RESOURCE ARTICLE

A genomic approach to inferring kinship reveals limited intergenerational dispersal in the yellow fever mosquito

Moshe Jasper1 | Thomas L. Schmidt1 | Nazni W. Ahmad2 | Steven P. Sinkins3 | Ary A. Hoffmann1

1School of BioSciences, Bio21 Institute, University of Melbourne, Parkville, Victoria, Australia
2Institute for Medical Research, Ministry of Health Malaysia, Kuala Lumpur, Malaysia
3Centre for Virus Research, University of Glasgow, Glasgow, UK

Correspondence
Thomas L. Schmidt, School of BioSciences, Bio21 Institute, University of Melbourne, Parkville, Vic, Australia.
Email: tom.schmidt@unimelb.edu.au

Funding information
National Health and Medical Research Council, Grant/Award Number: 1118640 and 1132412; Wellcome Trust, Grant/Award Number: 108508

Abstract
Understanding past dispersal and breeding events can provide insight into ecology and evolution and can help inform strategies for conservation and the control of pest species. However, parent–offspring dispersal can be difficult to investigate in rare species and in small pest species such as mosquitoes. Here, we develop a methodology for estimating parent–offspring dispersal from the spatial distribution of close kin, using pairwise kinship estimates derived from genome-wide single nucleotide polymorphisms (SNPs). SNPs were scored in 162 Aedes aegypti (yellow fever mosquito) collected from eight close-set, high-rise apartment buildings in an area of Malaysia with high dengue incidence. We used the SNPs to reconstruct kinship groups across three orders of kinship. We transformed the geographical distances between all kin pairs within each kinship category into axial standard deviations of these distances, then decomposed these into components representing past dispersal events. From these components, we isolated the axial standard deviation of parent–offspring dispersal and estimated neighbourhood area (129 m), median parent–offspring dispersal distance (75 m) and oviposition dispersal radius within a gonotrophic cycle (36 m). We also analysed genetic structure using distance-based redundancy analysis and linear regression, finding isolation by distance both within and between buildings and estimating neighbourhood size at 268 individuals. These findings indicate the scale required to suppress local outbreaks of arboviral disease and to target releases of modified mosquitoes for mosquito and disease control. Our methodology is readily implementable for studies of other species, including pests and species of conservation significance.

KEYWORDS
Aedes aegypti, conservation, dispersal, genome-wide single nucleotide polymorphisms, kinship, neighbourhood size
INTRODUCTION

Dispersal is a key trait in ecology and evolution, allowing species to evade stressful areas and locate favourable new areas and determining levels of gene flow that influence the adaptability of species (Clobert, Baguette, Benton, & Bullock, 2012). Knowledge of dispersal can also be vital to applied research that seeks to determine the conservation status of threatened species or the biosecurity risk of pest species, or for managing outbreaks of pests and disease vectors once invasions have taken place (Killeen, Knols, & Gu, 2003; Ouborg, Piquot, & Van Groenendael, 1999). Dispersal has traditionally been investigated by marking and tracking individuals across a landscape. However, these methods are often laborious and do not capture past dispersal patterns that produce gene flow. New methods and tools in landscape genetics can help understand past dispersal and gene flow patterns, including potential environmental parameters that limit dispersal (Schmidt, Filipović, Hoffmann, & Rašić, 2018; Watts et al., 2007). The advent of high-density sequencing technologies has also provided increased power for landscape genomic studies conducted at spatial scales fine enough to investigate dispersal discretely over generations, and to measure individual acts of movement (Schmidt, Rašić, et al., 2017).

Under Wright’s isolation-by-distance framework, the impact of dispersal on genetic variation in continuously distributed populations is described in terms of the neighbourhood size (NS), a product of parent–offspring dispersal variance and the ideal density of breeding adults within the dispersal area (Wright, 1946). Wright’s equation, \( NS = 4\pi \sigma^2 d \), represents a key bridge between local demographic processes and broader patterns of genetic differentiation. As regards the ecology and genomics of a species, this bridge can be crossed in both directions, either using genomic differentiation to understand processes of dispersal and density or alternatively using estimates of density and dispersal to gain an insight into likely processes of genomic differentiation.

The link between NS and dispersal has been investigated using composite methodologies incorporating both individual genetic distances and mark–release–recapture (MRR) (Sahlsten, Thorngren, & Hoglund, 2008; Watts et al., 2007). However, for many species it is difficult to estimate density or dispersal using MRR methods. For instance, the small size of mosquitoes necessitates the use of field-sampled or laboratory-reared mosquitoes that are released for recapture from a central point (Honorio et al., 2003; Reiter, Amador, Anderson, & Clark, 1995). If marking materials or sampling protocols impact the fitness of released individuals, or if rearing conditions do not adequately represent release conditions, this methodology risks overestimating or underestimating dispersal distances. For example, *Aedes aegypti* (yellow fever mosquito) reared in suboptimal larval conditions may settle for ovipositing in suboptimal sites more readily than those reared in favourable larval conditions (Kaur, Lai, & Giger, 2003). Thus, if rearing conditions are unfavourable compared to conditions at the release site, released mosquitoes may disperse less than mosquitoes from the wild population under investigation.

