.

DISCUSSION

In this study, the two forms of CYPM10 haplotype associated with pyrethroid resistance, D-Cu(+) and Cu(+), were found to be distributed in Asian and African populations of C. quinquefasciatus (Table 2). In addition, the reported cDNA sequence of the HAmCq^{G08} strain (Liu et al., 2011), which was established by selection from a colony collected in Alabama, USA, in 2002, appeared to have an

identical sequence to the *2 allele that was linked to the Cu(+) form among our samples (Table 3). Given that CYP9M10 was over-expressed in the HAmCq^{G08} strain (Liu et al., 2011), it is highly possible that overexpressing CYP9M10 haplotypes derived from the same lineage of Cu(+) and D-Cu(+) also spread in America. Raymond et al. (1991) previously reported that an amplified carboxyl esterase haplotype (conferring resistance against organophosphate insecticides) of single origin have spread into worldwide populations of the C. pipiens species complex including C. quinquefasciatus. The report by Raymond et al. (1991) and our present study indicate that there is little genetic barrier for resistance genes among the sub-populations of this mosquito species around the world, probably because of passive migration though human activity such as transportation.

The D-Cu(+) haplotypes in the Kenya and Singapore colonies were associated with higher resistance levels than the Cu(+) haplotypes within the same colony. This difference of the phenotypic effects was correlated with transcriptional level of CYP9M10, such that D-Cu(+) haplotypes consistently associated with higher transcription levels than did Cu(+) haplotypes (Figure 2). Such allelic (haplotypic) variation in resistance and expression level was previously suggested in the comparison of the laboratory strains JPP and JNB (Itokawa et al., 2011), but first confirmed in this study for a field population of various geographic origins. The segregation of these CYP9M10 regulatory variants in the field population is considered to consist of some component of the phenotypic variance in pyrethroid susceptibility of the wild C. quinquefasciatus population.

A strong recent positive selection on a beneficial mutation, such as one conferring insecticide resistance, causes a rapid expansion of single-haplotype lineages resulting in reduction in nucleotide diversity around the mutation (Catania et al., 2004; Schlenke and Begun, 2004). Although several past recombinations have broken the stringent congruence between 'form' and 'allele', all D-Cu(+) and Cu(+) haplotypes sequences compared in this study were highly similar in either or both coding and upstream regions regardless in which geographic area they had been isolated. These Cu(+) and D-Cu(+) haplotypes thus share a recent common ancestral haplo-type, despite their global distribution suggesting the existence of positive selection on this haplotype lineage. Given that the JPP strain originated in 1981 (Amin and Hemingway, 1989), this haplotype is expected to have already existed >30 years ago. Several studies (Guillemaud et al., 1998; Field and Foster, 2002; Hartley et al., 2006; Schmidt et al., 2010) have observed replacements of previously dominating resistant alleles by new, fitter alleles over time. Despite the fitness advantage of D-Cu(+) over Cu(+) haplotypes under insecticide pressure, however, it is unclear whether the D-Cu(+) haplotype will sweep the Cu(+) haplotype out of the current gene pool of C. quinquefasciatus. Pyrethroid insecticides are usually used to combat adult mosquitoes in the field. However, the resistance associated with the D-Cu(+) and Cu(+) haplotypes decreases markedly in the adult stage, because the overexpression of CYP9M10 almost halts after the larval stage (Hardstone et al., 2007; Komagata et al., 2010; Li and Liu, 2010; Itokawa et al., 2011). Although non-targeted organisms, such as Drosophila (Catania et al., 2004; Schlenke and Begun, 2004), could be subject to selective pressure of insecti-cides, perhaps as off-target victims of agricultural pesticide usage (Antonio-Nkondjio et al., 2011), it is unclear how often in the field C. quinquefasciatus encounters aquatic environments with high pyre-throid concentrations, where the D-Cu(+) haplotype is advanta-geous over Cu(+) haplotypes. Furthermore, Hardstone et al. (2009) reported that the D-Cu(+) haplotype of the JPP strain was associated

with a fitness cost in the absence of the insecticide compared with the susceptible allele, presumably a burden incurred by the overproduction of CYP9M10 proteins. Although whether Cu(+) haplotypes can enjoy higher fitness relative to D-Cu(+) in the absence of insecticide pressure remains to be confirmed; if so, this advantage may further mitigate the spread of the D-Cu(+) haplotype. Therefore, further population genetic analysis will be required to investigate the current frequency of the D-Cu(+) and Cu(+) haplotypes, and the actual selective advantages of these haplotypes in the field.

