Using spatial genetics to quantify mosquito dispersal for control programs

Igor Filipović1*, Hapuarachchige Chanditha Hapuarachchi2†, Wei-Ping Tien2, Muhammad Aliff Bin Abdul Razak2, Caleb Lee2, Cheong Huat Tan2, Gregor J. Devine1 and Gordana Rašić1*

RESEARCH ARTICLE

Abstract

Background: Hundreds of millions of people get a mosquito-borne disease every year and nearly one million die. Transmission of these infections is primarily tackled through the control of mosquito vectors. The accurate quantification of mosquito dispersal is critical for the design and optimization of vector control programs, yet the measurement of dispersal using traditional mark-release-recapture (MRR) methods is logistically challenging and often unrepresentative of an insect’s true behavior. Using Aedes aegypti (a major arboviral vector) as a model and two study sites in Singapore, we show how mosquito dispersal can be characterized by the spatial analyses of genetic relatedness among individuals sampled over a short time span without interruption of their natural behaviors.

Results: Using simple oviposition traps, we captured adult female Ae. aegypti across high-rise apartment blocks and genotyped them using genome-wide SNP markers. We developed a methodology that produces a dispersal kernel for distance which results from one generation of successful breeding (effective dispersal), using the distance separating full siblings and 2nd- and 3rd-degree relatives (close kin). The estimated dispersal distance kernel was exponential (Laplacian), with a mean dispersal distance (and dispersal kernel spread $\sigma$) of 45.2 m (95% CI 39.7–51.3 m), and 10% probability of a dispersal > 100 m (95% CI 92–117 m). Our genetically derived estimates matched the parametrized dispersal kernels from previous MRR experiments. If few close kin are captured, a conventional genetic isolation-by-distance analysis can be used, as it can produce $\sigma$ estimates congruent with the close-kin method if effective population density is accurately estimated. Genetic patch size, estimated by spatial autocorrelation analysis, reflects the spatial extent of the dispersal kernel “tail” that influences, for example, the critical radii of release zones and the speed of Wolbachia spread in mosquito replacement programs.

Conclusions: We demonstrate that spatial genetics can provide a robust characterization of mosquito dispersal. With the decreasing cost of next-generation sequencing, the production of spatial genetic data is increasingly accessible. Given the challenges of conventional MRR methods, and the importance of quantified dispersal in operational vector control decisions, we recommend genetic-based dispersal characterization as the more desirable means of parameterization.

Keywords: Mosquito dispersal, Genome-wide SNPs, Close kin, Dispersal kernel, IBD, Spatial autocorrelation

© The Author(s). 2020 Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit http://creativecommons.org/licenses/by/4.0/. The Creative Commons Public Domain Dedication waiver (http://creativecommons.org/publicdomain/zero/1.0/) applies to the data made available in this article, unless otherwise stated in a credit line to the data.
Background
Mosquitoes’ ability to carry and transmit human pathogens (malaria and filarial parasites, arboviruses) causes hundreds of millions of infections and nearly one million fatalities every year [1]. Both prevention and mitigation of many mosquito-borne disease outbreaks are primarily reliant on the control of mosquito vectors. Most of these interventions are designed to impact mosquito abundance or daily survival by targeting immature and adult stages. In the case of the major arbovirus vector *Aedes aegypti*, an urban-dwelling container-breeding mosquito, conventional control approaches include removal and treatment of larval habitats, as well as elimination of adults through insecticide application (indoor-residual and space spraying) and lethal trapping [2]. New biocontrol strategies undergoing field evaluations include RIDL® and Wolbachia®-based population replacement and suppression [3–5].

Defining the optimal area for treatment or mosquito release is one of the key considerations when implementing a public health intervention or designing a field-trial for a new control approach. For example, to contain the spread of dengue virus during an outbreak, focal insecticide-based control of *Ae. aegypti* adults is typically conducted at and around the main and secondary residences of dengue cases. The radius of the area to be treated is informed by the average dispersal distance of potentially infected female *Ae. aegypti* [6]. Understanding the ability of released sterile male mosquitoes to disperse and mate in an area being targeted by a suppression strategy is essential for predicting the required release pattern [7]. Additionally, sustained suppression in a target zone is difficult if a surrounding buffer zone is too small to prevent immigration by gravid wild-type females from neighboring areas. Similarly, stable introduction of a virus-blocking *Wolbachia* may fail if the release area of *Wolbachia*-infected *Ae. aegypti* is too small and too vulnerable to immigration by wild-type mosquitoes [8]. For the emerging genetic-based control approaches such as gene drive systems [9], well-characterized mosquito dispersal is crucial for addressing the biosafety concerns around the systems’ confineability and reversibility in the field [10, 11].

Quantifying mosquito dispersal of both wild-type and introduced mosquito strains in any given landscape is, therefore, critical for the considerations of the size of the treatment area and the surrounding buffer zones. Those considerations complement practical operational deliberations of the availability of human and economic resources for implementation and the sample sizes required to capture epidemiological endpoints [12, 13].

Mosquito dispersal characteristics have been typically studied using conventional mark-release-recapture (MRR) experiments utilizing powders and paints on trapped or laboratory-reared adult mosquitoes [14]. Location of the recaptured marked insects relative to their release point is typically used to estimate the mean distance traveled (MDT), and the distance within which 50% or 90% of mosquitoes are expected to disperse (FR50 and FR90, respectively). Fewer MRR studies have incorporated the dispersal kernel theory to estimate the distribution of dispersal distances over the whole flight range [7, 15, 16]. MRR experiments in *Ae. aegypti* have reported the mean dispersal distance to range from tens to hundreds of meters [14], and this variation points to a need to characterize dispersal locally so that the optimal control can be deployed in a given landscape. However, MRR experiments are operationally demanding, and the rearing and marking procedure can alter mosquito fitness and movement behavior in the field [17]. Additionally, the release of biting vector-competent females might pose an unacceptable risk of increased pathogen transmission in endemic areas.

Here we show how information on mosquito dispersal characteristics can be obtained from the spatial patterns of genetic distance and relatedness among sampled individuals, providing an alternative to the MRR experiments for informing the mosquito control programs. In contrast to conventional MRR approaches that require insect trapping or rearing, followed by mark, release, and recapture, the genetic approach requires only insect capture, utilizing the information from genetic markers and spatial location of individuals sampled continuously across an area over a limited time span and without manipulation or interruption of their natural behaviors.