MRR techniques can also be expanded to incorporate information from genetic inferences of kinship (Bravington, Skaug, & Anderson, 2016). This framework of close-kin mark–recapture (CKMR) uses a genetic “mark” that extends beyond the individual by considering individuals to have genetically “marked” their close kin, such that any kin sampled constitute an extended “recapture” of the original individual. Kinship estimation can also be useful for investigating migration between populations, particularly in cases where migration rates have changed too recently to have produced corresponding changes in population genetic structure (Palsbøll, Zachariah, & Bérubé, 2010), such as that might be expected in invasive species. While studies using microsatellites can often only confidently identify first-order kin relations (parent-offspring or full-sibling), the
use of high-density, genome-wide molecular markers can enable reasonably accurate assignment of individuals to second-order (e.g. half-sibling) and third-order (e.g. first cousin) groupings (Phillips, Garcia-Magariños, Salas, Carracedo, & Lareu, 2012).

In this paper, we develop a methodology that uses genome-wide single nucleotide polymorphisms (SNPs) to estimate relatedness across three orders of kinship, and then uses the spatial distribution of these kin to estimate parent-offspring dispersal and related parameters. Our method inherits from CKMR, but exploits the fact that the spatial separation distance between each kin pair can be characterized as a set of distances representing past parental dispersal events (Figure 1), provided that the population is sampled at a specific point in time and from a set of sampling locations distributed continuously through space. By separating the composite life histories associated with each kinship category into discrete parent-offspring dispersal events, we derive estimates of parent-offspring dispersal distances. Also, as most populations sampled continuously through space should have more second-order than first-order relations, and more third-order than second-order relations, this approach provides a means of investigating dispersal using a greater number of data points than by using first-order relatives alone, as dispersal distance estimates from first-, second-, and third-order relatives can all be incorporated into a single analysis. By using thousands of SNPs to estimate kinship and by integrating three orders of kinship into a single analysis, our methodology can be conducted with smaller sample sizes than those often required for kinship studies, a requirement which may have prevented more widespread adoption of kinship-based methodologies (Palsbøll et al., 2010).

We use the above methodology to investigate dispersal in Ae. aegypti, the primary vector of arboviral diseases such as dengue, chikungunya and Zika (Morrison, Zelinski-Gutierrez, Scott, & Rosenberg, 2008). This highly anthropophilic mosquito is generally considered a weak disperser by flight (Harrington et al., 2005), though some MRR studies have reported flights over long distances (Honorio et al., 2003; Reiter et al., 1995). Given a local abundance of human hosts, the primary driver of dispersal in females is the search for suitable oviposition sites (Edman et al., 1998), and when undertaking “skip” oviposition, a female may oviposit at several sites during a single gonotrophic cycle (Reiter, 2007). Typically, females mate only once in their lives, while males may have multiple partners (Christophers, 1960). Despite its limited active dispersal, Ae. aegypti has managed to invade much of the global tropics over the past several centuries, dispersing passively along human trade and transport routes (Powell & Tabachnick, 2013).

We focus on dispersal across a residential site in Kuala Lumpur, Malaysia, as part of preparation for active interventions involving releases of Ae. aegypti transfected with the bacterium Wolbachia. These interventions aim to replace uninfected Ae. aegypti with Wolbachia-infected mosquitoes that have a reduced potential to transmit dengue (Hoffmann et al., 2011; O’Neill et al., 2018; Schmidt, Barton, et al., 2017). The long-term persistence and spread of Wolbachia in Ae. aegypti will strongly depend on local Ae. aegypti dispersal characteristics (Hoffmann et al., 2014; Schmidt, Barton, et al., 2017; Turelli & Barton, 2017). Understanding intergenerational movement will also help inform strategies for responding to newly detected Ae. aegypti incursions globally or to localized dengue outbreaks by chemical suppression of populations. However, applications of this methodology extend well beyond Ae. aegypti and other pest vectors. For instance, a key question in insecticide resistance management in agricultural pests concerns local movement of resistance alleles once resistance first arises, to help reduce resistance spread (Maino, Binns, & Umina, 2018). And in threatened species, understanding the spread of individuals across metapopulations can be vital in managing the species (Szczys, Oswald, & Arnold, 2017). Our approach to estimating dispersal and neighbourhood size therefore has numerous applications for both threatened and threatening species.