As D-Cu(+) haplotypes were clearly derived from a Cu(+) haplotype and the two haplotypes showed the same decreasing pattern of expression through the larval to adult stages (Itokawa et al., 2011), the two haplotypes may, at least partly, share the same mechanism for overexpression. In the present study, Cu(+) haplotypes of different geographic origins consistently showed high expression levels comparable to that of the JNB haplotype. However, we also found regulatory variations within each Cu(−) and Cu(+) haplotype, and the distributions of the expression levels partly overlapped. It is unclear whether Cu(−) haplotypes with relatively high expression confer an effect on the resistance phenotype equivalent to that of Cu(+) haplotypes with relatively low expression levels. Given that we measured CYP9M10 expression in the whole insect body while ignoring other important aspects such as the spatial localization of this enzyme, equal CYP9M10 expression levels are not necessarily associated with equivalent resistance level if the cis-acting regulatory mutations have different origins. This issue awaits elucidation in a future study. At least two different groups (the high and low) of Cu(+) haplotypes were associated with different expression levels (Figure 4). Interestingly, the Cu(+) haplotypes in the high group were genealogically closer to the D-Cu(+) haplotypes than those in the low group (Figure 3b and Figure 4), and the mean expression levels of the Cu(+) haplotypes in the high group and the D-Cu(+) haplotypes differed only by approximately two-fold (as did the difference in copy number) (Figure 4). Thus, we propose that D-Cu(+) haplotype arose from a Cu(+) haplotype with relatively high expression level by duplication and that this duplication has simply doubled the expression level. This suggests that there were at least two steps of regulatory cis-acting mutations that occurred successively before the duplication. Pinpointing the location of the regulatory mutation(s) responsible for the overexpression of CYP9M10 is necessary to decipher the exact history of molecular evolution in this haplotype lineage. This achievement will extend our knowledge about the origins and evolutionary patterns of insecticide resistance genes in insects.

DATA ARCHIVING

The sequences of the haplotypes or cDNA of CYP9M10 newly analyzed in this study were deposited to DNA Data Bank of Japan (DDBJ); the accession numbers are AB724260 for D-Cu(+)^{KNY09}[*1–*1], AB724261 for Cu(+)^{KNY09}[*2], AB724262 for D-Cu(+)^{SNG09}[*1–*3], AB724263 for Cu(+)^{SNG09}[*1], AB724264 for D-Cu(+)^{SNG09}[*1–*3], AB724265 for Cu(+)^{SNG09}[*2], AB724266 for Cu(+)^{VTN10}[*6], AB724267 for Cu(+)^{VTN10}[*7], AB724268 for Cu(−)^{VTN10}[*8], AB724269 for Cu(−)^{VTN10}[*9], AB724270 for Cu(−)^{VTN10}[*10], AB724271 for Cu(−)^{VTN10}[*11], AB724272 for Cu(−)^{VTN10}[*12], AB724273 for Cu(−)^{VTN10}[*13], AB724274 for Cu(−)^{VTN10}[*14], AB724275 for Cu(−)^{VTN10}[*15], AB724276 for Cu(−)^{VTN10}[*16], AB724277 for Cu(−)^{VTN10}[*17], AB724278 for cDNA of allele *13, AB724279 for cDNA allele *14 and AB724280 for cDNA of allele *15.

Amin AM, Hemingway J (1989). Preliminary investigation of the mechanisms of DDT and pyrethroid resistance in Culex quinquefasciatus Say (Diptera: Culicidae) from Saudi Arabia. Bull Entomol Res 79: 361–366.