In social insects like bumblebees, queen dispersal distance has been estimated by comparing the locations of workers (sampled in summer) and queens (sampled in the following spring) that were identified as sisters through sibling reconstruction analysis with genetic markers [18]. Inferences about mosquito ovipositing behavior have also been made using the genetic reconstruction of sibling groups, where the distance between full siblings sampled from different larval sites directly reflects mother’s movement distance during her skip-oviposition [19–21]. Here we show that the distance separating not only full siblings (1st-degree relatives), but also 2nd- and 3rd-degree relatives (close kin), can be used to estimate the dispersal distance over one generation of successful breeding (i.e., effective dispersal distance) in insects like *Ae. aegypti*.

Our newly developed method decomposes the observed separation distances between close kin (sampled as breeding adults) to generate the distribution of potential effective dispersal distances and to parametrize the dispersal distance kernel. This dispersal kernel provides the density of probability that a dispersal event ends at a given distance away from the source, regardless of the direction.
Importantly, it refers to the dispersal distance achieved over one generation of successful reproduction, such as distance between the birthplace and the ovipositing place of a female (i.e., distance between the ovipositing place of a mother and a daughter). We demonstrate the robustness of the method to produce dispersal kernel parameters consistent among different subsets of data (with one or multiple kinship categories) and congruent with the estimates of dispersal characteristics from the previously published MRR experiments in *Ae. aegypti*.

When few close kin are captured, the conventional genetic analysis of isolation-by-distance (IBD) between unrelated individuals can be used to estimate the spread of the dispersal kernel from the slope of the IBD relationship and the effective density, and we show that its results can be congruent with the new close-kin method. Finally, we show that the estimated genetic patch size (area of high local dispersal and breeding) from the spatial autocorrelation analysis reflects the spatial extent of the effective dispersal distance kernel’s "tail" that cannot be ascertained with IBD analysis alone.

We analyzed the genotyped and geo-referenced *Aedes aegypti* individuals collected in two densely populated areas of Singapore with a homogeneous distribution of high-rise apartment buildings. *Aedes aegypti* is the primary vector of dengue virus in Singapore that, despite having a low *Aedes* house index (2%) and an extensive vector surveillance and control program [22], continues to experience regular dengue outbreaks [23, 24]. This dataset offered an opportunity to characterize *Ae. aegypti* dispersal in a highly urbanized landscape with a prominent vertical dimension, but the spatio-genetic analytical approach (the new close-kin method, the conventional IBD and spatial autocorrelation analyses) can be applied across different landscapes and vector species.

**Results**

**Entomological sampling**

Adult *Aedes aegypti* females were collected using the simple sticky traps (Gravitraps) [25] that were positioned for vertical sampling in high-rise apartment buildings (ground level, 4th–5th floors, ≥ 9th floor) in two public areas of Singapore: Tampines and Yishun. Horizontal violin plots show the density distribution of separation distance for full siblings, 2nd-degree relatives, 3rd-degree relatives, all close kin combined ("all CK"), non-close kin ("all non CK"), null distribution ("non CK random"), and traps. The box within each violin plot shows the interquartile range and the location of the median.
Spatial distribution of close-kin pairs

Reduced representation genome sequences for individual *Ae. aegypti* females were obtained using the double-digest RAD sequencing [26] following the library preparation protocol by Rašić et al. [27]. Genotype likelihoods at 24–60 k genome-wide SNP positions (median = 57,947) were used to estimate the relationship between all pairs of individuals, using the combinations of three statistics (R0, R1, and KING-robust) [28, 29] to determine the kinship category as parent-offspring, full siblings, 2nd-degree relatives (half siblings, avuncular, grandparent-grandoffspring), 3rd-degree relatives (first cousins), and non-close kin.

In the total dataset, we detected 76 close-kin pairs: 19 full-sibling pairs, 18 pairs of 2nd-degree relatives, and 39 pairs of 3rd-degree relatives. We did not detect any parent-offspring (po) pairs, and 26.3% (20/76) of close-kin pairs were found on the same floor or 4–5 floors apart (19/20 pairs). Nearly half (47%) of all detected full-sibling pairs were caught within a building, in comparison to 27.8% of all 2nd-degree relatives and 15.4% of all 3rd-degree relatives. The members of each close-kin pair were sampled within 8 days (median = 1 day apart).

The median pairwise distance between close kin (CK) was 111.6 m (90th percentile = 264 m), with the maximum distance of 531 m, giving a positively skewed distribution of distances (skewness = 1.24, kurtosis = 1.90; Fig. 1c “all CK”). To compare this distribution to a null distribution for randomly spaced individuals across the matrix of traps, we created 100 random subsamples (with replacement) of recorded pairwise distances between individuals not assigned to any kinship category (non-close kin). The null distribution (Fig. 1c “non CK random”) had a median of 263.6 m and was significantly different from the distribution of pairwise distances between close-kin (the permutation test of equality of two distribution densities \( p < 0.001 \)).

Dispersal kernel parametrization from close-kin data

A distance separating close kin is a result of dispersal and successful breeding over multiple generations, and here we show how it can be used to infer a dispersal distance achieved over one generation of successful breeding. This distance is also known as the “effective dispersal distance” [32] and can be defined as a distance between the birthplace and the ovipositing place of a female (or a distance between the ovipositing place of a mother and a daughter). Every close-kin category contains information about the number of possible dispersal and breeding events over one generation that can be represented as a set. This set of numbers is used to divide the observed spatial distance between each close-kin pair to create a set of possible effective dispersal distances. By combining the values from all pairwise sets into one dataset, we can characterize the resulting distribution of possible effective dispersal distances. This is done by finding the “best-fitting” function (e.g., exponential, Weibull, log-normal) to generate a probability density function (pdf) of effective dispersal distance, i.e., effective dispersal distance kernel. A detailed description of the method is found in the “Methods” section.

The “goodness of the fit” criteria (AIC, BIC, the Kolmogorov-Smirnov statistic [33], Table 1) and the Q-Q plot (Additional file 2: Fig. S3) indicated that the distribution of possible effective dispersal distance derived from the close-kin data in Tampines and Yishun is well described by the Weibull and negative exponential (Laplacian) distribution. Given that the estimated Weibull shape parameter \( k \) was close to 1 (median 1.11, 95% CI 1.01–1.22), the Weibull distribution can be reduced to an exponential distribution with the estimated rate parameter \( \lambda = 0.022 \) (95% CI 0.020–0.025) (Table 1). This rate parameter for the combined dataset (“all CK”) was highly congruent.
with the estimates from separate CK categories: full siblings $\lambda = 0.019$ (95% CI 0.014–0.027), 2nd-degree relatives $\lambda = 0.021$ (95% CI 0.016–0.027), 3rd-degree relatives $\lambda = 0.024$ (95% CI 0.020–0.028) (Fig. 2). Both the mean and standard deviation ($\sigma$) of the exponentially distributed effective dispersal distance are $1/\lambda = 45.2$ m (95% CI 39.7–51.3 m), and the inferred dispersal kernel gives the 95% probability of effective dispersal distance up to 136 m (95% CI 120–152 m).