# Materials and Methods

## Mosquito sampling

*Aedes aegypti* were collected from Mentari Court in Petaling Jaya, Malaysia, between 19 September and 9 October 2017. Mentari Court consists of a set of 18-storey apartment blocks occupying a 500 × 250 m area, enclosing a vegetated area and a school (Figure 2). We deployed ovitraps on either the third or the fourth floor of each building, a sufficient height to avoid sampling *Ae. albopictus*, which are common at ground level in Mentari Court. We collected and stored *Ae. aegypti* larvae every week for 3 weeks. We divided the study region into nine sample sites across eight buildings (Figure 2) and sampled 18 mosquitoes from each site for DNA extraction, selecting six mosquitoes from different traps from each of the three weeks.

## DNA extraction and genotyping

DNA was extracted from the 162 *Ae. aegypti* using a Roche High Pure™ PCR Template Preparation Kit (Roche Molecular Systems, Inc., Pleasanton, CA, USA), with the addition of an RNase A digestion step. Extracted DNA was used to construct two RAD libraries containing 81 individuals each. We followed the double-digest restriction-site-associated DNA sequencing protocol for *Ae. aegypti* developed by Rašić, Filipović, Weeks, and Hoffmann (2014), but selected a smaller size range of DNA fragments (350–450 bp) to accommodate the larger number of individuals in each library. The libraries were sequenced on the Illumina HiSeq3000 platform at the Monash Health Translation Facility, obtaining 100 bp paired-end reads and using a 20% Phi-X spike-in to reduce the impact of low diversity RADtags.

Following sequencing, barcoded reads were filtered, truncated to 80 bp, and demultiplexed using the *process_radtags* program in *stacks* v2.0 (Catchen, Hohenlohe, Bassham, Amores, & Cresko, 2013). Paired and single-end reads were concatenated and aligned to the *Ae. aegypti* nuclear genome assembly *aaegl4* (Dudchenko et al., 2017) with Bowtie 2 (Langmead & Salzberg, 2012) using – very
sensitive alignment settings. We used the program ref_map to build a Stacks catalog and the program Populations to select SNP loci that were scored in at least 70% of mosquitoes. We filtered further with VCFtools (Danecek et al., 2011), retaining SNPs in Hardy–Weinberg equilibrium ($p = 1e10^{-5}$) and with minor allele frequencies $>0.05$, and thinned the remaining SNPs so that none was within 250 kbp of another. Thinning at this threshold in Ae. aegypti retains approximately 8 SNPs per map unit, a sampling density shown to largely eradicate linkage effects in SNPs (Cho & Dupuis, 2009). This set of 3,939 SNPs was used for all downstream analyses.

### 2.3 Estimation of kinship categories and coefficients

To estimate the kinship category of each pair of mosquitoes, we calculated Loiselle’s kinship coefficient $k$ (Loiselle, Sork, Nason, & Graham, 1995) in the program $spagedi$ (Hardy & Vekemans, 2002). Kinship coefficients represent the probability that any allele scored in both individuals is identical by descent (Wright, 1922), with theoretical mean $k$ values for each kinship category as follows: full-siblings $= 0.25$, half-siblings $= 0.125$, full-cousins $= 0.0625$, half-cousins $= 0.0313$, second cousins $= 0.0156$ and unrelated $= 0$. We ignored intergenerational kinship categories (e.g. uncle-niece), as the maximum of two weeks between the sampling of any pair of individuals would be insufficient time for the ontogeny of an additional generation (Christophers, 1960).

To assign pairs of individuals to relatedness categories across three orders of kinship (i.e. first cousins), we first used maximum-likelihood estimation in the program ML-RELATE (Kalinowski, Wagner, & Taper, 2006) to identify first-order (full-sibling) and second-order (half-sibling) pairs. We used the $k$ scores of pairs within the full-sibling and half-sibling data sets to calculate standard deviations for these categories.

However, ML-RELATE is not configured to determine third-order relationships (e.g. cousins). To determine cousins, we estimated a lower bound of $k$ that separated first cousins from unrelated pairs and those of more distant kin groups. Here, we define first cousins as including both full-cousins and half-cousins (Figure 1). We then produced simulated $k$ scores for each kinship category, assuming that the $k$ scores within each kinship category followed a normal distribution with a unique mean and standard deviation and that these scores combined produced the entire distribution of $k$ scores. Standard deviations for full-cousins, half-cousins and second cousins were assumed to correspond to the standard deviation of the entire population after full-sibling and half-sibling pairs were removed.

