Tracking genetic invasions: genome-wide SNPs reveal the source of pyrethroid-resistant *Aedes aegypti* (yellow fever mosquito) incursions at international ports

Tom L. Schmidt[1], Anthony R. van Rooyen[2], Jessica Chung[1,3], Nancy M. Endersby-Harshman[1], Philippa C. Griffin[1], Angus Sly[4], Ary A. Hoffmann[1]*^, Andrew R. Weeks[1,2]*^

[1]Bio21 Institute, School of BioSciences, The University of Melbourne, Victoria 3010, Australia

[2]cesar Pty Ltd, 293 Royal Parade, Parkville, Victoria 3052, Australia

[3]Melbourne Bioinformatics, The University of Melbourne, Victoria 3010, Australia

[4]Department of Agriculture and Water Resources, Qantas Drive, Brisbane Airport, Queensland 4008, Australia

*Equal senior authors

^Corresponding authors

Abstract

Biological invasions are increasing globally in number and extent despite concerted efforts to restrict their spread. Knowledge of incursion pathways is necessary to prevent new invasions and to design more effective biosecurity protocols at source and recipient locations. In this study, we use genome-wide single-nucleotide polymorphisms (SNPs) and reference mosquitoes from 15 locations to determine the origin of 115 incursive *Aedes aegypti* (yellow fever mosquito) collected at international ports in Australia and New Zealand. We also genotyped mosquitoes at three point mutations in the voltage-sensitive sodium channel (*Vssc*) gene. These mutations are known to confer knockdown resistance to synthetic pyrethroid insecticides, which are widely used for controlling invertebrate pests. We first delineated reference populations using a set of 188 unrelated *Ae. aegypti* collected from the 15 locations across Southeast and East Asia, South America, Australia and the Pacific Islands. Incursives were assigned to these populations using discriminant analysis of principal components (DAPC) and an assignment test with a support vector machine predictive model, which showed strong agreement in assignment. We found that Bali, Indonesia, was the most common origin of *Ae. aegypti* intercepted at international ports in Australia, while *Ae. aegypti* detected at international ports in New Zealand originated from Pacific Islands such as Fiji. Genotyping at the *Vssc* gene showed that most incursives detected at Australian international ports had the same allelic genotype across the three point mutations, a genotype which is known to confer strong resistance to synthetic pyrethroids, the only insecticide class used in current aircraft disinsection protocols. This was the only genotype found in mosquitoes from Bali and was also present in other reference mosquitoes across Southeast Asia and the Pacific Islands. This research highlights the inadequacy of current disinsection protocols endorsed by the World Health Organisation (WHO) for preventing invasions of *Ae. aegypti*, while demonstrating the utility of genomic datasets for managing global biosecurity objectives.

Introduction

Invasive alien species are a major threat to global biodiversity, agriculture, and human health, and confer a great economic burden upon developed and developing countries alike [1]. Invasive species have been shown to disrupt soil processes [2] and ecological community structure [3], and are linked to reductions in the abundance and richness of native species [4]. They are responsible for the global spread of human diseases such as malaria, dengue, Zika, and chikungunya that together cause over 700,000 deaths annually and create heavy burdens on health systems worldwide [5]. The spread of invasive species is closely tied to human economic activity, with many species transported as 'stowaways' on human trade and transport vessels [6, 7]. Following the expansion of trade and transport activity in recent decades, global distributions of invasive species have increased markedly [8], with current predictions still arguably underestimating their true extent [9].

Preventing the establishment of new alien species requires knowledge of incursion pathways [1]. Although border interceptions can be effective at stopping incursions [10, 11], resource limitations restrict the number of inspections on arriving vessels or the number of border sections that can be monitored. These strategies must therefore seek to maximise detections by incorporating knowledge of incursion pathways, such as more frequent inspection of vessels from likely source locations [12], or regular inspection of known entry points of highly destructive species [13, 14]. Incursion pathways can also be monitored to estimate parameters indicating the likelihood of future invasions, such as propagule pressure and opportunities for selection [15].

Genomic information can be a useful means for identifying source populations of incursions. Invading individuals of unknown source intercepted at borders (hereafter defined as "incursives") can be genotyped using genetic markers and compared with reference populations from the species' known distribution. Analyses can then be performed to assign incursives to reference populations, in order to identify the likely source population or geographic region of origin. Genome-wide single nucleotide polymorphisms (SNPs) in particular can provide clear geographical delineation of populations [16], which allows for confident assignment of intercepted individuals to their population or region of origin. This method can in theory enable the assignment of all incursive individuals, and thus improves upon previous methods that either rely on the interception of incursives on vessels as they arrive from a known location [14], or that use lower-resolution genetic markers that frequently cannot resolve individual origin to a fine geographic scale [17]. Genome-wide SNPs could also be used to identify the source of incursions post-border and improve the chance of eradication by using pathway management to prevent future incursions.

Aedes aegypti (yellow fever mosquito) is a highly invasive pest and a primary global vector of arboviruses such as dengue [18], Zika [19] and chikungunya [20]. Its pantropical distribution has grown rapidly over the last century [21], placing a greater part of the human population at risk of disease. In Australia, *Ae. aegypti* is currently only endemic to Far North Queensland, but once had a broader distribution that extended into New South Wales, South Australia and Western Australia [22]. Under current climate change scenarios, there is speculation that the species will increase its distribution in Australia again [22-25]. Furthermore, detections of *Ae. aegypti* at Australia's international airports increased markedly over the period from 2014 to 2016 (Figure 1), with 39 detections at Perth International Airport alone between October 2015 and April 2016. New Zealand

does not currently support a local *Ae. aegypti* population, but historically has received incursives from places such as Fiji [26].

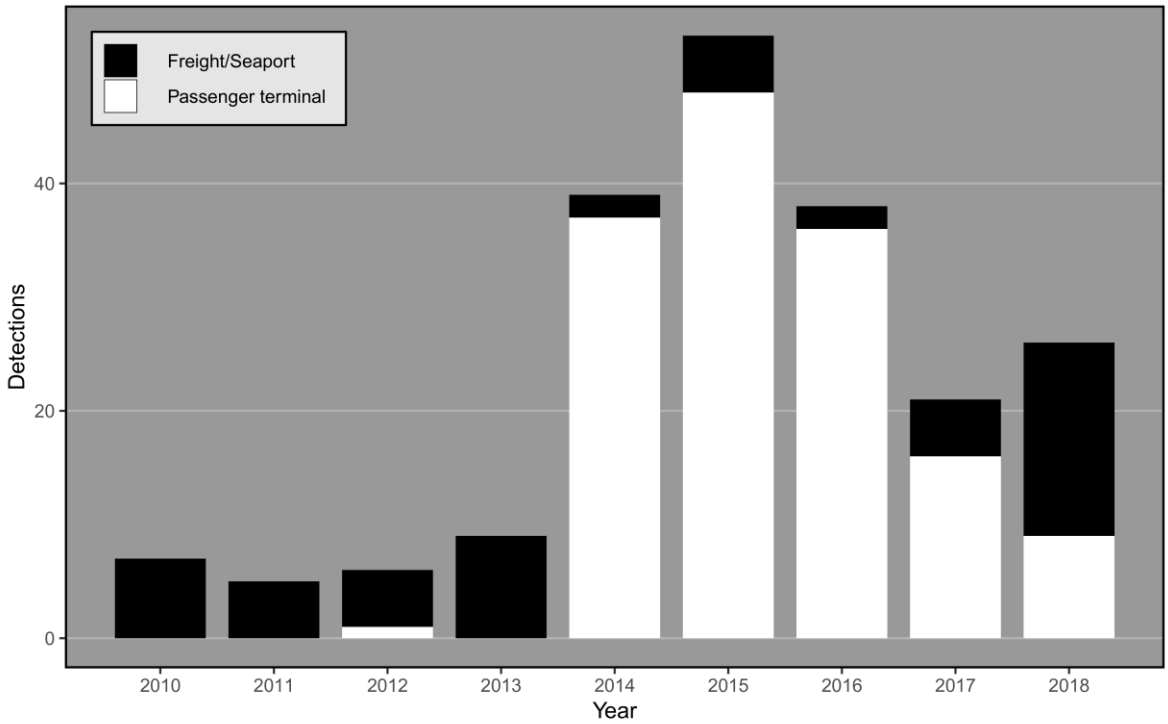


Figure 1: Number of *Ae. aegypti* detections per year at Australian international terminals. Detection at passenger airline terminals are shown in white, while those at seaports or freight terminals are in black. Each detection represents either a single adult retrieved from a trap or one or more larvae or pupae retrieved from an ovitrap.