We repeated the same analytical procedure for the null distribution (random sampling of pairwise distances between non-close kin). The null distribution produced a distance kernel with the Weibull shape $k$ of 1.35 (95% CI 1.34–1.36) and Weibull scale $\lambda$ of 111.63 (95% CI 110.50–112.70) (Fig. 2) and was significantly different from the close-kin-derived kernel (permutation test for equality of densities $p < 0.001$).

### Dispersal kernel spread from the isolation-by-distance (IBD) analysis

When the probability of dispersal declines with distance, a positive correlation between genetic and geographic distances between individuals is expected, and this relationship is known as isolation-by-distance (IBD) [34, 35]. IBD is best illustrated by the regression of pairwise genetic distances onto geographic distances between individuals [34]. Applying the IBD theory, we estimated the standard deviation of the dispersal kernel ($\sigma$), also known as the dispersal kernel spread, using the slope of the IBD relationship and the effective population density [36]. Three different genetic distance measures (PCA-based [37], Rousset’s $\hat{\alpha}$ [36], Loiselle’s kinship [38]) gave highly congruent results in the IBD analysis (Additional file 1: Table S5), and we focus on the results with the PCA genetic distance in the main text.

### Table 1 Goodness-of-fit statistic and criteria and parameter estimates from the distribution fitting analysis for close-kin data

| Goodness-of-fit test | Negaitive exponential (Laplacian) | Weibull | Log-normal |
|----------------------|-----------------------------------|---------|------------|
| KS                   | 0.084                             | 0.057   | 0.125      |
| AIC                  | 2387.913                          | 2386.055| 2429.829   |
| BIC                  | 2391.427                          | 2393.082| 2436.856   |
| Parameter            | $\lambda$                         | Shape   | Meanlog    |
| Median (95% CI)      | 0.022 (0.020–0.025)               | 1.108 (1.011–1.222) | 3.322 (3.170–3.462) |
| Scale                |                                   | sdlog   |            |
|                      | 46.939 (41.770–52.445)            | 1.162 (1.055–1.252) |            |

A lower value of the statistic (the Kolmogorov-Smirnov statistic, KS) or the criterion (AIC, BIC) indicates a better fit. The parameter median and 95% CI were generated with 1000 bootstraps.

![Effective dispersal distance kernel](image)
Significant correlation between PCA genetic distance and linear geographic distance (IBD) was detected between non-close kin in Tampines (Mantel r = 0.124, 95% CI 0.052–0.198; $R^2 = 0.015$, $p = 3.91 \times 10^{-8}$) and Yishun (Mantel r = 0.158, 95% CI 0.112–0.208; $R^2 = 0.024$, $p = 2.18 \times 10^{-21}$) (Fig. 3). Because genetic distance between individuals can also be driven by the time of sampling (temporal distance), we confirmed that correlation between genetic and geographic distance (IBD) remained significant after the effect of temporal distance was removed (Partial Mantel test $p \leq 0.01$ for Tampines and Yishun; Additional file 1: Table S6). Absence of temporal structuring in our data was not surprising, given that we trapped ≥80% of analyzed individuals within 1 week (>93% within 2 weeks) in each site.

The estimated IBD slope $b$ was 0.0037 (95% CI 0.0024–0.0050) for the Tampines data, and 0.0065 (95% CI 0.0051–0.0079) for the Yishun dataset (Table 2). The estimated effective population density $D$ varied from 0.0014 to 0.0074 for Tampines and from 0.0014 to

![Fig. 3 Isolation-by-distance analysis on non-close kin data from Tampines and Yishun. Mantel test and linear regression were applied to the matrices of PCA genetic distance and linear geographic distance between pairs of individuals in Tampines (upper) and Yishun (lower), with close kin removed from both datasets. The red line shows regression with 95% CI (dashed lines)](image)
Aedes aegypti data from Tampines and Yishun

Our newly developed method allows for the parametrization of the effective dispersal distance kernel, as it decomposes the observed distances between close kin to generate the distribution of potential effective dispersal distances (achieved over one generation of successful reproduction). It gives probabilities of dispersal distances in any direction, referred to as the “dispersal distance kernel” \( k_d(r) \) [44], and it should not be confused with the “dispersal location kernel” \( k_l(r) \) that gives probabilities for the end locations of dispersers relative to the source locations [44]. The location kernel can be derived from the distance kernel and vice versa, given their relation: \( k_d(r) = 2\pi r k_l(r) \) in a two-dimensional habitat [44], and \( k_d(r) = 4\pi r^2 k_l(r) \) in a three-dimensional habitat.

Our method has some obvious parallels to the work of Jasper et al. [20] that also utilizes separation distance between close kin to infer dispersal distance over one generation, and the protocol by Rašić et al. [27] to generate genome-wide SNP data in *Ae. aegypti* collected from high-rise buildings. However, there are weaknesses in the method by Jasper et al. [20] that are not present in our method. First, their underlying assumption is Gaussian dispersal (normal distribution of dispersal distance [35]), even though the observed dispersal kernel in mosquitoes and other insects tends to be more leptokurtic.
with a higher probability of short- and long-distance dispersal \cite{7, 44, 45}. Our method allows the inference of a best-fitting distribution (e.g., negative exponential, Weibull, log-normal, etc., Fig. 2) when parameterizing the dispersal kernel. Many modeling studies have shown that processes such as population spread behave differently when long-tailed dispersal distributions are used instead of Gaussian \cite{46}. For example, an incorrect assumption of Gaussian dispersal can have operational impact on a mosquito control campaign like the Wolbachia-based replacement, because it would cause a perception that Wolbachia spreads through a population more slowly than expected (by as much as 40%, depending on the true shape of non-Gaussian dispersal kernel) \cite{47}. Second, the Jasper et al. method \cite{20} produces very wide confidence intervals for the estimated dispersal kernel spread (95% CI for $\sigma = 23$–93 m), while our method achieves much greater precision (95% CI for $\sigma = 40$–51 m). Third, Jasper et al. \cite{20} produce a non-trivial number of cases where $\sigma$ has a value of an imaginary number (square root of a negative number), and this raises concerns about the fundamental properties of their method.