Antonio-Nkondjio C, Fossog BT, Ndo C, Djantio BM, Togouet SZ, Awono-Ambene P et al. (2011). Anopheles gambiae distribution and insecticide resistance in the cities of Douala and Yaounde (Cameroon): influence of urban agriculture and pollution. Malar J 10: 154.

Arensburger P, Megy K, Waterhouse RM, Abrudan J, Amedeo P, Antelo B et al. (2010). Sequencing of Culex quinquefasciatus establishes a platform for mosquito comparative genomics. Science 330: 86–88.

Berticat C, Marquine M, Raymond M, Chevillon C (2001). Recombination between two amplified esterase alleles in Culex pipiens. J Hered 92: 349–351.

Catania F, Kauer MO, Daborn PJ, Yen JL, Ffrench-Constant RH, Schlotterer C (2004). World-wide survey of an Accord insertion and its association with DDT resistance in Drosophila melanogaster. Mol Ecol 13: 2491–2504.

Chung H, Bogwitz MR, McCart C, Andrianopoulos A, Ffrench-Constant RH, Batterham P et al. (2007). Cis-regulatory elements in the Accord retrotransposon result in tissue-specific expression of the Drosophila melanogaster insecticide resistance gene Cyp6g1. Genetics 175: 1071–1077.

Clopper CJ, Pearson ES (1934). The use of confidence or fiducial limits illustrated in the case of the binomial. Biometrika 26: 404–413.

Cowles CR, Hirschhorn JN, Altshuler D, Lander ES (2002). Detection of regulatory variation in mouse genes. Nat Genet 32: 432–437.

Daborn PJ, Yen JL, Bogwitz MR, Le Goff G, Feil E, Jeffers S et al. (2002). A single p450 allele associated with insecticide resistance in Drosophila. Science 297: 2253–2256.

Edgar RC (2004). MUSCLE: multiple sequence alignment with high accuracy and high throughput. Nucleic Acids Res 32: 1792–1797.

Feyereisen R (2012). Insect CYP Genes and P450 Enzymes. In: Lawrence IG (ed) Insect Molecular Biology and Biochemistry. Academic Press: San Diego, USA, pp 236–316.

Field LM, Foster SP (2002). Amplified esterase genes and their relationship with other insecticide resistance mechanisms in English field populations of the aphid, Myzus persicae (Sulzer). Pest Manag Sci 58: 889–894.

Germer S, Holland MJ, Higuchi R (2000). High-throughput SNP allele-frequency determination in pooled DNA samples by kinetic PCR. Genome Res 10: 258–266.

Guillemaud T, Lenormand T, Bourguet D, Chevillon C, Pasteur N, Raymond M (1998). Evolution of resistance in Culex pipiens: allele replacement and changing environment. Evolution 52: 443–453.

Hall TA (1999). BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. Nucl Acids Symp Ser 41: 95–98.

Hardstone MC, Komagata O, Kasai S, Tomita T, Scott JG (2010). Use of isogenic strains indicates CYP9M10 is linked to permethrin resistance in Culex pipiens quinquefasciatus. Insect Mol Biol 19: 717–726.

Hardstone MC, Lazzaro BP, Scott JG (2009). The effect of three environmental conditions on the fitness of cytochrome P450 monooxygenase-mediated permethrin resistance in Culex pipiens quinquefasciatus. BMC Evol Biol 9: 42.

Hardstone MC, Leichter C, Harrington LC, Kasai S, Tomita T, Scott JG (2007). Cytochrome P450 monooxygenase-mediated permethrin resistance confers limited and larval specific cross-resistance in the southern house mosquito, Culex pipiens quinquefasciatus. Pesticide Biochem Physiol 89: 175–184.

Hartley CJ, Newcomb RD, Russell RJ, Yong CG, Stevens JR, Yeates DK et al. (2006). Amplification of DNA from preserved specimens shows blowflies were preadapted for the rapid evolution of insecticide resistance. Proc Natl Acad Sci USA 103: 8757–8762.

Hudson RR, Kaplan NL (1985). Statistical properties of the number of recombination events in the history of a sample of DNA sequences. Genetics 111: 147–164.