When dispersing by flight alone, *Ae. aegypti* is a weak disperser that prefers to remain close to human dwellings and sites of activity [27]. Its current distribution has been facilitated by several centuries of human transport, including terrestrial [28], marine [29] and aerial [30] conveyance. Studies of *Ae. aegypti* using genome-wide SNPs have found genetic structure at fine scales of < 5 km [31] and strong population differentiation at broader scales [16, 32]. Similar SNP datasets to that used in this study have proved useful for investigating the source of newly established *Ae. aegypti* populations [29, 33]. The high level of genetic structuring at local and regional scales in this species suggests that genome-wide SNP analyses are likely to be informative for determining the origin of incursive mosquitoes, provided adequate reference mosquitoes are available from potential source populations.

The increasing rate of incursive *Ae. aegypti* detections at Australian airports has occurred despite protocols that ensure all international flights into Australia undergo disinsection procedures [34],

wherein insecticides are used to kill incursives through residual and/or knockdown treatments before or after take-off. However, dispersing pests from populations resistant to insecticides may be less affected by these measures. Although such resistance is not present in Australian *Ae. aegypti* [35], it is common in *Ae. aegypti* populations globally [36-38]. Point mutations at the *Vssc* gene can confer target-site knockdown resistance to synthetic pyrethroids in a broad range of insects [39-41] including *Ae. aegypti* [42]. The only two chemicals registered for aircraft disinsection by the World Health Organisation (WHO), permethrin and d-phenothrin [43, 44], are synthetic pyrethroids.

In this study, we use genome-wide SNPs to assign *Ae. aegypti* intercepted at international ports in Australia and New Zealand to their likely population, country, or region of origin by comparing incursives with 188 reference mosquitoes collected from 15 locations around the world. We also genotype incursive and reference mosquitoes for three different point mutations at the *Vssc* gene known to confer resistance to synthetic pyrethroids in *Ae. aegypti* [45]. Our results demonstrate how genomic data can elucidate risks associated with species and genetic incursions.

Results

Delineation of reference groupings

Monte-Carlo cross validation using the R package *assignPOP* [46] showed perfect assignment of resampled mosquitoes among 8 of the 18 reference samples, indicating that these samples were genetically distinct and could be treated as separate reference populations. However, for the time-separated pairs of reference samples from Bali, Kuala Lumpur, and Rio de Janeiro, resampled mosquitoes were frequently assigned to the incorrect sample of the pair. This indicates that geographical genetic structure observed at these locations was stable over time. Additionally, resampled mosquitoes from the Cairns and Townsville reference samples showed 10% incorrect assignment with each other, which is not surprising given weak genetic structure across this area [47], while 7% of resampled mosquitoes from the Ho Chi Minh City and Nha Trang reference samples were incorrectly assigned. Accordingly, each of these five pairs were combined into composite reference populations, leaving 13 reference populations for the assignment of incursives.

Geographical assignment of incursive *Ae. aegypti*

Discriminant analysis of principal components (DAPC [48]) cluster detection using 41,834 SNPs gave a range for the number of clusters (K) of $3 \leq K \leq 13$. At $K = 3$, populations were partitioned into a South American cluster, an Asian cluster, and a cluster containing Australia and the Pacific Islands. The Australian population split from the Pacific Islands at $K = 4$, following which new geographical partitions formed at each increase until $K = 13$ placed all reference populations into their own cluster. These same geographical partitions were established at $K = 13$ in each run, indicating that the inclusion of the single incursive was not biasing cluster detection.

Of the 115 incursives, 111 demonstrated clustering for $3 \leq K \leq 12$ that was consistent with their final cluster assignment at $K = 13$ (Appendix A). The 8 incursives collected at New Zealand terminals all showed consistent clustering with the Pacific Islands, while the incursives collected in Australia

clustered with Southeast Asia, Australia, and South America. Supplementary Figure A shows a typical stepwise delineation of clusters for $3 \leq K \leq 13$ for incursive IPT33.

Posterior probabilities of membership to the 13 reference populations calculated for incursives using *assignPOP* were of range 0.16 – 0.82. Relative probabilities (see Materials and Methods) ranged from 1.01 to 29.00. For incursives consistently assigned to either of the Asian or Pacific Islands clusters, we repeated assignment tests using only those reference populations, which allowed us to increase confidence in assignment by incorporating more SNPs: 55,443 and 146,142 respectively. Overall, 92 incursives (80%) had relative probabilities > 2, while 23 (20%) had relative probabilities < 2. Comparing assignment between DAPC and *assignPOP*, 112 of the 115 incursives (97.4%) had highest posterior probabilities for the same reference population they were assigned to at $K = 13$ (Appendix A), indicating strong agreement between the two assignment methods.

Figure 2 shows the relationship between posterior probability, relative probability, and proportion of missing data. The incursives with relative probabilities > 2 formed a group (white and black circles) in which posterior probability of membership decreased as missing data increased. Five incursives with relative probabilities > 2 were exceptions to this (black circles), grouping with the incursives with relative probabilities < 2 (black squares). This latter group had low posterior probabilities and mostly low missing data, with no correlation between the two ($R^2 = 0.023$, $F = 0.609$, $P = 0.442$). Among the incursives of the first group (white circles), the strong negative correlation between posterior probability of membership and proportion of missing data ($R^2 = 0.795$, $F = 329.6$, $P < 0.001$) indicates that lower posterior probabilities in this group were likely due to missing data (e.g. from imperfect preservation of specimens). Conversely, 25 of the 28 incursives in the second group (black circles and squares) had posterior probabilities < 0.5 despite having < 0.13 missing data (Figure 2). As the low posterior and relative probabilities of these incursives could not be explained by high missing data, it seems likely that these incursives were from source populations not sampled in this study.

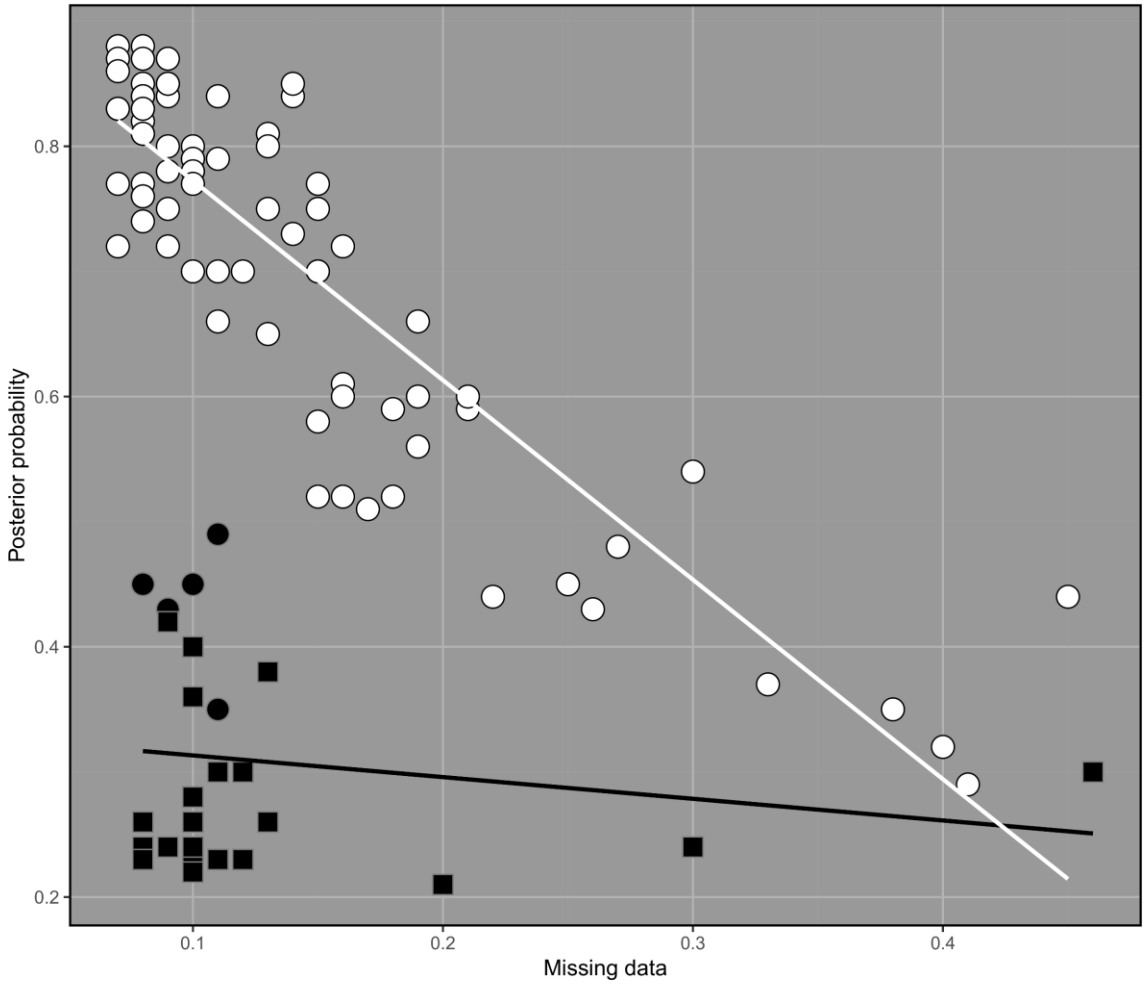


Figure 2: Posterior probabilities and proportion of missing data for the 115 incursive *Ae. aegypti*. White and black circles indicate incursives with relative probabilities > 2, while black squares indicate incursives with relative probabilities < 2. Incursives marked in white were considered well-assigned, while those marked in black were considered poorly-assigned. Among well-assigned incursives, lower posterior probabilities correlated strongly with higher missing data ($R^2 = 0.795$, $F = 329.6$, $P < 0.001$). Among poorly-assigned incursives there was no similar relationship ($R^2 = 0.023$, $F = 0.609$, $P = 0.442$).

