

## Genomic Analyses of Anopheline Mosquitoes: Micro- and Macro-Geographic Population Structure of Two Important Malaria Vectors

Rachel M. Wiltshire

### Publication Date

10-04-2019

### License

This work is made available under a All Rights Reserved license and should only be used in accordance with that license.

### Citation for this work (American Psychological Association 7th edition)

Wiltshire, R. M. (2019). *Genomic Analyses of Anopheline Mosquitoes: Micro- and Macro-Geographic Population Structure of Two Important Malaria Vectors* (Version 1). University of Notre Dame.  
<https://doi.org/10.7274/v692t438t8k>

This work was downloaded from CurateND, the University of Notre Dame's institutional repository.

For more information about this work, to report or an issue, or to preserve and share your original work, please contact the CurateND team for assistance at [curate@nd.edu](mailto:curate@nd.edu).

GENOMIC ANALYSES OF ANOPHELINE MOSQUITOES: MICRO- AND MACRO-  
GEOGRAPHIC POPULATION STRUCTURE OF TWO IMPORTANT MALARIA VECTORS

A Dissertation

Submitted to the Graduate School  
of the University of Notre Dame  
in Partial Fulfillment of the Requirements  
for the Degree of

Doctor of Philosophy

by

Rachel M. Wiltshire

---

Frank H. Collins, Director

Graduate Program in Biological Sciences

Notre Dame, Indiana

April 2019

© Copyright by  
RACHEL M. WILTSHIRE  
2019  
All rights reserved

GENOMIC ANALYSES OF ANOPHELINE MOSQUITOES: MICRO- AND MACRO-  
GEOGRAPHIC POPULATION STRUCTURE OF TWO IMPORTANT MALARIA VECTORS

Abstract

by

Rachel M. Wiltshire

Malaria continues to be a major public health issue. 219 million cases and 435,000 deaths were estimated to have occurred in 2017. Disease transmission is a complex epidemiological interaction comprising causal factors that are dependent on local conditions; however, the host-vector-parasite association remains constant.

This dissertation explored how the genetic and bionomic behaviors of two important anophelines might respond to alternative control interventions in attempts to interrupt sustained malaria transmission.

Genetically modified (GM) mosquitoes have been demonstrated as a form of innovative vector control. To advance into field trials, a genetic assessment of the target population must be conducted. Six *Anopheles gambiae* populations were sampled from the Lake Victoria basin and sequenced. 5,175 single nucleotide polymorphism (SNP) markers were analyzed. Principal components analysis (PCA) illustrated individuals clustered according to geography with some intersection. Genetic differentiation ( $F_{ST}$ ) was variable with inter-island comparisons having the

highest values (0.0480-0.0846). Estimates of effective size were small (124.2-1920.3). These results indicated that three island populations could be candidates for small-scale field testing of GM mosquitoes.

A genome-wide approach to the analysis of population structure has the power to identify genetic processes (such as restricted gene flow or an insecticide-resistant genotype) and the genomic regions that regulate them, providing greater insight into how these mechanisms will impact upon control interventions. A preliminary analysis of 17,757 SNPs across four *An. farauti* populations in the southwest Pacific confirmed their geographic isolation from each other but also detected a previously unidentified genetic association between, and lack of diversity within, Queensland (Australia) and Vanuatu (PCA, ADMIXTURE). Further detailed analysis of the SNPs contributing to this relationship is recommended.

In Western Province, Solomon Islands, a proportion of *An. farauti* mosquitoes, which may be contributing to residual transmission, avoid lethal doses of insecticide from long-lasting insecticidal nets. Sugar-fermented yeast was evaluated as an organic source of carbon dioxide (CO<sub>2</sub>) trap attractant to host-seeking and resting mosquitoes alongside human-produced CO<sub>2</sub> and a control. Although the human attracted the greatest numbers (n=349), sugar-fermented yeast (n=210) demonstrated attractiveness, and improvements to the source design i.e. plume composition and delivery, could further enhance its appeal.

## DEDICATION

This is for N, M, and J for allowing me the freedom to chase my dreams;  
#superJericho, my trusty steed; and the #wondercats—Oscar and Scarlett—who  
generated the extensive use of this hashtag: #crazycatlady.