Using the close-kin data from the Tampines and Yishun districts in Singapore, our close-kin method produced the negative exponential (Laplacian) kernel that gives 50% probability (kernel median) of effective dispersal occurring within 32 m. This indicates that, in this landscape, we can expect half of the successfully breeding individuals to stay within the high-rise building where they emerged or move to the adjacent building. The mean effective dispersal distance (and dispersal kernel spread $\sigma$) was estimated at 45.2 m (95% CI 39.7–51.3 m), with a 10% probability of a dispersal distance greater than 100 m (95% CI 92–117 m). Our genetic-based estimates match the parametrized dispersal kernel from mark-release-recapture (MRR) experiments performed in Brazil and Malaysia with \textit{Ae. aegypti} males from a genetically engineered line OX513A \cite{7}. Namely, their dispersal kernel gave estimates of a high level of dispersal to up 33 m, a mean distance traveled of 52.8 m (95% CI 49.9–56.8 m) in Brazil and 58 m (95% CI 52.1–71 m) in Malaysia, with 10% of dispersers moving $>105.7$ m (95% CI 86.5–141.1 m) \cite{7}. Moreover, globally collated MRR experimental data for \textit{Ae. aegypti} \cite{14} produced an exponential kernel with $\sigma = 54.1$ m \cite{10}. Given that we used a trapping system that preferentially catches ovipositing females, our close-kin-based kernel is more influenced by female dispersal. However, a high congruence with the MMR-based kernel for males from similar landscapes suggests similar dispersal in both sexes, as well as the robustness of spatial genetic patterns in reflecting the dispersal characteristics of this mosquito.

The IBD pattern reflects effective dispersal of both sexes averaged out over many generations, and our $\sigma$ estimates from the IBD analysis indicate that they can match the short-term measures derived from the close-kin analysis. Given that the close-kin method requires more intensive sampling in order to capture enough close-kin pairs for the reliable kernel parametrization, the use of IBD analysis as an alternative is appealing, particularly under budgetary limitations for genome-wide genotyping. However, IBD-based estimation of $\sigma$ requires accurate estimation of effective population size,
which it is not easily achievable \[48, 49\], and the uncertainty about this parameter has more impact than the uncertainty in the IBD slope \[50\]. The highest congruence between the IBD and the close-kin analysis was obtained using the $Ne_{est}$ estimates from the PWoP method \[39\] (method 1, Table 2), which gave $\sigma$ of 54.1 m (95% CI 40.8–67.7 m) for Tampines and 44 m (95% CI 37.6–49.7 m) for Yishun. The second best match was obtained using the entomological effective density estimate (method 3) that gave $\sigma$ of 66.8 m (95% CI 48.8–105.6 m) for Tampines and 74.4 m (95% CI 57.9–104.9 m) for Yishun. This is interesting, as it suggests that the gravitrapp data could be used as an entomological proxy for the effective population size, complementing the genetic-based estimation of this population parameter.

Fig. 5 Spatial genetic autocorrelation in Tampines and Yishun. The ending point of a distance class is on the x-axis, and spatial autocorrelation coefficient ($r$) of genotypes in Tampines (107 individuals) and Yishun (108 individuals) is on the y-axis. Two dashed lines along the x-axis are the permuted 95% CI of autocorrelations under the null hypothesis of a random distribution of genotypes in space. Vertical lines are the bootstrapped 95% CIs with the mean genetic autocorrelation.

The linkage disequilibrium-based $Ne_{est}$ estimate \[40\] (method 2) produced the widest range of $\sigma$ values (Tampines 54.7–206.2 m, Yishun 73–120.6 m) (Fig. 4). This LD-based estimator is expected to be less precise than the PWoP estimator \[39\] and also downward biased when IBD is present in a continuously distributed population, because it reflects the local genetic neighborhood size ($Nb$, \[35\]) rather than the effective size of the global population ($Ne$) \[51\]. We observed this trend in Ae. aegypti from Tampines and Yishun (with LD $Ne_{est}$ being similar to $Nb$; Additional file 1: Table S5). In addition to the limitations related to $Ne_{est}$ estimation, the IBD method assumes long-term stability of mosquito dispersal patterns and abundance, making it a meaningful alternative to the close-kin method in populations that do not
experience strong seasonal fluctuations, landscape alterations, intensive control campaigns, etc.

In Tampines and Yishun, that are largely homogeneous landscapes with multi-storey apartment buildings, 47% of all detected full siblings were found on the same floor or 4–5 floors apart, and the inferred dispersal kernel predicts high level of effective dispersal within a building or between adjacent buildings. This agrees with a previous study in Singapore where Ae. aegypti females marked with rubidium via spiked blood meal were released from middle floors and moved readily towards the top or bottom of multi-storey buildings in search of oviposition sites [52]. In Trinidad, significantly more eggs were collected in ovitraps 13–24 m above ground level than at any other elevation [53], again suggesting that vertical movement is common.

Given that high-rise apartment blocks can provide an abundance of hosts, oviposition, and resting sites, the tendency of Ae. aegypti to remain close to the birthplace is not unexpected. The tail of the dispersal kernel, however, provides insight into rarer long-range dispersal events that are consequential for the control strategies that rely on the releases of modified mosquitoes for population suppression or replacement. For example, the Wolbachia spread is expected to be slower in Ae. aegypti populations with longer-tail dispersal kernels, but this dispersal pattern also allows the initiation of the Wolbachia spread with smaller local releases [47]. In a population with a dispersal kernel like in Tampines and Yishun, theoretical approximation [47] predicts that the spread of Wolbachia (with a fitness cost equivalent to wMel strain) could be initiated if Wolbachia-infected Ae. aegypti are released in an area with a radius of at least 100–130 m (2.51σ for Laplacian kernel [47]). The diameter of this release area (200–260 m) matches the end of the effective dispersal kernel’s “tail” (99th percentile = 206 m, 95% CI 184–234 m). This diameter could also be roughly estimated from the spatial autocorrelation analysis, given that it corresponds to the distance at which spatial correlation stops being significantly positive (x-intercept = 200–300 m).