We considered the 87 incursives (75.7%) represented as white circles in Figure 2 as well-assigned. These all showed consistent assignment for DAPC and *assignPOP* and consistent cluster assignment for $3 \leq K \leq 13$, and had relative probabilities > 2. Figure 3 shows a DAPC of the well-assigned incursives plotted with the 13 reference populations. The well-assigned incursives were assigned to the following source populations: 62 to Bali, 13 to Kuala Lumpur, 7 to Rio de Janeiro, 3 to Australia, 1 to Fiji and 1 to Bangkok (Appendix A). The incursive that was assigned to Fiji was collected in New Zealand and was the only well-assigned New Zealand incursive; the incursives assigned elsewhere were all collected in Australia. Supplementary Figure B shows a DAPC of the seven Asian reference populations (Bali, Bangkok, Kuala Lumpur, Singapore, Taiwan, Vietnam, Yogyakarta) plotted with the 76 well-assigned incursives putatively from Asia.

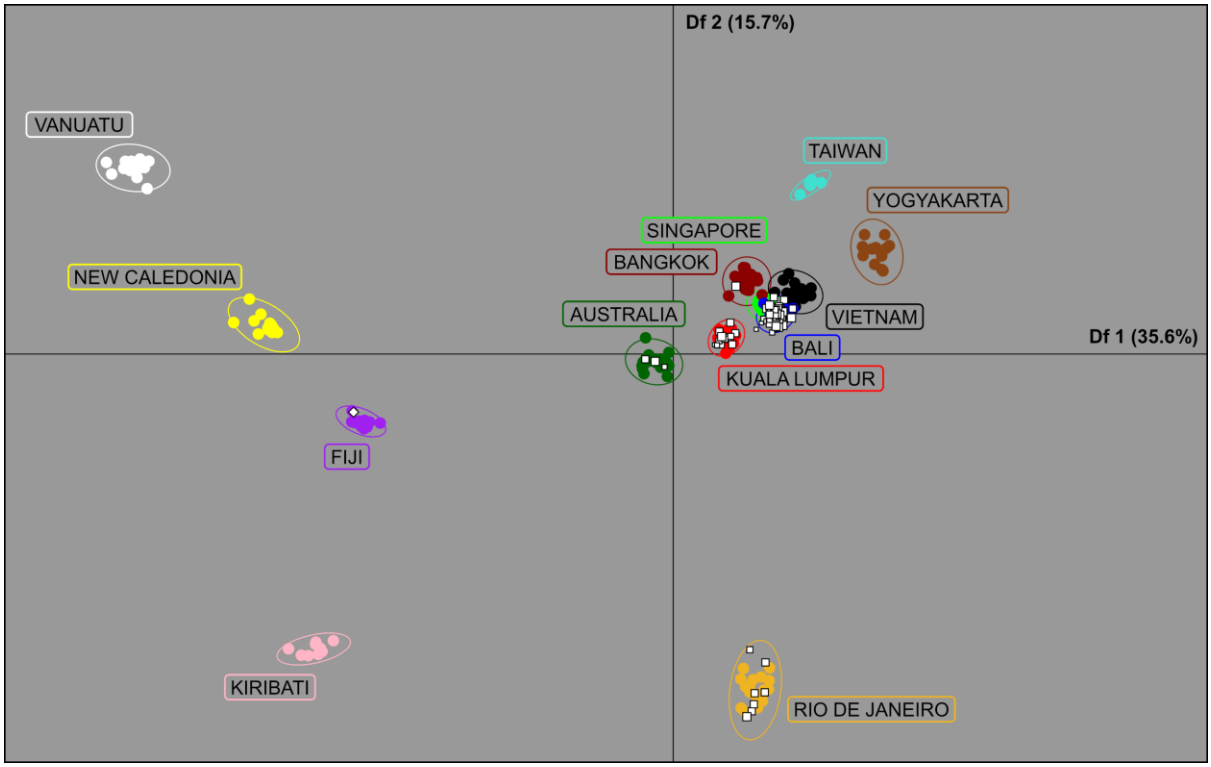


Figure 3: Discriminant analysis of principal components (DAPC). The plot shows the first and second discriminant functions of a DAPC of all 188 reference *Ae. aegypti* and all 87 well-assigned incursive *Ae. aegypti*, using 41,834 SNPs and 100 principal components. White squares in normal orientation show incursives intercepted at Australian terminals, while white squares rotated 45° show incursives intercepted at New Zealand terminals. Squares are sized by a logarithmic function indicating relative probability of membership (see Materials and Methods).

Vssc Genotyping

TaqMan® (Life Technologies, California, USA) assays for the three target site mutations showed that the most common genotype found in incursive mosquitoes across the three loci was homozygous resistant at V1016G, homozygous susceptible at F1534C, homozygous resistant at S989P (GG/TT/CC – the letters refer to the mutant or wildtype bases at each of the loci, rather than the amino acid notation). This genotype putatively confers strong resistance to Type I and Type II synthetic pyrethroids [45]. Of the 72 incursives with the GG/TT/CC genotype, 62 were assigned to Bali, 5 were assigned to Kuala Lumpur, 1 was assigned to Bangkok, and 4 were uncertain but had a likely Asian origin. As well as occurring in 100% of putatively Balinese incursives, this genotype was also found in all reference mosquitoes from Bali.

Another common genotype among incursives was homozygous susceptible at V1016G, homozygous resistant at F1534C, homozygous susceptible at S989P (TT/GG/TT), which putatively confers strong resistance to Type I synthetic pyrethroids [49]. This genotype was found in 16 incursive *Ae. aegypti* and 9 of the 11 Fijian reference mosquitoes. Of these incursives, seven were assigned to Rio de Janeiro, and one each to Bangkok and Fiji; the remainder were uncertain. Overall, 93 incursives (80.9%) had genotypes putatively conferring strong resistance to Type I synthetic pyrethroids.

Among all incursives, only three had the genotype that lacks all three resistance-associated point mutations (TT/TT/TT). These incursives were all well-assigned to the Australian reference population, indicating that they were likely transported on domestic rather than international flights. This genotype was also found in one of the Fijian reference mosquitoes.

Discussion

Assignment tests using genome-wide SNPs were able to identify common source populations from which pyrethroid-resistant *Ae. aegypti* disperse into Australia and New Zealand. We observed evidence for different incursion pathways into Australia and into New Zealand, as all the incursives intercepted in Australia were assigned to Southeast Asia, South America, and Australia, while all the incursives intercepted in New Zealand were assigned to the Pacific Islands. Among well-assigned incursives, most (71.2%) were assigned to Bali. All incursives assigned to Bali were collected airside at international passenger terminals in Australia, with the majority coming from Perth (56.3%), one of the most frequented routes between Australia and Denpasar International Airport in Bali (Figure 4). All *Ae. aegypti* suspected (incursives) and known (reference mosquitoes) to be from Bali shared a *Vssc* genotype (1016/1534/989; GG/TT/CC) linked to strong resistance to Type I and Type II synthetic pyrethroids [45]. Most of the suspected and known Fijian *Ae. aegypti* shared a genotype (TT/GG/TT) linked to strong resistance to Type I synthetic pyrethroids [49]. If the resistance observed in laboratory studies is indicative of field resistance, then WHO biosecurity protocols that rely upon disinsection with Type I synthetic pyrethroids [43, 44] are unlikely to stop the dispersal of *Ae. aegypti*, and will therefore be insufficient for preventing future incursions. The protocols used in this study may be implementable for *Ae. aegypti* and other invasive species internationally, which may help indicate the extent to which pyrethroid-based protocols are insufficient for achieving global biosecurity objectives.