Using the theoretical means and standard deviations of $k$, we randomly sampled 100,000 simulated $k$ scores from each kinship category. In the initial pool of 13,041 empirical mosquito pairings, ML-RELATE identified approximately 50 full-sibling and half-sibling pairs. Assuming that the data contained twice as many first cousin (full and half) pairings as sibling (full and half) pairings, and twice as many second cousin pairings as first cousin pairings, final sampling distributions were developed as follows: 100,000 unrelated, 2,000 s cousins, 500 full-cousins, 500 half-cousins, 250 half-siblings and 250 full-siblings, giving a ratio of 400:8:2:2:1:1. This assumption is reasonable if the local population size is approximately constant; for a diploid population of constant size, an average of two offspring from any one individual will themselves have offspring. If we take $n$ to be the number of larvae parented by one individual, the ratio of first cousin to sibling pairings would be approximately $\frac{n_{\text{first cousins}}}{n_{\text{siblings}}} = \frac{2^2}{2(2^n-1)}$, which ranges from 1.1 to 4 for $2 \leq n \leq 10$. If we allow a slight population expansion with an average of 2.4 reproducing offspring, ratios from 1 $\leq n \leq 10$ produce values from 2.2 to 2.4.

To analyse how closely this distribution approximates the field data, we randomly sampled 10,000 simulated $k$ scores from the above sampling distribution and plotted a histogram of this combined distribution and a histogram of the unrelated distribution against a histogram of 10,000 $k$ scores from the empirical data. As the combined distribution matched the empirical distribution much more closely than the unrelated distribution (Figure S1), we adopted it for kinship inference. Note that this unrelated distribution is not completely reflective of a natural population, as it assumes total random mating and zero population structure. It is a convenient null from which various models including this one can be developed and compared.

To determine a lower threshold of $k$ to define a kinship category for first cousins, we plotted histograms of $k$ scores for each kinship category, following the ratio described above (Figure S2). We assigned $k = 0.06$ as the lower threshold to describe first cousin relationships. Individual pairs of $k > 0.06$ that were neither full-siblings nor half-siblings were much more likely to be first cousins (either full or half) than any other category.

### 2.4 Inference of dispersal distributions

We treated separation distances between pairs of sampled kin as representative of multiple past dispersal events, following Figure 1. Full-sibling separation distances represent the dispersal of the mother during oviposition. Half-sibling separation distances represent the breeding dispersal of the father between matings, plus the postmating dispersal of each mother between mating and oviposition, plus the ovipositional dispersal of each mother. Full-cousin separation distances represent the ovipositional dispersal of the grandmother, plus the premating dispersal of each parent, plus the postmating dispersal and ovipositional dispersal of each mother. Half-cousin separation distances are as those for full-cousins, plus the breeding dispersal of the father between matings and the postmating dispersal of each mother between mating and oviposition. In all cases, we considered dispersal to operate across a two-dimensional plane and ignored any differences in sampling altitude between buildings.

---

**FIGURE 2** Mentari Court geography, trap placement and kinship networks. Lines indicate pairs of full-siblings (a), half-siblings (b) and cousins (c). Rotated squares indicate ovitrap locations. Dashed lines show the radii within which 86.5% of pairs in that kinship category are expected to be found [Colour figure can be viewed at wileyonlinelibrary.com]
Axial standard deviations correspond in form to the dispersal component of NS (Wright, 1946). We calculated axial standard deviations from the distributions of separation distances for each kinship category. To do this, we (a) projected distances onto a polar coordinate system with a random angle of rotation, (b) converted the distances to one-dimensional vectors by multiplying each distance by the cosine of its rotation angle and (c) calculated the standard deviation of the resulting distribution. The final axial standard deviations were estimated by applying steps (a)–(c) to each kinship category. See Appendix S1.

2.5 Determination of parent–offspring dispersal

Parent–offspring dispersal is described by Wright (1946) as the distribution of parents at some phase of the life cycle relative to offspring at the same phase. These events serve to produce geographical distributions of not only full-siblings and half-siblings but also full-cousins and half-cousins as well (Figure 1). Specifically, the full-cousin and half-cousin dispersal distributions are produced by combining the parent–offspring dispersal distributions with the full-sibling and half-sibling distributions, respectively. As the variance of the normal distribution formed by the combination of two normal distributions is equal to the sum of the component variances, we can infer parent–offspring axial dispersal, σ_{PO}, with the following equations:

\[ \sigma_{PO}(full\text{-}cousins) = \sqrt{\sigma^2_{full\text{-}siblings} - \sigma^2_{full\text{-}cousins}} \] (1)

\[ \sigma_{PO}(half\text{-}cousins) = \sqrt{\sigma^2_{half\text{-}siblings} - \sigma^2_{half\text{-}cousins}} \] (2)

As we assume a mixed distribution of full-cousins and half-cousins in the kinship pairings, these represent the upper (full-cousin) and lower (half-cousin) bounds of parent–offspring dispersal estimates using the first cousin distribution. Assuming a 50:50 ratio of full-cousins and half-cousins in our data, the mean parent–offspring axial dispersal can be approximated by the following equation:

\[ \sigma_{PO(first\text{-}cousins)} = \sqrt{\sigma^2_{first\text{-}cousins} - \frac{1}{2} (\sigma^2_{full\text{-}siblings} + \sigma^2_{half\text{-}siblings})} \] (3)