Accurately estimated dispersal range is also critical when determining the size of the area targeted by sterile or incompatible mosquito male releases for population elimination (SIT or IIT programs). In Fresno County, California, a high suppression (~95%) but not local elimination of Ae. aegypti was achieved with very high release numbers of Wolbachia-infected males [54]. The inability to achieve complete local elimination was probably the result of immigration of inseminated females from the nearby untreated areas [54]. In this landscape, released males were recaptured in large numbers up to 200 m from the nearest release point, and females were increasingly abundant closer they were to the untreated area (particularly up to 200 m from the treatment edge) [54]. Therefore, expanding the treatment area to account for a buffer of more than 200 m around the core area would likely be needed to permit complete elimination in the core. In Tampines and Yishun, spatial genetic analyses (tail of the dispersal kernel, x-intercept) suggest this buffer zone to be ~250 m.

The theoretical approximation of the conditions for Wolbachia-based mosquito replacement [47] and simulation modeling of this and other mosquito control strategies [10, 55–58] have all been developed assuming isotropic dispersal (invariable based on direction) in a 2-dimensional landscape. The approximation of the dispersal kernel for high-rise uniform landscapes could be achieved by considering the releases of mosquitoes from multiple floors rather than from the ground level only. Clearly, further theoretical and simulation modeling development that incorporates mosquito dispersal that is anisotropic (variable based on direction) in a 3-dimensional habitat is needed to precisely predict the requirements and outcomes of different mosquito control strategies in complex landscapes with a prominent vertical dimension.

Conclusions

Accurate and precise characterization of dispersal in the field is critical for the optimization of surveillance and control of disease vectors like Ae. aegypti. Knowledge of the dispersal kernel parameters enables operational teams to design and implement optimal surveillance, control, and release strategies in a given landscape, facilitating the efficient utilization of resources and maximizing the impacts of interventions. We show that spatial genetic analyses can provide robust estimates of mosquito dispersal patterns. Our newly developed method for the construction of the effective dispersal distance kernel through close-kin analysis enables the most comprehensive estimation of relevant parameters. The indirect inference from the IBD framework that requires less intensive sampling than close-kin analysis can also provide estimates of dispersal kernel spread; however, this approach is sensitive to the inaccurate estimates of effective population size and is uninformative about the probabilities of long-range dispersal that have important implications for control programs. Spatial autocorrelation analysis can complement the IBD analysis to ascertain the spatial extent of the effective dispersal kernel tail through estimation of the genetic patch size. With the decreasing cost of next-generation sequencing, acquisition of spatial genetic data is increasingly accessible. Given the complexities and criticisms of conventional MRR methods, and the central role of dispersal measures in essential vector
control programs, we recommend genetic-based dispersal characterization as the more desirable means of parameterization.

**Methods**

**Field collections**

Adult *Aedes aegypti* females were collected using the sticky traps developed by the Environmental Health Institute, National Environment Agency of Singapore (EHI, NEA), known as the Gravitraps—simple, hay infusion-filled cylindrical traps with a sticky inner surface that preferentially catch gravid females in search of suitable ovipositing sites [25]. The Gravitraps have been deployed in public housing estates island-wide since 2017 as part of the vector surveillance program. For the current study, we chose two public housing estates: Tampines (30 acre sampling area in patch 1, 10 acres in patch 2) and Yishun (46 acre sampling area) that are 14 km apart (Fig. 1a, b). Each unit in the high-rise apartment blocks has open windows and entrance from an open-air corridor suitable for mosquito movement. There are no heating or cooling ducts between apartment blocks and floors that could serve as movement corridors for mosquitoes. The Gravitraps were positioned for vertical sampling in each block: ground level (1st–2nd floors), mid-level (4th–5th floors), and high level (9th floor and above). In both sites, the median number of traps per block per week was 6 (mean = 6.4–6.7). Sticky linings from each Gravitrap were collected by the entomological surveillance team once a week. All adult mosquitoes were identified to species/genus using the taxonomic keys. Mosquitoes identified as *Ae. aegypti* were transferred to 2-mL tubes according to their collection date, residential block, and floor and were stored in 100% ethanol at 4 °C until further processing.

**DNA extraction, sequencing, and genotyping of individual mosquitoes**

Total genomic DNA was extracted from individual mosquitoes using Dneasy Blood and Tissue DNA extraction kit (Qiagen, Hilden, Germany) according to the manufacturer’s instructions. Individual mosquitoes were homogenized manually in 180 μl of ATL buffer using sterile plastic pestles, and the proteinase digestion (with 20 μl of proteinase K) was carried out overnight. The DNA was quantified by using the Qubit High Sensitivity DNA kit (Thermo Fisher Scientific, Waltham, MA, USA), and only the specimens that yielded a DNA concentration of at least 4 ng/μl were included in RADseq libraries. Reduced-genome representation sequence data were generated for each individual using the double-digest RAD sequencing approach by Peterson et al. [26], with the sample processing and library preparation protocol as described in Rašić et al. [27]. ddRAD-seq libraries were sequenced on the Illumina HiSeq4000 platform. The sequencing data were demultiplexed [59] and processed (trimmed to 90 bp and filtered for quality) using the bash script/pipeline from Rašić et al. [27]. The percentages of raw reads passing the quality filtering threshold (phred score ≥ 20) were high for all individuals (96.52–98.29%), suggesting no substantial DNA degradation [60]. The high-quality reads were aligned to the AaegL5 version of the *Ae. aegypti* genome assembly [61] using Bowtie [62]. Unambiguously mapped high-quality reads were converted to bam format and processed in SAMtools [63] to generate sorted bam files that were used to produce genotype likelihood and VCF files using the SAMtools variant calling method as implemented in ANGSD [64]. The final dataset included 107 mosquitoes from Tampines and 108 from Yishun that had < 30% missing data at 83,255 and 69,051 variable sites (SNPs) for the Tampines and Yishun datasets, respectively.

**Inference of familial relationship (kinship estimation)**

Relationships between individuals were determined using the recently developed approach by Waples et al. [28] as implemented in the program NGSRelate [29]. This method shows improved accuracy and precision when compared to related approaches, given that (a) it does not require population allele frequency estimates; instead, the framework calculates two-dimensional site-frequency-spectrum for each pair of individuals, and (b) it is applied directly to sequencing data (via genotype likelihoods) rather than the called genotypes, which is particularly suitable for lower-depth sequencing data acquired in RAD-seq experiments [28]. Like most other relatedness inference methods, it can be biased upward or downward for different violations of underlying assumptions (inbreeding, population structure or admixture) [28]. Analogous to the KING-robust method [65], this method is expected to generally be negatively biased for pairs of individuals that have different ancestries, but it is also fairly robust in separating close kin with similar ancestry from unrelated individuals under population structure [30, 31].