Our results strongly suggest that resistance to synthetic pyrethroids used during disinsection is likely to be facilitating the dispersal of *Ae. aegypti* into Australia and New Zealand. There are currently only two chemicals registered for aircraft disinsection by the WHO, permethrin and d-phenothrin [43, 44]; both are Type I synthetic pyrethroids (permethrin is first generation, d-phenothrin is second generation). Of the 115 incursive mosquitoes analysed in this study, only three incursives had the wholly-susceptible genotype (TT/TT/TT). These were collected at Brisbane Airport, Australia, and were assigned to the Australian population, and thus are likely to have been transported on a domestic flight and will not have undergone disinsection as required on all international flights to Australia [34]. There is substantial geographical variability in pyrethroid resistance among *Ae. aegypti* populations [38], and considerable diversity in resistance genotypes within some of the reference populations in this study [36]. Therefore, our finding that all incursive mosquitoes from international sources had alleles associated with resistance to Type I pyrethroids, and that 80.9% were likely to have displayed strong levels of resistance, suggests that aircraft disinsection may be successfully eliminating pyrethroid-susceptible mosquitoes, while pyrethroid resistance is permitting others to survive and reach the Australian and New Zealand mainlands.

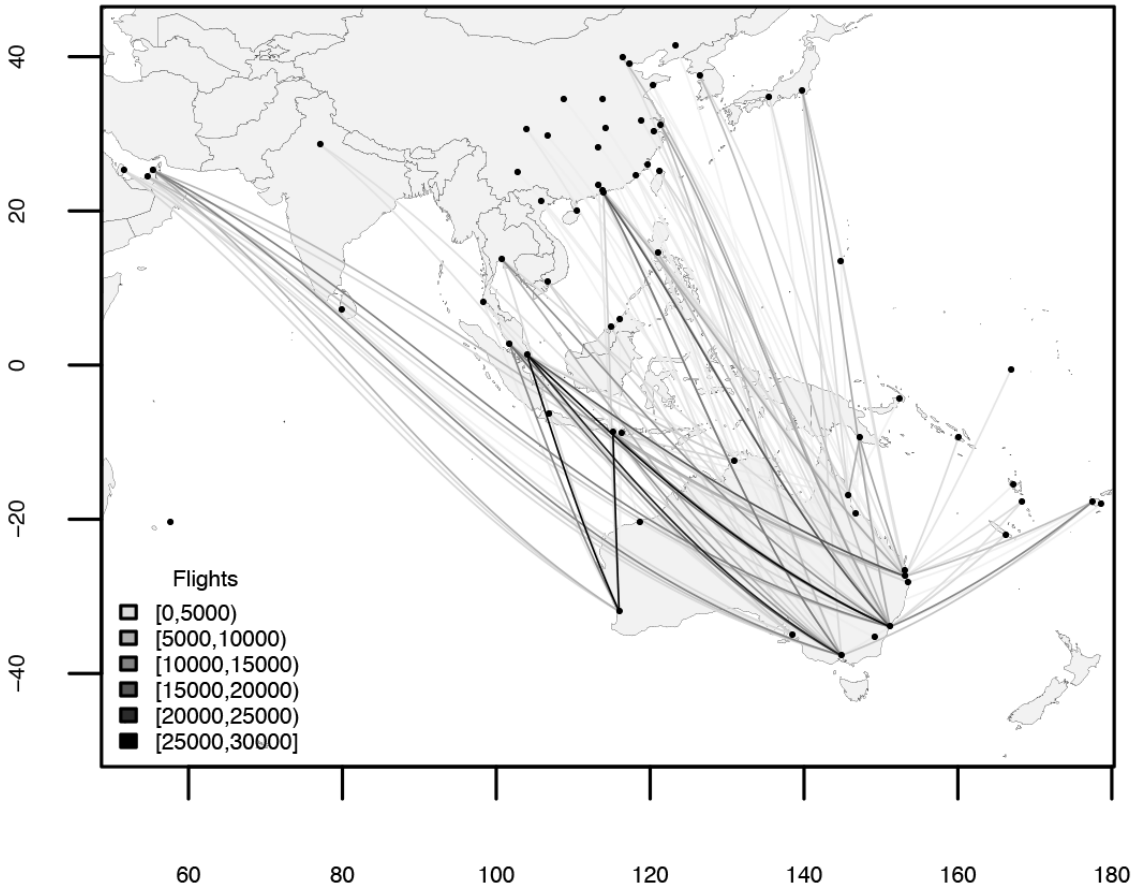


Figure 4: International passenger flights into Australia, 2014 – 2018. Total number of international passenger flights into Australian airports between January, 2014, and June, 2018, from regions in Asia, the South Pacific, and the Middle East.

International incursions of pyrethroid-resistant *Ae. aegypti* into Australia present three main management challenges. They could introduce alleles conferring pyrethroid resistance into the current, pyrethroid-susceptible Australian population and thereby reduce the effectiveness of current control programs; they could bring disease into non-endemic regions [50]; and they could lead to the establishment of new Australian *Ae. aegypti* populations outside the current native range of northern Queensland. Considering the current high frequency at which resistant *Ae. aegypti* are arriving in Australia, it is critical that future research focuses on determining whether the target site mutations provide 'operational resistance' to disinsection and whether this differs for the various methods used for disinsection (e.g. pre-flight and top-of-descent versus pre-embarkation/residual treatment; DAWR/MPI [34]). If operational resistance is shown to disinsection, then management of *Ae. aegypti* may need to focus more on suppressing populations at source locations. Interestingly, there has been a large reduction in the number of *Ae. aegypti* incursives detected at Australian ports over the last two years (Figure 1), which coincides with discussions between the Australian and Indonesian Governments as the results of this study were becoming available and may indicate a

change in density of local mosquito populations in Bali either through local control efforts or other local dynamics. Nevertheless, due to the likely high incidence of pyrethroid resistance in these populations, it may be more effective to use non-insecticidal methods to suppress them, such as by releasing *Ae. aegypti* males that are sterile or that carry an introduced infection of endosymbiotic bacteria [51]. Local suppression around Denpasar International Airport may also help reduce the dengue burden in this area, which has one of the highest incidences of dengue in Bali [52]. Non-insecticidal suppression may also be useful for eliminating small populations of resistant mosquitoes in newly-invaded regions.

This study has applied protocols for the assignment of incursive individuals to their likely population of origin. Following their application to *Ae. aegypti*, our protocols could be applied to other invasive species that are also frequently detected at borders. One candidate would be the related dengue vector *Ae. albopictus*, which has invaded and continues to invade tropical and temperate regions globally [21, 53], including the Torres Strait Islands in northern Australia [54], from which it threatens invasion of the mainland. This approach may be limited by the weak genetic structure between *Ae. albopictus* populations compared with *Ae. aegypti* [55], even when investigated with high-density markers [56, 57]. However, genome-level population genetic studies on *Ae. albopictus* are currently scarce, and a large database of reference samples may be necessary to identify distinct populations and produce confident assignments as in this study.

For invasive species with population differentiation equivalent to or greater than that of *Ae. aegypti*, the methods used in this study are likely to provide similar sensitivity for determining the origin of incursions. These methods could also be applied for post-border detections, where newly invasive species are detected and there is a need to determine their likely source. For instance, *Diuraphis noxia* (Russian wheat aphid) was first detected in Australia in 2016, and successful control of this species now depends, in part, on identifying past and present sources of incursions [58]. Likewise, genome-wide SNPs may help untangle the complexity of marine invasion narratives, in which many processes remain cryptic when analysed without the appropriate molecular tools [59].

Conclusions

Current WHO protocols for preventing the unintentional dispersal of invasive invertebrates rely upon Type I synthetic pyrethroids for disinsection of vessels. This study demonstrates that these measures are likely to have been insufficient for stopping the dispersal of a major disease vector, *Ae. aegypti*, into Australia and New Zealand through airports and seaports. Most *Ae. aegypti* collected in Australia came from Bali, while all those collected in New Zealand came from Pacific Islands like Fiji. A large majority of incursives (80.9%) had point mutations at the *Vssc* gene that putatively confer strong resistance to Type I or Type I and Type II synthetic pyrethroids. These mutations were found at high frequency in the greater Balinese and Fijian *Ae. aegypti* populations. Our findings indicate that insecticide-based strategies are likely to be limited in their capacity to stop biological invasions, and alternative control strategies should be considered, particularly for important disease vectors such as *Ae. aegypti.*

Materials and Methods

Mosquito collection

A total of 115 incursive *Ae. aegypti* were collected from passenger and freight terminals at airports and seaports in Australia and New Zealand between December 12th, 2012 and February 26th, 2018. Mosquitoes were considered to be incursive if they fit either of two conditions: i) they were collected at a terminal outside the native range of Far North Queensland; or ii) they were collected within the native range, but had resistance alleles at the *Vssc* gene not normally found in Australian *Ae. aegypti* [35]. Overall, only a single incursive mosquito from Far North Queensland was included (collected from Townsville seaport); all other incursives were collected outside the native range. The majority of incursives were collected at international airport passenger terminals (77.4%), generally around baggage areas airside. Appendix A lists data for each of the 115 incursives.