As dbRDA requires independent variables to have one-dimensional inputs, we first transformed the matrix of pairwise geographical distances into eight principal components (PCs) using the function pcnm. Sampling time was modelled as a continuous variable of integers corresponding to the first, second and third week of sampling, while building residency was modelled categorically. We used a pairwise matrix of Rousset’s a scores (Rousset, 2000) as the dependent variable, calculated in SPAGEDi. For geographical distance in the dbRDAs “dist|env” and “env|dist”, we used only the PCs that were significant in the dbRDA “dist” (p < 0.006 following Bonferroni correction).

In a population experiencing isolation by distance, NS is estimated as the inverse of the slope of the regression of pairwise genetic distance against the natural logarithm of geographical distance (Rousset, 2000). That is, \( NS = b^{-1} \), where Genetic distance = \( a + \ln(\text{Geographical distance}) \times b \). We used the R function “lm” to perform three of these regressions. The first, using all nonsibling pairs of individuals, we used to estimate NS following the above. For the other two regressions, we considered only within-building pairings and between-building pairings, to investigate isolation by distance within and between buildings, respectively.

3 RESULTS

3.1 Kinship inference

Using the 3,939 SNPs retained after filtering, we identified 13 full-sibling pairs and 34 half-sibling pairs using ML-Relate. We also designated 51 pairs with k > 0.06 as first cousins. Mean separation distance for full-siblings was 18.1 m; for half-siblings, it was 48.6 m; and for cousins, it was 75.1 m (Table 1). Figure S3 shows the spatial
distribution of $k$ scores across geographical distance for all kinship categories as scatterplots (3A) and density histograms (3B).

### 3.2 | Determination of parent–offspring dispersal via kinship distributions

For each kinship category, we used geographical distances of separation between kin to construct a bivariate normal distribution, from which we calculated the axial standard deviation of dispersal (Table 1). Using the full-sibling, half-sibling and first cousin standard deviations, we estimated parent–offspring axial dispersal as 64 m (23–93 m; 95% C.I.). The radius of effective parent–offspring dispersal is equal to double the parent–offspring axial standard deviation, coming to 129 m (47–185 m; 95% C.I.). This corresponds to neighbourhood area, the circle within which 86.5% of parent–offspring dispersal occurs.

Following the process for parent–offspring dispersal, we generated dispersal radii for each kinship distribution. Referencing the axial SD in Table 1, we derived a dispersal radius of 36 m (12–59 m; 95% C.I.) for full-siblings, representing ovipositional dispersal; a dispersal radius of 52 m (35–101 m; 95% C.I.) for half-siblings, representing breeding and ovipositional dispersal; and a dispersal radius of 75 m (52–101 m; 95% C.I.) for first cousins, representing ovipositional dispersal. A dispersal radius of 116 m (53–180 m; 95% C.I.) for half-siblings, representing ovipositional dispersal; a dispersal radius of 129 m (47–185 m; 95% C.I.) for full-siblings, representing ovipositional dispersal; and a dispersal radius of 154 m (101–202 m; 95% C.I.) for first cousins. Dispersal radius describes the radius within which 86.5% of dispersed individuals are expected to be found.

### 3.3 | Geographical genetic structure and neighbourhood size

The dbRDA evaluating the effects of geographical distance on genetic structure ("dist") indicated that 3 of 8 PCs were within the significance threshold (Bonferroni-corrected critical value: $p < 0.006$). These PCs were used as the geographical distance covariate in the dbRDAs env|dist and dist|env (Table 2). After conditioning for the effects of geographical distance (Bonferroni-corrected critical value: $p < 0.025$), env|dist showed that building residency had a significant effect on patterns of genetic structure ($p = 0.002; F = 1.67; 7 df$), while sampling week did not ($p = 0.854; F = 0.690; 2 df$). After conditioning for the effects of building residency and sampling week, the 3 PCs in dist|env showed no significant association between geographical distance and genetic distance ($p = 0.090; F = 1.155; 3 df$). As building residency and geographical distance were necessarily correlated, these results indicate a significant effect of building residency independent of geographical distance, and an uncertain effect of geographical distance.

We explored the effects of geographical distance further using linear regressions. The three linear regressions calculated for all nonsibling pairs ($p < 2e^{-16}; R^2 = 0.0090; \text{slope} = 0.0037$), all nonsibling pairs within the same building ($p = 0.012; R^2 = 0.0062; \text{slope} = 0.0041$), and all nonsibling pairs in different buildings ($p = 1.3e^{-09}; R^2 = 0.0048; \text{slope} = 0.00484$), all showed significant positive associations between geographical and genetic distances.