Itokawa K, Komagata O, Kasai S, Masada M, Tomita T (2011). Cis-acting mutation and duplication: History of molecular evolution in a P450 haplotype responsible for insecticide resistance in Culex quinquefasciatus. Insect Biochem Mol Biol 41: l503–512.

226

Itokawa K, Komagata O, Kasai S, Okamura Y, Masada M, Tomita T (2010). Genomic structures of *Cyp9m10* in pyrethroid resistant and susceptible strains of *Culex quinquefasciatus*. *Insect Biochem Mol Biol* **40**: 631–640.

Kasai S, Weerashinghe IS, Shono T (1998). P450 monooxygenases are an important mechanism of permethrin resistance in *Culex quinquefasciatus* Say larvae. *Arch Insect Biochem Physiol* **37**: 47–56.

Kojima KK, Jurka J (2011). DNA transposons from the southern house mosquito. *Repbase Reports* **11**: 55–55.

Komagata O, Kasai S, Tomita T (2010). Overexpression of cytochrome P450 genes in pyrethroid-resistant *Culex quinquefasciatus*. *Insect Biochem Mol Biol* **40**: 146–152.

Labbe P, Sidos N, Raymond M, Lenormand T (2009). Resistance gene replacement in the mosquito *Culex pipiens*: fitness estimation from long-term cline series. *Genetics* **182**: 303–312.

Li T, Liu N (2010). Inheritance of permethrin resistance in *Culex quinquefasciatus*. *J Med Entomol* **47**: 1127–1134.

Librado P, Rozas J (2009). DnaSP v5: a software for comprehensive analysis of DNA polymorphism data. *Bioinformatics* **25**: 1451–1452.

Liu H, Cupp EW, Micher KM, Guo A, Liu N (2004). Insecticide resistance and cross-resistance in Alabama and Florida strains of *Culex quinquefasciatus*. *J Med Entomol* **41**: 408–413.

Liu N, Li T, Reid WR, Yang T, Zhang L (2011). Multiple cytochrome P450 genes: their constitutive overexpression and permethrin induction in insecticide resistant mosquitoes, *Culex quinquefasciatus*. *PLoS One* **6**: e23403.

Livak KJ, Schmittgen TD (2001). Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T)) Method. *Methods* **25**: 402–408.

Portaro JK, Barr AR (1975). 'Curing' Wolbachia infections in *Culex pipiens*. *J Med Entomol* **12**: 265.

R Development Core Team (2010). R: a language and environment for statistical computing. *R Foundation for Statistical Computing,* Vienna, Austria.

Raymond M, Callaghan A, Fort P, Pasteur N (1991). Worldwide migration of amplified insecticide resistance genes in mosquitoes. *Nature* **350**: 151–153.

Schlenke TA, Begun DJ (2004). Strong selective sweep associated with a transposon insertion in Drosophila simulans. *Proc Natl Acad Sci USA* **101**: 1626–1631.

Schmidt JM, Good RT, Appleton B, Sherrard J, Raymant GC, Bogwitz MR *et al.* (2010). Copy number variation and transposable elements feature in recent, ongoing adaptation at the *Cyp6g1* locus. *PLoS Genet* **6**: e1000998.

Smith JL, Fonseca DM (2004). Rapid assays for identification of members of the *Culex (Culex) pipiens* complex, their hybrids, and other sibling species (Diptera: culicidae). *Am J Trop Med Hyg* **70**: 339–345.

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.

Wilding CS, Smith I, Lynd A, Yawson AE, Weetman D, Paine MJ *et al.* (2012). A *cis*-regulatory sequence driving metabolic insecticide resistance in mosquitoes: functional characterisation and signatures of selection. *Insect Biochem Mol Biol* **42**: 699–707.

Yebakima A, Marquine M, Rosine J, Yp-Tcha MM, Pasteur N (2004). Evolution of resistance under insecticide selection pressure in *Culex pipiens quinquefasciatus* (Diptera, Culicidae) from Martinique. *J Med Entomol* **41**: 718–725.

Supplementary Information accompanies this paper on Heredity website (http://www.nature.com/hdy)