For the reference populations, we used 18 samples of *Ae. aegypti* collected throughout its range, with emphasis placed on regions with high frequency of passenger air traffic to Australia and New Zealand (Figure 4). This included samples from 15 locations: Bali (Indonesia), Yogyakarta (Indonesia), Kuala Lumpur (Malaysia), Bangkok (Thailand), Rio de Janeiro (Brazil), Cairns (Australia), Townsville (Australia), Nha Trang (Vietnam), Ho Chi Minh City (Vietnam), Fiji, Vanuatu, Kiribati, New Caledonia, Taiwan and Singapore. Bali, Kuala Lumpur and Rio de Janeiro were each sampled twice at different time points to determine the stability of genetic patterns. Where possible, we sampled close to airports with the highest frequency of passenger traffic into Australia. Appendix B lists data for each of the 13 reference samples.

DNA was extracted from mosquitoes using either Qiagen DNeasy Blood & Tissue Kits (Qiagen, Hilden, Germany) or Roche High Pure™ PCR Template Preparation Kit (Roche Molecular Systems, Inc., Pleasanton, CA, USA), each with an RNase treatment step. Extracted DNA was used for SNP detection and genotyping at the *Vssc* gene.

SNP detection and filtering

We applied the double digest restriction-site associated DNA sequencing (ddRADseq) protocol for *Ae. aegypti* developed by Rašić et al. [16] to construct RAD libraries. The ddRADseq approach provides markers with excellent resolution at intraspecific scales [60], and is therefore suitable for both broad and fine scale population assignment in this species, as well as for deciphering patterns of pairwise relatedness to help optimise analyses of genetic structure. Libraries were sequenced at either the Australian Genome Research Facility (AGRF, Melbourne, Australia) or the University of Melbourne (Pathology) on a HiSeq 2500 (Illumina, California, USA) in either High Output or Rapid Run modes.

We used the process_radtags program in Stacks v2.0 [61] to demultiplex sequence reads and trim the reads to 80 bp in length. Using a 15 bp sliding window, low quality reads were discarded if the average phred score dropped below 20. Reads were aligned to the *Ae. aegypti* nuclear genome assembly AaegL4 [62] with Bowtie v1.2.1.1 [62], allowing for a maximum of three mismatches (-v 3 --tryhard --best). The Stacks pipeline was used in three stages: initial filtering of reference mosquitoes;

establishment of reference populations from the 18 reference samples; and building a final catalog of incursive and reference mosquitoes for population assignment.

Initial filtering of reference samples aimed to remove closely related mosquitoes and those with high missing data, and to normalise the number of mosquitoes in each reference population. We used the Stacks ref_map pipeline to build individual Stacks catalogs for each of the 18 reference samples, from which we called genotypes at RAD stacks at a 0.05 significance level. We generated VCF files for each catalog with the Stacks program *populations*. SNPs were required to be present in ≥ 75% of the mosquitoes and have a minor allele frequency ≥ 0.05 (-r 0.75 --min_maf 0.05 --vcf). We used VCFtools [64] to remove loci that deviated from Hardy-Weinberg equilibrium, then thinned the data so that no two SNPs were within 250 kbp of one another. As the *Ae. aegypti* genome contains approximately 2.1 Mbp sequence per centimorgan [63], thinning the data in this way retains up to 8 SNPs per centimorgan, a density shown to sufficiently filter out linked loci [64]. Reference mosquitoes with ≥ 30% of missing genotypes were omitted from all further analyses. We used SPAGeDi [65] to calculate Loiselle's *k* [66] among reference mosquitoes, identifying pairs with putative first-degree relatedness ($k \geq 0.1875$ [67]) and omitting related mosquitoes in order of missing data so that all remaining pairs had $k < 0.1875$. Finally, for reference samples with more than 18 mosquitoes, we omitted mosquitoes in order of missing data to reduce the number to 18.

To establish reference populations from the reference samples, we generated a new catalog containing all of the filtered reference mosquitoes. We retained SNPs that were present in ≥ 75% of the mosquitoes in each reference sample and that had minor allele frequencies ≥ 0.05. We used the R package assignPOP v1.1.4 [68] to perform Monte-Carlo cross-validation on reference samples (functions "assign.MC" and "accuracy.MC"), treating all samples as putative populations. In this process, a proportion of mosquitoes was resampled from each putative population, and this subset of mosquitoes were assigned to populations using the remaining mosquitoes from each putative population. We ran this cross-validation multiple times, in turn resampling a proportion of 0.1, 0.2, and 0.3 of the mosquitoes from each reference sample, with 200 replicates run for each proportion. Reference samples were treated as distinct reference populations if resampled mosquitoes were consistently assigned to their correct sample. In cases where resampled mosquitoes were assigned across multiple reference samples, we created composite reference populations from the multiple samples. If any composite reference populations contained more than 18 mosquitoes, we omitted mosquitoes in order of missing data proportion to reduce the number to 18.

A final catalog was built containing the 115 incursives and the 188 reference mosquitoes retained after filtering. This catalog was used to produce subsets of SNPs for the analyses delineated below. In each subset, we retained SNPs with minor allele frequency ≥ 0.05 that were present in ≥ 75% of the mosquitoes in each reference population and in ≥ 75% of incursives.

Geographical assignment of incursive *Ae. aegypti*

Population assignment comprised two separate processes: cluster detection with DAPC, performed in the R package *adegenet* v2.1.1 [69]; and a Monte-Carlo assignment test using a support vector machine predictive model, performed in the R package *assignPOP* v1.1.4. We used DAPC to partition reference mosquitoes and incursives into sets of clusters, and from these partitions infer population

assignment. We used *assignPOP* to generate for each incursive a posterior probability of assignment to each of the reference populations, and from these probabilities infer confidence in assignment. Incursives showing inconsistent assignment for DAPC and *assignPOP* were not considered well-assigned.

For population assignment with DAPC, we used the *adegenet* function "find.clusters" to establish partitions within 115 subsets of mosquitoes. Each subset included all 188 reference mosquitoes and a single incursive. We identified an optimal range for the number of clusters (*K*) to use for partitioning, using 1,000 principal components and 10^9 iterations for each *K*. The lower bound of this range was set as the value of *K* with the lowest Bayesian information criterion (BIC) and the upper bound was the value with the lowest Akaike information criterion (AIC), which provided a range of *K* from the conservative BIC to the less conservative AIC [70]. For each subset and for each value of *K* within this range, the single incursive formed a cluster containing the incursive and one or more of the reference populations. This clustering indicated the assignment of the incursive for that value of *K*. Following this, we performed cluster detection among closely related reference populations, which allowed the use of more SNPs and thus more confident assignment among populations within a subregion. For incursives initially assigned to Asian populations (Bali, Bangkok, Kuala Lumpur, Singapore, Taiwan, Vietnam, and Yogyakarta), we performed cluster detection using only the Asian reference populations, and did likewise for incursives initially assigned to the Pacific Islands (Fiji, Kiribati, New Caledonia, and Vanuatu).

For population assignment in *assignPOP*, we used the function "assign.X" to perform assignment tests on incursives. This analysis used the 188 reference mosquitoes to build a predictive model with which to assign the 115 incursives, using all principal components with eigenvalues > 1, and a linear support vector machine classification function. This generated posterior probabilities of membership for each individual to each reference population. From each of these scores we calculated 'relative probability' of membership, which was defined as the probability of membership to the most likely population divided by the probability of membership to the second most likely population. We used these relative probabilities as an additional measure of confidence in assignment. For an incursive to be considered well-assigned it would need to have a relative probability > 2; that is, the posterior probability of its assigned reference population would have to be at least twice that of the next most likely population.

Vssc Genotyping

All incursives were genotyped for the presence of three target site mutations (V1016G, F1534C, S989P) in the *Vssc* gene known to provide resistance to synthetic pyrethroids [45, 49]. We also genotyped mosquitoes from the Bali and Fiji reference populations. The mutation site amino acid positions in this study are labelled according to the sequence of the most abundant splice variant of the house fly, *Musca domestica*, *Vssc* (GenBank accession nos. AAB47604 and AAB47605) [71]. The V1016G mutation is widespread in Southeast Asia and is known to provide resistance to Type I and Type II synthetic pyrethroids when in the homozygous (GG) state [45]. The F1534C mutation is thought to confer resistance to Type I synthetic pyrethroids [49] when homozygous. The S989P mutation alone has no effect, but in the presence of the V1016G mutation it is thought to act

synergistically, increasing resistance to Type II synthetic pyrethroids (but not Type I synthetic pyrethroids) [72]. We developed TaqMan® assays for all three target site mutations based on the high resolution melt genotyping assays in Wuliandari et al. [45]. Genotyping of each SNP was then undertaken on a LightCycler® 480 (Roche, Basel, Switzerland) 480 real time PCR machine with three replicates in a 384-well plate format.