## CONTENTS

FIGURES.....	vii
TABLES.....	xi
ACKNOWLEDGMENTS .....	xiii
LIST OF ABBREVIATIONS.....	xx
CHAPTER 1: MALARIA: A MOST PERSISTENT PARASITIC PROBLEM .....	1
1.1 Biology, transmission, disease .....	1
1.2 A numbers game.....	2
1.3 Disease management: present day.....	3
1.3.1 Asymptomatic infections.....	3
1.3.2 Clinical malaria.....	4
1.3.3 Strategy .....	5
1.3.4 Prevention .....	6
1.3.4.1 Intermittent preventive therapy (IPT) .....	6
1.3.4.2 Seasonal malaria chemoprevention (SMC) .....	7
1.3.4.3 Vector control (adults).....	7
1.3.4.4 Supplemental techniques (juvenile stages) .....	8
1.3.4.5 Communication and behavioral impact .....	9
1.3.5 Treatment.....	9
1.4 Vector control .....	12
1.4.1 Preservation and improvement of the utility of existing insecticide-based interventions .....	13
1.4.1.1 Target site insensitivity .....	13
1.4.1.2 Metabolic resistance .....	14
1.4.2 Interventions that affect vector species not effectively targeted by current tools.....	16
1.4.3 Innovative approaches that will permanently reduce the very high vectorial capacities of the dominant malaria vectors in sub-Saharan Africa .....	17
1.5 Integration of current trial products into future programs .....	17
1.6 An extended research agenda .....	18
1.7 Going global .....	20

CHAPTER 2: DETERMINING THE POPULATION GENETIC STRUCTURE OF <i>ANOPHELES GAMBIAE</i> IN THE NORTHWESTERN LAKE VICTORIA BASIN BY REDUCED-REPRESENTATION SEQUENCING .....	21
2.1 Abstract .....	22
2.2 Introduction.....	24
2.2.1 Global burden .....	24
2.2.2 Insecticide-based vector control tools.....	24
2.2.3 GM mosquitoes as a vector control solution .....	25
2.2.4 <i>An. gambiae</i> species complex .....	27
2.2.5 <i>An. gambiae</i> population structure across the African continent	27
2.2.6 Previous Lake Victoria <i>An. gambiae</i> population genetic studies	28
2.2.7 Current objective.....	29
2.3 Materials and methods.....	30
2.3.1 Study area.....	30
2.3.2 Entomological sampling .....	34
2.3.3 Species identification and preservation.....	36
2.3.4 Genomic DNA extraction .....	37
2.3.5 RADseq library construction and sequencing .....	38
2.3.6 Determination of 2L chromosomal karyotypes .....	40
2.3.7 Bioinformatics processing .....	40
2.3.8 Population genetics .....	42
2.4 Results.....	43
2.4.1 Species identification and dataset composition.....	43
2.4.2 Chromosomal mapping and distribution of SNPs .....	45
2.4.3 Population structure.....	45
2.4.4 Genetic differentiation .....	52
2.4.5 Effective population size .....	55
2.5 Discussion.....	57
2.5.1 There are limited malaria vector species in the Ssese Islands ...	57
2.5.2 The 2La inversion confounds population genetic structure .....	57
2.5.3 Moderate but significant genetic differentiation is observed in island populations .....	58
2.5.4 Island populations have small effective sizes .....	60
2.6 Conclusion.....	61
CHAPTER 3: IDENTIFYING THE GEOGRAPHIC STRUCTURE OF <i>ANOPHELES FARAUTI</i> POPULATIONS IN THE SOUTHWEST PACIFIC USING SINGLE NUCLEOTIDE POLYMORPHISM MARKERS.....	62
3.1 Systematics of malaria vectors in the southwest Pacific.....	64
3.2 Current objective .....	70
3.3 Materials and methods.....	71
3.3.1 Entomological sampling .....	71
3.3.2 Species identification and preservation.....	72
3.3.3 Whole genome library construction and sequencing.....	73
3.3.3.1 Reference genome.....	73



3.3.3.2 Samples from natural populations .....	74
3.3.4 Bioinformatics processing .....	74
3.3.4.1 AfarF1 assembly .....	74
3.3.4.2 Wild-type sequence data .....	75
3.3.5 Population structure .....	75
3.4 Results.....	76
3.4.1 Species identification .....	76
3.4.2 <i>An. farauti</i> reference genome assembly .....	77
3.4.3 Wild-type <i>An. farauti</i> samples .....	77
3.4.4 Population structure .....	80
3.5 Discussion.....	83
3.5.1 Australia and Vanuatu populations are genetically similar .....	84
3.5.2 Australia and Vanuatu individuals demonstrate little genomic variation in the generated SNP dataset.....	86
3.5.3 Greater structure is observed in the Solomon Islands population.....	87
3.6 Conclusion.....	88

#### CHAPTER 4: FIELD EVALUATION OF SUGAR-FERMENTED YEAST AS AN ORGANIC SOURCE OF CARBON DIOXIDE TO ATTRACT *ANOPHELES FARAUTI*