For the spatial analyses, we considered close kin as pairs with an inferred category of 1st-, 2nd-, or 3rd-degree relatives. Kinship categories were determined based on the combination of three statistics (R0, R1, and KING-robust kinship), which allows the distinction between the parent-offspring and the full-sibling relationship within the category of 1st-degree relatives [28]. Second-degree relatives include half-siblings, avuncular, and grandparent-grandoffspring pairs that cannot be genetically distinguished, but the grandparent-grandoffspring category is the least likely in our sampling scheme (collection of gravid females, ≥ 80% collected in 1 week and > 93% collected during 2 weeks in...
each site). Also, we can assume that half-siblings are paternally (i.e., half-sisters share a father, not a mother) given that *Ae. aegypti* females rarely mate more than once [66, 67]. We assume that 3rd-degree relatives are first cousins, given that a category like great-grandparent/great-grandoffspring is unlikely in our sampling scheme.

**Genetic and geographic distance among individuals**

We used different individual-based genetic distances among individuals within each area (Yishun or Tampines). The PCA-based genetic distance was estimated by first performing the principal component analysis (PCA) from genotype data in the R package “adegenet” [68] and then creating a distance matrix from the Euclidean distance among the maximal number of PC axes. PCA genetic distance does not assume any particular microevolutionary processes, and it exhibits a linear relationship with Euclidean geographic distance, showing the highest model selection accuracy in landscape genetic studies, especially when dispersal rates are high across the examined area [37]. We also estimated Rousset’s genetic distance $\delta$ [36] and Loiselle’s kinship coefficient [38] in the program SPAgeDI [69].

Pairwise spatial (geographic) distance between mosquitoes was calculated as the shortest straight line (Euclidean) distance in three dimensions, based on the $X/Y$ (long/lat) and $Z$ (height) coordinates of their collection point, here called the Euclidean 3D distance, representing a linear geographic distance in meters (m). Natural logarithm (ln) of this distance was used in the analyses where Rousset’s $\delta$ or Loiselle’s kinship coefficient was applied (see below), given that both of these genetic coefficients exhibit approximate linear relationship with ln-geographic distance [69].

**Estimation of mosquito dispersal characteristics**

**Dispersal kernel estimation from close-kin data**

In our sampling scheme, adult females were caught after landing on a lethal ovipositing site (Gravitrap), and we assume that this is a result of their active flight (and not passive, human-assisted movement). We consider pairs of females caught in different traps (non-zero separation distance) that could be genetically assigned to one of the following kinship categories: parent-offspring (po), full siblings (fs), 2nd-degree relatives (2nd) (half siblings, hs; avuncular, av; grandparent-grandoffspring, go), 3rd-degree relatives (3rd) (first cousins, fc), and non-close kin. Every close-kin category contains information about the number of possible effective dispersal events. For example, a pair of full siblings (fs) could have originated from the same breeding site from which each sibling moved into a gravid trap (two events), or they could have originated from different breeding sites (three events, including mother’s skip oviposition). Therefore, the corresponding number of possible dispersal events, for each case, can be calculated as the number of such breeding sites ($n$) plus one ($n + 1$). Based on this, we constructed the sets with elements that represent the number of possible dispersal events for each case as $n_{\text{min}} + 1,.., n_{\text{max}} + 1$. For the fs category, this set is FS = {2, 3}. In the case of a parent-offspring (po) pair, the minimum and maximum number of breeding sites is $n_{\text{min}} = n_{\text{max}} = 1$, giving a set PO = {2}. For the kinship category of 2nd-degree relatives, we have the following subsets: half-siblings HS = {2, 3, 4, 5}, avuncular AV = {3, 4}, and grandparent-grandoffspring GO = {2, 3}. We constructed the full set for 2nd-degree relatives as the union of these three subsets (containing unique elements): 2ND = HS $\cup$ AV $\cup$ GO = {2, 3, 4, 5}. In the case of 3rd-degree relatives (first cousins fc), the minimum number of contributing breeding sites is $n_{\text{min}} = 1$ (first cousins and their mothers all originate from the same breeding site), while the maximum is $n_{\text{max}} = 4$ (first cousins and their mothers each originate from a unique breeding site). Therefore, for the category of 3rd-degree relatives, we can construct a set as 3RD = FC = {2, 3, 4, 5}.

We then created a set of possible effective dispersal distances for each detected close-kin pair based on their distance and assigned kinship category. This set of distances contains the same number of elements as the corresponding set of possible effective dispersal events (described above), and its values are obtained by dividing the detected spatial distance between a pair ($d$) with the corresponding set element. For example, if a collected pair AB, separated by distance $d_{AB}$, falls into the category of 3rd-degree relatives, then a set of possible effective dispersal distances for this pair would be $d_{3\text{rd},AB} = [d_{AB}/2, d_{AB}/3, d_{AB}/4, d_{AB}/5]$. For a full-sibling pair BC separated by spatial distance $d_{BC}$, the set of possible effective dispersal distances will be $d_{3\text{rd},BC} = [d_{BC}/2, d_{BC}/3]$.

By combining the values from all pairwise sets of possible effective dispersal distances into one dataset, we characterized the resulting distribution of possible effective dispersal distances. This dataset was used to generate a probability density function (pdf) of effective dispersal distance (i.e., effective dispersal distance kernel) by fitting different functions (exponential, Weibull, log-normal) using the R package “fitdistrplus” [33] that incorporates maximum likelihood estimation and parametric bootstrapping to generate median as well as 2.5 and 97.5 percentiles of each distribution parameter. To determine the “best fitting” of the tested distributions, we assessed the Q-Q plots and computed goodness-of-fit statistics with an approximate Kolmogorov-Smirnov test, Akaike Information Criterion (AIC), and Bayesian Information Criterion (BIC) [33].
To estimate a pdf for randomly spaced individuals across the sampled area (null distribution), we simulated 100 datasets where pairs had a randomly assigned kinship category and a distance randomly sampled from the recorded distances between non-close kin. The number of simulated pairs in each kinship category matched the number of observed pairs in the empirical dataset. We then applied the analytical procedure described above on all simulated datasets and compared the simulated (random) and empirical distributions in the R package “sm” [70] using the permutation test of equality of two distribution densities.