Acknowledgements

We thank Ashley Callahan, Jason Axford, Tim Hurst, Stephen Doggett, Elizabeth Valerie, Craig Williams and Joe Davis for the collection of reference samples used in this study. We also thank Ashley Callahan for conducting some of the DNA extraction and library preparation. We thank Gordana Rašić for input in conceptualisation and study design. We thank staff from the Department of Agriculture and Water Resources (DAWR), Australian Government, for the collection of all incursive mosquitoes and DAWR for providing funding for this work. AAH was supported by Program and Fellowship grants from the National Health and Medical Research Council (NHMRC). TLS and AAH were also supported by the Wellcome Trust UK. These grants provided the necessary support to establish the reference database used in this study.

Author contributions

Tom L. Schmidt

Data Curation, Formal Analysis, Investigation, Methodology, Software, Validation, Visualization, Writing – Original Draft Preparation, Writing – Review & Editing

Anthony R. van Rooyen

Data Curation, Investigation, Methodology, Validation, Writing – Review & Editing

Jessica Chung

Data Curation, Resources, Software, Writing – Review & Editing

Nancy M. Endersby-Harshman

Data Curation, Investigation, Methodology, Project Administration, Validation, Writing – Original Draft Preparation, Writing – Review & Editing

Philippa C. Griffin

Conceptualization, Data Curation, Formal Analysis, Methodology, Writing – Review & Editing

Angus Sly

Funding Acquisition, Project Administration, Resources

Ary A. Hoffmann

Conceptualization, Funding Acquisition, Methodology, Project Administration, Resources, Supervision, Writing – Original Draft Preparation, Writing – Review & Editing

Andrew R. Weeks

Conceptualization, Funding Acquisition, Methodology, Project Administration, Resources, Supervision, Visualization, Writing – Original Draft Preparation, Writing – Review & Editing

References

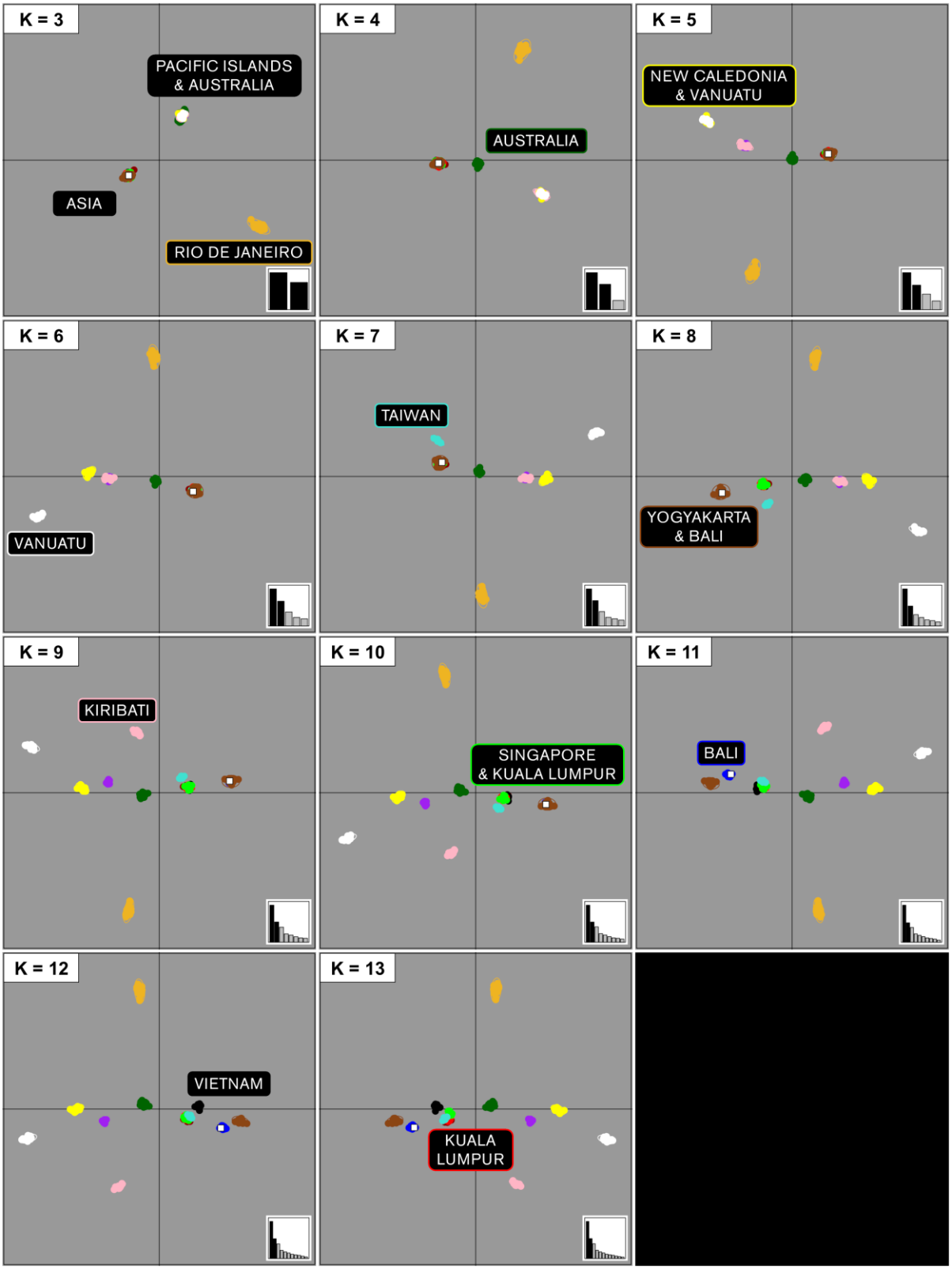
1. IUCN, editor Guidelines for the prevention of biodiversity loss caused by alien invasive species. International Union for the Conservation of Nature; 2000; Gland, Switzerland.
2. Raizada P, Raghubanshi A, Singh J. Impact of invasive alien plant species on soil processes: a review. Proc Natl Acad Sci India Sect B Biol Sci. 2008;78:288-98.
3. Hejda M, Pyšek P, Jarošík V. Impact of invasive plants on the species richness, diversity and composition of invaded communities. J Ecol. 2009;97(3):393-403.
4. Blackburn TM, Cassey P, Duncan RP, Evans KL, Gaston KJ. Avian extinction and mammalian introductions on oceanic islands. Science. 2004;305(5692):1955-8.
5. Lounibos LP. Invasions by insect vectors of human disease. Annu Rev Entomol. 2002;47(1):233-66.
6. Levine JM, D'antonio CM. Forecasting biological invasions with increasing international trade. Conserv Biol. 2003;17(1):322-6.
7. Westphal MI, Browne M, MacKinnon K, Noble I. The link between international trade and the global distribution of invasive alien species. Biol Invasions. 2008;10(4):391-8.
8. Hulme PE. Trade, transport and trouble: managing invasive species pathways in an era of globalization. J Appl Ecol. 2009;46(1):10-8.
9. McGeoch MA, Butchart SH, Spear D, Marais E, Kleynhans EJ, Symes A, et al. Global indicators of biological invasion: species numbers, biodiversity impact and policy responses. Divers Distrib. 2010;16(1):95-108.
10. Caley P, Ingram R, De Barro P. Entry of exotic insects into Australia: Does border interception count match incursion risk? Biol Invasions. 2015;17(4):1087-94.
11. Bacon SJ, Bacher S, Aebi A. Gaps in border controls are related to quarantine alien insect invasions in Europe. PLoS One. 2012;7(10):e47689.
12. Robinson A, Burgman MA, Cannon R. Allocating surveillance resources to reduce ecological invasions: maximizing detections and information about the threat. Ecol Appl. 2011;21(4):1410-7.
13. Myers JH, Simberloff D, Kuris AM, Carey JR. Eradication revisited: dealing with exotic species. Trends Ecol Evol. 2000;15(8):316-20.
14. Mehta SV, Haight RG, Homans FR, Polasky S, Venette RC. Optimal detection and control strategies for invasive species management. Ecol Econ. 2007;61(2-3):237-45.
15. Wilson JR, Dormontt EE, Prentis PJ, Lowe AJ, Richardson DM. Something in the way you move: dispersal pathways affect invasion success. Trends Ecol Evol. 2009;24(3):136-44.
16. Rašić G, Filipović I, Weeks AR, Hoffmann AA. Genome-wide SNPs lead to strong signals of geographic structure and relatedness patterns in the major arbovirus vector, *Aedes aegypti*. BMC Genomics. 2014;15(1).
17. Puckett EE, Eggert LS. Comparison of SNP and microsatellite genotyping panels for spatial assignment of individuals to natal range: A case study using the American black bear (*Ursus americanus*). Biol Conserv. 2016;193:86-93.
18. Smith C. The history of dengue in tropical Asia and its probable relationship to the mosquito *Aedes aegypti*. J Trop Med Hyg. 1956;59(10):243-51.
19. Bogoch II, Brady OJ, Kraemer M, German M, Creatore MI, Kulkarni MA, et al. Anticipating the international spread of Zika virus from Brazil. Lancet. 2016;387(10016):335-6.
20. Weaver SC, Lecuit M. Chikungunya virus and the global spread of a mosquito-borne disease. N Engl J Med. 2015;372(13):1231-9.