---

**Table 1** Estimates of dispersal distances in metres for each kinship category. Mean distance describes empirical distances.

<table>
<thead>
<tr>
<th>Category</th>
<th>Mean distance</th>
<th>Axial distance</th>
<th>95% C.I.</th>
<th>Dispersal radius</th>
<th>95% C.I.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Full-siblings</td>
<td>18.1</td>
<td>18</td>
<td>5.9–30</td>
<td>36</td>
<td>12–59</td>
</tr>
<tr>
<td>Half-siblings</td>
<td>48.6</td>
<td>58</td>
<td>27–90</td>
<td>116</td>
<td>53–180</td>
</tr>
<tr>
<td>First cousins</td>
<td>75.1</td>
<td>77</td>
<td>52–101</td>
<td>154</td>
<td>101–202</td>
</tr>
</tbody>
</table>

**Table 2** Results of dbRDAs. These test effects of building residency and sampling week on genetic distance after conditioning for geographical distance (env|dist), and effects of geographical distance after conditioning for building residency and sampling week (dist|env).

<table>
<thead>
<tr>
<th>Factor</th>
<th>df</th>
<th>Sum of squares</th>
<th>$F$ statistic</th>
<th>$p$ value</th>
</tr>
</thead>
<tbody>
<tr>
<td>env</td>
<td>dist</td>
<td>Building</td>
<td>7</td>
<td>0.1116</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Week</td>
<td>2</td>
<td>0.0132</td>
</tr>
<tr>
<td>dist</td>
<td>env</td>
<td>Distance (3 PCs)</td>
<td>3</td>
<td>0.0648</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Residual</td>
<td>119</td>
<td>2.2256</td>
</tr>
</tbody>
</table>

*Significance at the Bonferroni-corrected critical value of $p < 0.025$ (env|dist) or $p < 0.05$ (dist|env).
inhabiting a circle of radius 258 m, which is two times the axial standard deviation of parent–offspring dispersal (Wright, 1946).

4 | DISCUSSION

This study has developed a methodology that uses genomic inferences of kinship to estimate parent–offspring dispersal and related population parameters, and applied this methodology to investigate an urban population of the invasive disease vector Ae. aegypti. From 162 sequenced individuals, we identified 98 pairs of kin across three orders of kinship. By limiting the duration of sampling and by sampling only eggs, we were able to ignore intergenerational kin categories and thus assigned the 98 pairs to intragenerational categories: full-sibling, half-sibling, full-cousin and half-cousin. After removing the spatial distances separating kin pairs within each category into axial standard deviations, we isolated the parent–offspring component of dispersal, which corresponds to the difference between the axial standard deviations of full-siblings and full-cousins and of half-siblings and half-cousins. Using our final estimate of parent–offspring dispersal, we calculated neighborhood area (129 m), median parent–offspring dispersal distance (75 m) and radius of ovipositional dispersal within a gonotrophic cycle (36 m). Our additional analyses with dbRDA and linear regression revealed isolation by distance both within and between buildings and provided an estimate of NS (Wright, 1946) of 268 individuals. These results indicate the scale of intergenerational dispersal and population structure in Ae. aegypti, and thus also the scale at which efforts must be focused to restrict new invasions, suppress or transform established populations, and respond strategically to arboviral outbreaks. With appropriate modifications where required, our methodology will be readily applicable to studies of other species of interest such as pests and those of conservation significance.

Our estimates of parent–offspring dispersal, combined with the results of the dbRDAs and linear regressions, suggest that the Ae. aegypti population of Mentari Court is hierarchically structured. The linear regressions indicate isolation-by-distance patterns both within (<200 m) and between buildings, while the dbRDAs showed that building residency had an effect on structure even after accounting for geographical distance. These results suggest dispersal within buildings is less restricted than dispersal between buildings, and that dispersal between adjacent buildings is more common than between nonadjacent buildings. This interpretation accords with kin observations, wherein most kin pairs found in nonadjacent buildings were cousins (Figure 2), which are a result of multiple parent–offspring dispersal events (Figure 1). This also accords with our estimate that 86.5% of parent–offspring dispersal occurs within a circle of radius 129 m, which corresponds in scale to movement within buildings and between adjacent buildings but not between nonadjacent buildings. Thus, dispersal of Ae. aegypti between nonadjacent buildings at Mentari Court is likely to be a multigenerational process. This equates to a matter of months for Ae. aegypti, which typically has 10–12 generations per year (Christophers, 1960).