**Isolation-by-distance analysis (IBD) and estimation of the dispersal kernel spread**

IBD analysis can be used even when few or no close kin are captured. In fact, highly related individuals should be removed from the IBD analysis in order for it to reflect the long-term population processes [71], and we created a subsample for each area by removing individuals identified as close kin, leaving 63 and 85 individuals in the Tampines and Yishun subsample, respectively. The significance of IBD was tested separately in Tampines and Yishun using Mantel’s correlation test with 1000 permutations, as implemented in the R package “ecodist” [72].

IBD is best illustrated by the regression of pairwise genetic distances onto geographic distances among individuals [34]. The slope of this linear regression and the effective density can be used to estimate the standard deviation of the dispersal kernel (σ) that is also known as the dispersal kernel spread [50]. The dispersal kernel spread can be calculated as σ = √(1/4πDb), where b is the slope of the linear regression and D is the effective density of reproducing individuals.

The slope of the linear relationship was estimated using the lm() function in R (R Core Team) for three different sets of genetic and geographic matrices. The first set included a matrix of PCA-based genetic distances against the matrix of linear geographic distances. A matrix of Rousset’s d or Loiselle’s kinship coefficient was tested against the matrix of I-transformed geographic distances, given that both genetic estimators are expected to vary approximately linearly with the natural logarithm of the distance [37].

The effective density D is defined as Ne/study area, where Ne is the effective population size. We estimated Ne using two genetic methods based on a single sample. The first method is Ne estimation by Waples and Waples [39], based on the “parentage analysis without parents” (method 1, PWoP) that uses the frequency of full- and half-siblings in a population sample to reconstruct the number of parents that contributed to such a sample. The median and the 95% confidence interval were calculated using 100 resamples with a random replacement of one individual. The second method represents Ne estimation by Waples and England [40] that is based on the linkage disequilibrium data (method 2, LDN_e), with the 95% confidence interval calculated using the jacknifing procedure over loci implemented in the program Ne estimator v.2.1 [73]. Finally, effective density was estimated using the entomological survey data from the Gravitrap sentinel trap system across the study areas for the period from January 2018 through May 2018 (method 3, Gravitr)ap). During those months, the surveillance system contained a minimum of 1357 traps distributed over a 45 acre area in Tampines and a minimum of 1048 traps over a 65 acre area in Yishun. An average of 441 adult females (range 391–760) were caught per month in Tampines, and 280 (range 202–404) in Yishun (Additional file 1: Table S7). The average number of breeding females per square meter was multiplied by 2 to give the effective number of breeding individuals per unit area, as we assume 1:1 sex ratio in an Ae. aegypti population [74]. For Tampines, we considered patch 1 (larger sampling area) as a more representative population sample for the calculations of local Ne and density than the smaller patch 2 (Fig. 1a).

**Spatial autocorrelation analysis**

To compute the correlogram curve for each sampling site, we used PCA genetic distance among all genotyped mosquitoes in a site, and the spatial grouping within distance classes that were incrementally increased by 50 m. The analysis was done in GenAlEx v.6.501 [75]. The autocorrelation coefficient under the null hypothesis of no spatial structure was generated by the permutation procedure that shuffles all individuals among the geographic locations within a site 1000 times and generates 95% CI with the 25th and 975th ranked permuted values. 95% CI for the observed autocorrelation coefficient for each distance class was obtained from 1000 random draws of individuals with replacement.

**Supplementary information**

Supplementary information accompanies this paper at https://doi.org/10.1186/s12915-020-00841-0.

**Additional file 1: Table S1.** Tampines: the number of female Aedes aegypti caught per week per building. Table S2. Tampines: the number of recovered traps per building. Table S3. Yishun: the number of female Aedes aegypti caught per building. Table S4. Yishun: the number of recovered traps per building. Table S5. IBD-based estimates of the dispersal kernel spread (σ) in Aedes aegypti from Singapore. Table S6. Mantel and partial Mantel tests for non-close-kin data in Tampines and Yishun. Table S7. Entomological data used for the estimation of effective population density in Aedes aegypti (method 3).

**Additional file 2: Fig. S1.** Spatial network of close kin in Tampines (A) and Yishun (B). Fig. S2. Frequency histogram of separation distance for each kinship category. Fig. S3. Distribution fitting (dispersal kernel parametrization) analysis.
Acknowledgements

We would like to thank Dr. Chee Seng Chong and the colleagues from the Applied Entomology and Informatics Sections at EHI, for providing mosquito specimens and entomological data; and, Mr. Chew Ming Fai, Director-General for Public Health, National Environment Agency (NEA) and Associate Prof. Lee Ching Ng, Director, EHI, for their support and approval to publish the study. We thank Dr. Paola Florez de Sessions, Dr. Yuan Zhu, Dr. Chiea Chuen Khor, Dr. Chih Chuan Shih, and colleagues at Genome Institute Singapore, A*Star, Singapore, for their assistance in accessing the RADseq libraries. We also thank Sean L. Wu for insightful discussions on statistical analyses and two anonymous reviewers for their helpful suggestions and comments.

Authors’ contributions

G.R., H.C.H., and I.F. conceived the study; I.F. and G.R. developed the methodology and analyzed the spatio-genomics data; W.P.T. prepared the RADseq libraries; MABAR, C.L., C.H.T., and H.C.H. performed the field work and analyzed the entomological data; H.C.H. and I.F. curated the data; H.C.H. performed the project administration; G.R. and I.F. wrote the manuscript; G.J.D. and H.C.H. edited the manuscript. All authors read and approved the final manuscript.

Funding

The study received financial support from the Ministry of Finance, Singapore, through the Climate Resilience Studies Fund, and partial support from a commercial contract on insecticide resistance awarded to QIMRB by the by the Australian Department of Health. This research was supported by use of the Nectar Research Cloud, a collaborative Australian research platform supported by the National Collaborative Research Infrastructure Strategy (NCRIS). The funding bodies had no role in the design of the study and collection, analysis, and interpretation of data, and in writing the manuscript.

Availability of data and materials

The datasets generated and analyzed during the current study are available in the NCBI Sequence Read Archive (SRA) repository under the accession number PRJNA6393937, https://www.ncbi.nlm.nih.gov/bioproject/PRJNA6393937 [59], and are included in this published article and its Additional files.

Ethics approval and consent to participate

Not applicable.

Competing interests

The authors declare that they have no competing interests.

Author details

1Mosquito Control Laboratory, QIMR Berghofer Medical Research Institute, 300 Herston Road, Herston, QLD 4006, Australia. 2Environmental Health Institute, National Environment Agency, 11, Biopolis Way, #06-05-08, Singapore 138667, Singapore.