21. Kraemer MU, Sinka ME, Duda KA, Mylne AQ, Shearer FM, Barker CM, et al. The global distribution of the arbovirus vectors *Aedes aegypti* and *Ae. albopictus*. Elife. 2015;4.
22. Russell RC, Currie BJ, Lindsay MD, Mackenzie JS, Ritchie SA, Whelan PI. Dengue and climate change in Australia: predictions for the future should incorporate knowledge from the past. Med J Aust. 2009;190:265-8.
23. Kearney M, Porter WP, Williams C, Ritchie S, Hoffmann AA. Integrating biophysical models and evolutionary theory to predict climatic impacts on species' ranges: the dengue mosquito *Aedes aegypti* in Australia. Funct Ecol. 2009;23(3):528-38.
24. Preston BL, Jones RN. Climate change impacts on Australia and the benefits of early action to reduce global greenhouse gas emissions: CSIRO Canberra; 2006.
25. McMichael AJ, Woodruff RE, Hales S. Climate change and human health: present and future risks. Lancet. 2006;367(9513):859-69.
26. Laird M. Insects collected from Aircraft arriving in New Zealand from Abroad. Zoology Publications, Victoria University Collection. 1951;11:30.
27. Harrington LC, Scott TW, Lerdthusnee K, Coleman RC, Costero A, Clark GG, et al. Dispersal of the dengue vector *Aedes aegypti* within and between rural communities. Am J Trop Med Hyg. 2005;72(2):209-20.
28. Guagliardo SA, Barboza JL, Morrison AC, Astete H, Vazquez-Prokopec G, Kitron U. Patterns of geographic expansion of *Aedes aegypti* in the Peruvian Amazon. PLoS Negl Trop Dis. 2014;8(8):e3033.
29. Brown JE, Evans BR, Zheng W, Obas V, Barrera-Martinez L, Egizi A, et al. Human impacts have shaped historical and recent evolution in *Aedes aegypti*, the dengue and yellow fever mosquito. Evol. 2014;68(2):514-25.
30. Sukehiro N, Kida N, Umezawa M, Murakami T, Arai N, Jinnai T, et al. First report on invasion of yellow fever mosquito, *Aedes aegypti,* at Narita International Airport, Japan in August 2012. Jpn J Infect Dis. 2013;66(3):189-94.
31. Schmidt TL, Filipović I, Hoffmann AA, Rašić G. Fine-scale landscape genomics helps explain the slow spatial spread of *Wolbachia* through the *Aedes aegypti* population in Cairns, Australia. Heredity. 2018;120.
32. Rašić G, Endersby-Harshman N, Tantowijoyo W, Goundar A, White V, Yang Q, et al. *Aedes aegypti* has spatially structured and seasonally stable populations in Yogyakarta, Indonesia. Parasit Vectors. 2015;8(1):610.
33. Gloria-Soria A, Lima A, Lovin DD, Cunningham JM, Severson DW, Powell JR. Origin of a High-Latitude Population of *Aedes aegypti* in Washington, DC. Am J Trop Med Hyg. 2018;98(2):445-52.
34. DAWR/MPI. Schedule of aircraft disinsection procedures for flights into Australia and New Zealand. Version 4.0. Department of Agriculture and Water Resources, Canberra, and Ministry for Primary Industries, Wellington; 2016.
35. Endersby-Harshman NM, Wuliandari JR, Harshman LG, Frohn V, Johnson BJ, Ritchie SA, et al. Pyrethroid Susceptibility Has Been Maintained in the Dengue Vector, *Aedes aegypti* (Diptera: Culicidae), in Queensland, Australia. J Med Entomol. 2017;54(6):1649-58.
36. Endersby-Harshman NM, Weeks AR, Hoffmann AA. The detection and significance of emerging insecticide resistance in mosquitoes. Microbiol Aust. 2018.
37. Ranson H, Burhani J, Lumjuan N, Black IV WC. Insecticide resistance in dengue vectors. TropIKAnet. 2010;1(1):0-.