These findings are broadly consistent with previous estimates from MRR studies which describe a tendency of Ae. aegypti to stay within the building of its release, and when dispersing farther rarely travel more than 150 m (Harrington et al., 2005; Maciel-de-Freitas, Codeco, & Lourenco-de-Oliveira, 2007; Ordonez-Gonzalez, Mercado-Hernandez, Flores-Suarez, & Fernandez-Salas, 2001; Russell, Webb, Williams, & Ritchie, 2005). Occasional long-distance dispersal >500 m has also been recorded through MRR (Honorio et al., 2003; Reiter et al., 1995); while these dispersal distances are larger than the sampling scale of this study, they have been observed in genomic studies (Schmidt et al., 2018). Likewise, our finding that buildings or the space between them acts as a dispersal barrier is consistent with observations of fine-scale dispersal barriers in Ae. aegypti (Hemme, Thomas, Chadee, & Severson, 2010; Schmidt et al., 2018). The 18-storey high-rise blocks of Mentari Court are populated with Ae. aegypti up to the top floor and are likely to provide adequate within-building opportunities for mating, human blood-feeding, and oviposition, with dispersal within buildings likely to be safer and easier than crossing the more "hostile" open spaces between buildings. Future work could examine dispersal patterns in building complexes of fewer storeys to test whether these lead to higher rates of movement between buildings, which could result from smaller buildings having fewer oviposition sites. Future work could also investigate vertical dispersal between floors, which was not considered in this study.

If movement patterns of Ae. aegypti through Mentari Court are representative of high-rise building complexes generally, we expect these findings to have implications for understanding Ae. aegypti as an infectious disease vector, as an invasive species, and as a target of biological control initiatives. In the event of an arbovirus outbreak, the initial spread of the virus by female Ae. aegypti will likely be restricted to within buildings or between nearby buildings. Spread by females to more distant buildings would likely take multiple generations, or multiple gonotrophic cycles of a single female, and thus proceed more slowly. Following a new detection of invasive Ae. aegypti, we expect limited active dispersal from the point of introduction, though passive dispersal by human transport may spread the incursion more quickly. Finally, considering releases of Wolbachia-infected Ae. aegypti at Mentari Court, we expect that successful invasion of one building with Wolbachia is unlikely to result in the invasion of other buildings, as movement between buildings may be insufficiently common for the invasion to surpass a critical threshold frequency of 0.25–0.35 (Schmidt, Barton, et al., 2017; Turelli & Barton, 2017). However, limited dispersal between buildings is also likely to ensure that a successfully invaded building remains so, as few infected Ae. aegypti will leave and few uninfected Ae. aegypti will arrive.
Using the regression between individual genetic distance and the natural logarithm of geographical distance (Rousset, 2000), we estimated \textit{Ae. aegypti} NS at 268 individuals. Combining this with our parameter for parent–offspring dispersal (\( \rho \)) of 129 m, we can estimate the ideal density \( d \) of \textit{Ae. aegypti} breeding adults as 1,281 km\(^{-2}\), using Wright’s equation for NS under isolation by distance, \( NS = 4\pi \rho^2 d \) (Wright, 1946). In interpreting these values, it is essential to be clear about the distinction between NS and effective population size (\( N_e \)). NS is a parameter representing breeding structure in an isolation-by-distance framework and relates directly to dispersal, while \( N_e \) is dependent on local species abundance and thus habitat availability (Nunney, 2016). These parameters are therefore expected to estimate different numbers of individuals. An extensive study of \textit{Ae. aegypti} estimated average \( N_e \) of 400 – 600 across different sites and time points (Saarman et al., 2017), while census size estimates range from 900–5,500 individuals (Carvalho et al., 2015; Lounibos, 2003; Sheppard, Macdonald, Tonn, & Grab, 1969). Relative census sizes of populations can be estimated using variations in \( N_e \) and may improve our ability to understand \textit{Ae. aegypti} population characteristics. Census size estimates are often five- to ten-fold larger than those of Ne (Luikart, Ryman, Tallmon, Schwartz, & Allendorf, 2010; Palstra & Fraser, 2012); were a similar ratio to hold for NS, we would expect to find between 1,340 and 2,680 \textit{Ae. aegypti} within the dispersal radius, a figure consistent with typical adult census population sizes for this species. All of this emphasizes how combining kinship-based dispersal estimates with more widely used, distance-based investigations of genetic structure can improve our understanding of local population and breeding processes in \textit{Ae. aegypti}.

Apart from the introduction of novel oviposition sites through ovitraps, our methodology of inferring dispersal through genetic relatedness leaves dispersal processes essentially undisturbed, an advantage over MRR studies. Accordingly, genomic inferences of dispersal may more accurately reflect the true biological processes occurring within the study site. An additional advantage of kinship-based methodologies is their capacity to investigate population processes such as dispersal among populations in which migration rates have changed too recently to have produced corresponding changes in population genetic structure (Palsbøll et al., 2010), a situation likely to be common in invasive populations and those with habitats impacted by humans. Ongoing decreases in sequencing costs mean that the approach is becoming feasible for a wider range of species, particularly as our methodology parameterizes dispersal using estimates from three orders of kinship and does not depend on large sample sizes required for many kinship-based methodologies. When stringent requirements on sample size are relaxed, more widespread adoption of kinship-based methodologies seems likely (Palsbøll et al., 2010), particularly since current genetic markers allow for more accurate estimation of kinship categories than previously used markers (Hauser, Baird, Hilborn, Seeb, & Seeb, 2011). In this study, we sequenced 162 \textit{Ae. aegypti} on only two sequencing lanes, which provided sufficient read depth and breadth to identify 98 pairs of close kin.