Received: 7 April 2020 Accepted: 5 August 2020

Published online: 20 August 2020

References

1. World Health Assembly. World Health Organization. 2018. http://www.who.int/vector-control/publications/global-control-response/en/. Accessed 10 Feb 2020.

2. Achee NL, Gould F, Perkins TA, Reiner RC Jr, Morrison AC, Ritchie SA, et al. A critical assessment of vector control for dengue prevention. PLoS Negl Trop Dis. 2015;9(5):e0003655.

3. Hoffmann AA, Montgomery BL, Popovic J, Iturbe-Ormaetxe I, Johnson PH, Muzi F, et al. Successful establishment of Wolbachia in Aedes populations to suppress dengue transmission. Nature. 2011;476(7361):454–7.

4. Kittayapong P, Kaengthong N-Q, Ninphanomchai S, Limophasmanee W. Combined sterile insect technique and incompatible insect technique: sex separation and quality of sterile Aedes aegypti male mosquitoes released in a pilot population suppression trial in Thailand. Parasit Vectors. 2018;11:657.

5. Lacroix R, McKemey AR, Raduan N, Kwee Wee L, Hong Ming W, Guat Ney T, et al. Open field release of genetically engineered sterile male Aedes aegypti in Malaysia. PLoS One. 2012;7(8):e42771.

6. Vázquez-Prokopec GM, Montgomery BL, Horne P, Clennon JA, Ritchie SA. Contacting combination with targeted indoor residual spraying significantly reduces dengue transmission. Sci Adv. 2017;3(2):e1602024.

7. Winskill P, Carvalho DO, Capurro ML, Alphey L, Donnelly CA, McKemey AR. Dispersal of engineered male Aedes aegypti mosquitoes. PLoS Negl Trop Dis. 2015;9(11):e0004156.

8. Schmidt TL, Barton NH, Račić G, Turley AF, Montgomery BL, Iturbe-Ormaetxe I, et al. Local introduction and heterogeneous spatial spread of dengue-suppressing Wolbachia through an urban population of Aedes aegypti. PLoS Biol. 2017;15(5):e2001894.

9. Marshall JM, Akbari OS. Gene drive strategies for population replacement. In: Adelman ZN, editor. Genetic control of malaria and dengue. Cambridge: Academic Press; 2016. p. 169–200.

10. Sánchez CHM, Bennett JB, Wu SL, Račić G, Akbari OS, Marshall JM. Modelling confinement and reversibility of threshold-dependent gene drive systems in spatially-explicit Aedes aegypti populations. BMC Biol. 2020;18(1):150.

11. Marshall JM, Akbari OS. Can CRISPR-based gene drive be confined in the wild? A question for molecular and population biology. ACS Chem Biol. 2018;13(2):424–30.

12. Stone CM, Schwab SR, Fonseca DM, Fefferman NH. Contrasting the value of targeted versus area-wide mosquito control scenarios to limit arbovirus transmission with human mobility patterns based on different tropical urban population centers. PLoS Negl Trop Dis. 2019;13(7):e0007479.

13. Roiz D, Wilson AL, Scott TW, Fonseca DM, Jourdain F, Müller P, et al. Integrated Aedes management for the control of Aedes-borne diseases. PLoS Negl Trop Dis. 2018;12(12):e0006845.

14. Guerra CA, Reiner RC Jr, Perkins TA, Lindsay SW, Midega JT, Brady OJ, et al. A global assembly of adult female mosquito mark-release-recapture data to inform the control of mosquito-borne pathogens. Parasit Vectors. 2014;7:276.

15. Trewin BJ, Pagendarm DE, Zalucki MP, Darbo JM, Devine GJ, Jansen CC, et al. Urban landscape features influence the movement and distribution of the Australian container-inhabiting mosquito vectors Aedes aegypti (Diptera: Culicidae) and Aedes notoscriptus (Diptera: Culicidae). J Med Entomol. 2019;56(2):443–53.

16. Marcondes M, Reyes T, Bander CM. Quantifying Aedes aegypti dispersal in space and time: a modeling approach. Ecosphere. 2019;10(ec2):2977.

17. Dickens BL, Brant HL. Effects of marking methods and fluorescent dyes on Aedes aegypti survival. Parasit Vectors. 2014;7:65.

18. Lepais O, Davill B, O’Connor S, Osborne JL, Sanderson RA, Cussans J, et al. Estimation of bumblebee queen dispersal distances using sibship reconstruction method. Mol Ecol. 2010;19(4):819–31.

19. Odero JO, Fillinger U, Rippon EJ, Nisiga DK, Weetman D. Using sibship reconstructions to understand the relationship between larval habitat productivity and oviposition behaviour in Kenyan Anopheles arabiensis. Malar J. 2019;18:286.

20. Jasper M, Schmidt TL, Ahmad NW, Sinkins SP, Hoffmann AA. A genomic approach to informing kinship reveals limited intergenerational dispersal in the yellow fever mosquito. Mol Ecol Resour. 2019;19(5):1254–64.

21. Colton YM, Chadee DD, Severson DW. Natural skip oviposition of the mosquito Aedes aegypti is codetermined by multiple genetic factors. Med Vet Entomol. 2003;17(2):195–204.

22. Ong J, Liu X, Rajarethinam J, Yap G, Ho D, Ng LC. A novel entomological index, Aedes aegypti Breeding Percentage, reveals the geographical spread of the dengue vector in Singapore and serves as a spatial risk indicator for dengue. Parasit Vectors. 2019;12:17.

23. Hapuarachchi HC, Koo C, Rajarethinam J, Chong C-S, Lin C, Yap G, et al. Epidemic resurgence of dengue fever in Singapore in 2013–2014: a virological and entomological perspective. BMC Infect Dis. 2016;16:300.

24. Rajarethinam J, Ang L-W, Ong J, Yasas C, Hapuarachchi HC, Yap G, et al. Dengue in Singapore from 2004 to 2016: cyclical epidemic patterns dominated by serotypes 1 and 2. Am J Trop Med Hyg. 2018;99(1):204–10.

25. Lee C, Vythilingam J, Chong C-S, Abdul Razak MA, Tan C-H, Liew C, et al. Gravitraps for management of dengue clusters in Singapore. Am J Trop Med Hyg. 2013;88(5):888–92.

26. 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.

27. Rajić 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;15:275.

28. Waples RK, Albrechtsen A, Moltke I. Allele frequency-free inference of close familial relationships from genotypes or low-depth sequencing data. Mol Ecol. 2019;28(1):35–48.