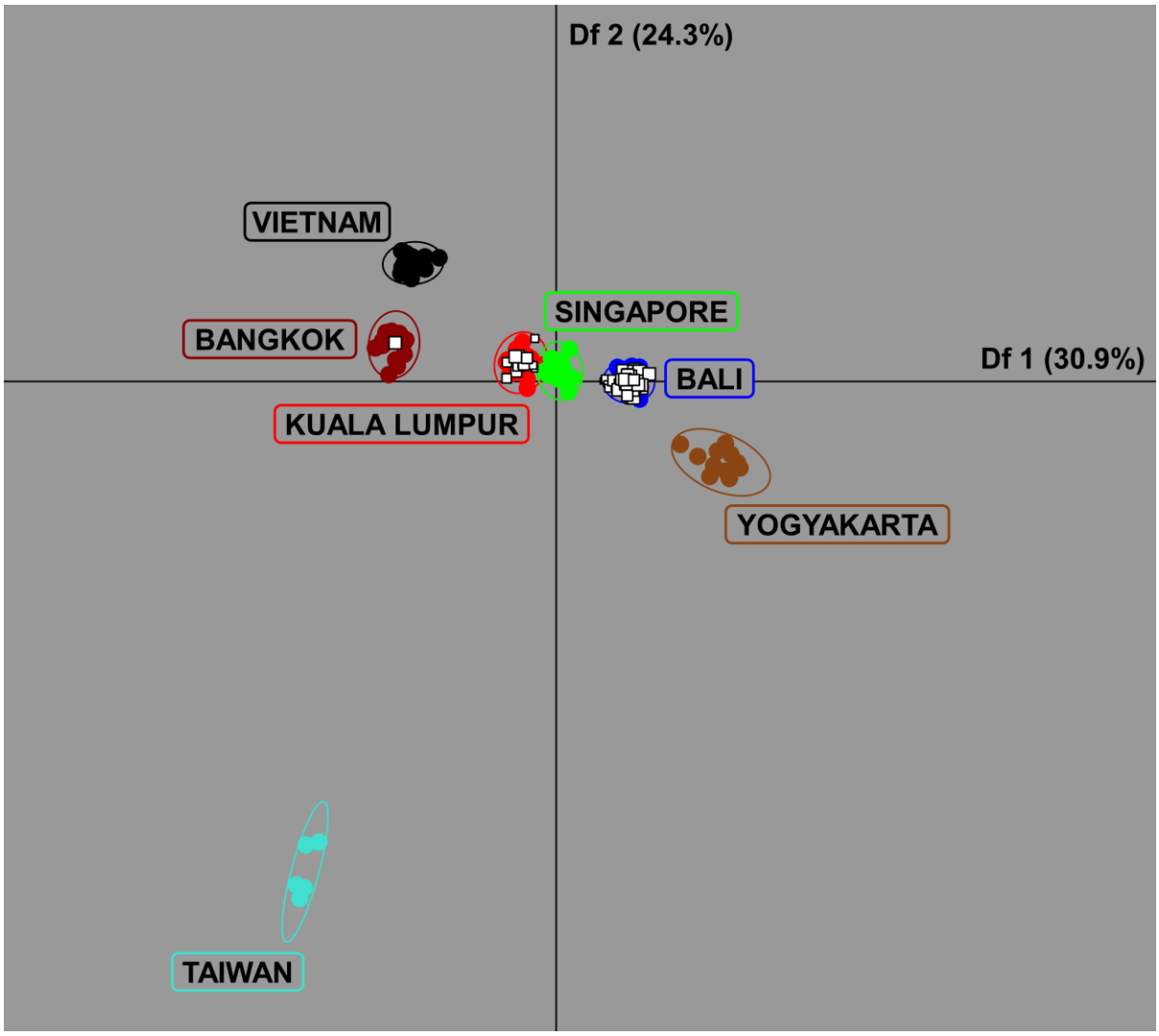
38. Smith LB, Kasai S, Scott JG. Pyrethroid resistance in *Aedes aegypti* and *Aedes albopictus*: Important mosquito vectors of human diseases. Pestic Biochem Physiol. 2016;133:12.
39. Scott JG, Leichter CA, Rinkevihc FD, Harris SA, Su C, Aberegg LC, et al. Insecticide resistance in house flies from the United States: Resistance levels and frequency of pyrethroid resistance alleles. Pestic Biochem Physiol. 2013;107(3):377-84.
40. Scott JG, Yoshimizu MH, Kasai S. Pyrethroid resistance in *Culex pipiens* mosquitoes. Pestic Biochem Physiol. 2015;120:68-76.
41. Seong KM, Lee D-Y, Yoon KS, Kwon DH, Kim HC, Klein TA, et al. Establishment of quantitative sequencing and filter contact vial bioassay for monitoring pyrethroid resistance in the common bed bug, *Cimex lectularius*. J Med Entomol. 2014;47(4):592-9.
42. Donnelly MJ, Corbel V, Weetman D, Wilding CS, Williamson MS, Black IV WC. Does kdr genotype predict insecticide-resistance phenotype in mosquitoes? Trends Parasitol. 2009;25(5):213-9.
43. WHO. Guidelines for testing the efficacy of insecticide products used in aircraft. Gevena: World Health Organization; 2012.
44. IPCS. Aicraft disinsection insecticides (Environmental health criteria 243). United Kingdom Department of Health: European Commission and the Policy Research Programme; 2013.
45. Wuliandari JR, Lee SF, White VL, Tantowijoyo W, Hoffmann AA, Endersby-Harshman NM. Association between three mutations, F1565C, V1023G and S996P, in the voltage-sensitive sodium channel gene and knockdown resistance in *Aedes aegypti* from Yogyakarta, Indonesia. Insects. 2015;6(3):658-85.
46. Chen K, Marschall E, Sovic M, Fries A, Gibbs H, Ludsin S. *assignPOP*: An R package for population assignment using genetic, non-genetic, or integrated data in a machine-learning framework. Methods Ecol Evol. 2018;9(2):439-46.
47. Endersby N, Hoffmann A, White V, Ritchie S, Johnson P, Weeks A. Changes in the genetic structure of *Aedes aegypti* (Diptera: Culicidae) populations in Queensland, Australia, across two seasons: implications for potential mosquito releases. J Med Entomol. 2011;48(5):999-1007.
48. Jombart T, Devillard S, Balloux F. Discriminant analysis of principal components: a new method for the analysis of genetically structured populations. BMC Genet. 2010;11(1):94.
49. Du Y, Nomura Y, Satar G, Hu Z, Nauen R, He SY, et al. Molecular evidence for dual pyrethroid-receptor sites on a mosquito sodium channel. Proc Natl Acad Sci. 2013;110(29):11785-90.
50. Siala E, Gamara D, Kallel K, Daaboub J, Zouiten F, Houzé S, et al. Airport malaria: report of four cases in Tunisia. Malar J. 2015;14(1):42.
51. McGraw EA, O'Neill SL. Beyond insecticides: new thinking on an ancient problem. Nat Rev Microbiol. 2013;11(3):181-93.
52. Purnama SG, Baskoro T. Maya Index and density of *Aedes aegypti* larvae of the dengue infection. Makara J Health Res. 2013;16(2):57-64.
53. Benedict MQ, Levine RS, Hawley WA, Lounibos LP. Spread of the tiger: global risk of invasion by the mosquito *Aedes albopictus*. Vector Borne Zoonotic Dis. 2007;7(1):76-85.
54. Ritchie SA, Moore P, Carruthers M, Williams C, Montgomery B, Foley P, et al. Discovery of a widespread infestation of *Aedes albopictus* in the Torres Strait, Australia. J Am Mosq Control Assoc. 2006;22(3):358-65.
55. Goubert C, Minard G, Vieira C, Boulesteix M. Population genetics of the Asian tiger mosquito *Aedes albopictus*, an invasive vector of human diseases. Heredity. 2016;117(3):125–34.

56. Kotsakiozi P, Richardson JB, Pichler V, Favia G, Martins AJ, Urbanelli S, et al. Population genomics of the Asian tiger mosquito, *Aedes albopictus*: insights into the recent worldwide invasion. Ecol Evol. 2017;7(23):10143-57.
57. Schmidt TL, Rašić G, Zhang D, Zheng X, Xi Z, Hoffmann AA. Genome-wide SNPs reveal the drivers of gene flow in an urban population of the Asian Tiger Mosquito, *Aedes albopictus*. PLoS Negl Trop Dis. 2017;11(10):e0006009.
58. Yazdani M, Baker G, DeGraaf H, Henry K, Hill K, Kimber B, et al. First detection of Russian wheat aphid *Diuraphis noxia* Kurdjumov (Hemiptera: Aphididae) in Australia: a major threat to cereal production. Aust Entomol. 2017.
59. Ojaveer H, Galil BS, Minchin D, Olenin S, Amorim A, Canning-Clode J, et al. Ten recommendations for advancing the assessment and management of non-indigenous species in marine ecosystems. Mar Policy. 2014;44:160-5.
60. Peterson BK, Weber JN, Kay EH, Fisher HS, Hoekstra HE. Double digest RADseq: an inexpensive method for de novo SNP discovery and genotyping in model and non-model species. PLoS One. 2012;7(5):e37135.
61. Catchen J, Hohenlohe PA, Bassham S, Amores A, Cresko WA. Stacks: an analysis tool set for population genomics. Mol Ecol. 2013;22(11):3124-40.
62. Dudchenko O, Batra SS, Omer AD, Nyquist SK, Hoeger M, Durand NC, et al. De novo assembly of the *Aedes aegypti* genome using Hi-C yields chromosome-length scaffolds. Science. 2017;356(6333):92-5.
63. Brown S, Severson D, Smith L, Knudson D. Integration of the *Aedes aegypti* mosquito genetic linkage and physical maps. Genetics. 2001;157(3):1299-305.
64. Cho K, Dupuis J. Handling linkage disequilibrium in qualitative trait linkage analysis using dense SNPs: a two-step strategy. BMC Genet. 2009;10(1):44.
65. Hardy OJ, Vekemans X. SPAGeDi: a versatile computer program to analyse spatial genetic structure at the individual or population levels. Mol Ecol Notes. 2002;2(4):618-20.
66. Loiselle BA, Sork VL, Nason J, Graham C. Spatial genetic structure of a tropical understory shrub, *Psychotria officinalis* (Rubiaceae). Am J Bot. 1995;82:1420-5.
67. Iacchei M, Ben-Horin T, Selkoe KA, Bird CE, García-Rodríguez FJ, Toonen RJ. Combined analyses of kinship and FST suggest potential drivers of chaotic genetic patchiness in high gene-flow populations. Mol Ecol. 2013;22(13):3476-94.
68. Chen KY, Marschall EA, Sovic MG, Fries AC, Gibbs HL, Ludsin SA. assignPOP: An r package for population assignment using genetic, non-genetic, or integrated data in a machine-learning framework. Methods Ecol Evol. 2018;9(2):439-46.
69. Jombart T. adegenet: a R package for the multivariate analysis of genetic markers. Bioinformatics. 2008;24(11):1403-5.
70. Raftery AE. Bayes factors and BIC: Comment on “A critique of the Bayesian information criterion for model selection”. Sociol Methods Res. 1999;27(3):411-27.
71. Kasai S, Komagata O, Itokawa K, Shono T, Ng LC, Kobayashi M, et al. Mechanisms of pyrethroid resistance in the dengue mosquito vector, *Aedes aegypti*: target site insensitivity, penetration, and metabolism. PLOS Neglected Tropical Diseases. 2014;8(6):e2948.
72. Hirata K, Komagata O, Itokawa K, Yamamoto A, Tomita T, Kasai S. A single crossing-over event in voltage-sensitive Na+ channel genes may cause critical failure of dengue mosquito control by insecticides. PLoS Negl Trop Dis. 2014;8(8):e3085.

Supplementary Figure A: A typical stepwise delineation of clusters for 3 ≤ *K* ≤ 13 for incursive IPT33.



Supplementary Figure B: DAPC of the Asian reference populations and the 76 well-assigned incursives putatively from Asia.



Appendix A: Relevant data for each of the 115 incursives.
Appendix B: Relevant data for each of the 13 reference samples.