Our methodology is applicable to a range of species, including those of medical, economic or conservation significance. However, differences in breeding and dispersal ecologies between species will alter the way that the method is applied. For instance, for species where females mate more than once, additional steps will be required to separate half-sibling dispersal from maternal dispersal. Similarly, species with sex-biased dispersal will require additional steps when decomposing axial standard deviations into components. Also, while our method is well-suited to studying rare species of conservation significance, in which direct population manipulation through MRR may not be possible, these studies will likely have low sample size and involve inbred individuals (Hedrick & Kalinowski, 2000). These conditions may make it more difficult to estimate kinship using genomics, though we note that parts of our methodology can equally be applied using kinship estimates derived through other means such as visual tracking of relatives.

The kinship-based inference processes described here have close affinities with the work of Bravington et al. (2016) and constitute an important extension to CKMR. Using genome-wide SNPs, we have been able to extend the CKMR methodology to cover three orders of kinship, estimating relatedness to the level of half-cousin. Likewise, this study has developed an approach for estimating intergenerational dispersal through the genomic marking of unsampled parents and grandparents. This will be particularly helpful for studies of species with traits such as small size or high population density that make them unsuitable for standard MRR experiments. For example, monitoring of rare mammals is often carried out by the scats they leave, with recent work using DNA from scats to infer \( N_e \) from kinship (Skrbinšek et al., 2012). As methods for retrieving DNA from scats improve (Schultz, Cristescu, Littleford-Colquhoun, Jaccoud, & Frère, 2018), investigating kin networks and intergenerational dispersal with high-density markers may become increasingly viable for achieving conservation objectives.

5 | CONCLUSIONS

We developed and applied a methodology to investigate dispersal and breeding dynamics in \textit{Ae. aegypti} from the spatial distribution of 98 pairs of kin across three orders of kinship. We also observed genetic structure within and between buildings at Mentari Court, with most movement between nonadjacent buildings taking place over multiple generations. These findings are being considered in the design of future \textit{Wolbachia} releases at Mentari Court and other high-density urban sites, and they will also inform protocols for the effective deployment of resources in response to disease outbreaks. Our methodological approach for investigating dispersal and breeding dynamics will be readily adaptable to future studies of dispersal in a range of species.

ACKNOWLEDGEMENTS

This work was funded by the National Health and Medical Research Council (Program Grant no. 1132412; Fellowship Grant no. 1118640), and the Wellcome Trust (Grant no. 108508). We thank three anonymous reviewers for their comments.
AUTHOR CONTRIBUTIONS

Research was designed by N.W.A., A.A.H. and S.P.S.; research was performed by M.J., T.L.S. and N.W.A.; manuscript was written by M.J. and T.L.S.; new analytical tools were provided by M.J.; data were analysed by M.J.; supervision was done by T.L.S. and A.A.H.; and funding was provided by S.P.S. and A.A.H.

DATA ACCESSIBILITY

Sorted .bam files and geographical locations for 162 Aedes aegypti have been archived at the Sequence Read Archive at NCBI Genbank, with accession number PRJNA542421.

ORCID

Moshe Jasper https://orcid.org/0000-0003-4541-3223
Thomas L. Schmidt https://orcid.org/0000-0003-4695-075X
Ary A. Hoffmann http://orcid.org/0000-0001-9497-7645

REFERENCES

Hoffmann, A. A., Montgomery, B. L., Popovici, J., Iturbe-Ormaetxe, I., Johnson, P. H., Muzzi, F., ... O'Neill, S. L. (2011). Stability of the wMel Wolbachia Infection following invasion into Aedes aegypti populations. PLOS Neglected Tropical Diseases, 8(9), e3115. https://doi.org/10.1371/journal.pntd.0003115
Maino, J. L., Binns, M., & Umina, P. (2018). No longer a west-side story—pesticide resistance discovered in the eastern range of a major


How to cite this article: Jasper M, Schmidt TL, Ahmad NW, Sinkins SP, Hoffmann AA. A genomic approach to inferring kinship reveals limited intergenerational dispersal in the yellow fever mosquito. Mol Ecol Resour. 2019;19:1254–1264. https://doi.org/10.1111/1755-0998.13043

SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section at the end of the article.