MOSQUITOES IN WESTERN PROVINCE, SOLOMON ISLANDS .....	89
4.1 Malaria in the Solomon Islands archipelago .....	90
4.1.1 Site selection .....	92
4.1.2 Physical geography .....	92
4.1.3 Climate .....	93
4.1.4 Sociodemography .....	95
4.1.5 <i>Plasmodium</i> species prevalence .....	99
4.1.6 Glucose-6-phosphate dehydrogenase deficiency .....	100
4.2 Mosquito sensory physiology .....	102
4.3 CO <sub>2</sub> plume structure and dynamics .....	104
4.4 Current objective .....	104
4.5 Materials and methods: study site.....	105
4.5.1 Physical geography .....	105
4.5.2 Climate .....	107
4.5.3 Sociodemography .....	109
4.5.4 Malaria vectors .....	110
4.5.5 Transmission setting.....	110
4.6 Materials and methods: trap structure.....	112
4.6.1 CO <sub>2</sub> sources .....	112
4.6.1.1 Control .....	113
4.6.2 Trap design and assembly .....	114
4.7 Materials and methods: study design .....	118
4.7.1 Experimental model .....	118
4.7.2 Statistical analysis .....	118
4.7.3 Entomological sampling .....	119
4.7.4 Species identification and preservation .....	120

4.7.5 Climatic data.....	120
4.7.6 Ethics.....	120
4.8 Results.....	121
4.8.1 Trap placement.....	121
4.8.2 Sugar-fermented yeast CO <sub>2</sub> preparation.....	124
4.8.3 Climatic data.....	125
4.8.4 Morphological species identification .....	128
4.8.5 Attractiveness of CO <sub>2</sub> sources .....	128
4.8.6 Host-seeking/resting patterns.....	131
4.9 Discussion.....	132
4.9.1 Human-generated CO <sub>2</sub> attracted greater numbers of <i>An. farauti</i> .....	133
4.9.2 Peak host-seeking/resting behavior occurred between 21:00- 23:00 hours .....	134
4.10 Conclusion .....	135
CHAPTER 5: CONCLUSIONS.....	136
5.1 Research summary .....	136
5.1.1 Micro-population structure of <i>An. gambiae</i> in the Ugandan Lake Victoria basin .....	136
5.1.2 Macro-population structure of <i>An. farauti</i> in the southwest Pacific region .....	137
5.1.3 Sugar-fermented yeast as a source of CO <sub>2</sub> to attract <i>An. farauti</i> mosquitoes in Western Province, Solomon Islands.....	138
5.2 Contribution that extends the research agenda.....	138
5.3 Application to malaria vector control approaches .....	139
APPENDIX A: GLOBAL POSITIONING SYSTEM (GPS) COORDINATES FOR AND DISTANCES BETWEEN THE SIX <i>ANOPHELES GAMBIAE</i> ENTOMOLOGICAL SAMPLING SITES IN THE NORTHWESTERN LAKE VICTORIA BASIN .....	140
APPENDIX B: BIOINFORMATICS PIPELINE DETAILING SOFTWARE PROGRAMS AND PARAMETERS IMPLEMENTED IN THE ANALYSIS OF THE <i>ANOPHELES</i> <i>GAMBIAE</i> RADSEQ DATASET .....	143
APPENDIX C: DETERMINATION OF THE 2LA INVERSION KARYOTYPE IN NATURAL POPULATIONS OF <i>ANOPHELES GAMBIAE</i> MOSQUITOES SAMPLED FROM THE NORTHWESTERN LAKE VICTORIA BASIN.....	147
APPENDIX D: BIOINFORMATICS PIPELINE DETAILING SOFTWARE PROGRAMS AND PARAMETERS IMPLEMENTED IN THE ANALYSIS OF THE <i>ANOPHELES</i> <i>FARAUTI</i> WHOLE GENOME SEQUENCED DATASET .....	152
BIBLIOGRAPHY .....	156

## FIGURES

- Figure 2.1 **Sampling locations in northwestern Lake Victoria and the southern Uganda peninsular.** Inset top left: Uganda is highlighted in black to illustrate its location within continental Africa. Inset bottom left: Legend lists entomological sampling sites corresponding to numbered black crosses in the main schematic. The black star marks Kampala, the capital city. .... 31
- Figure 2.2 **Typical housing associated with villages, and small towns in the Ssesse Islands.** Walls are constructed from wooden boards and roofs are corrugated sheet metal (left), which are sometimes covered with a combination of plastic and thatched materials (right). Adult female *Anopheles* mosquitoes can be found resting indoors on walls and objects i.e. bed nets (usually near children) after taking a blood meal. They can also be seen higher up in the eaves of the property where the size of the gaps between the walls and roof are extensive allowing for frequent entry and exit to the exterior. .... 35
- Figure 2.3 **Principal components analysis (PCA) plot of the 5,175 genome-wide SNPs shared between the six *An. gambiae* populations (N=79).** Each circle represents an individual mosquito. Entomological sampling sites are color-coded according to the legend (Site) as follows: BK (Bukasa); BL (Bugala); EB (Entebbe); NZ (Nsadzi); SY (Sserinya), and WL (Wamala). The first principal component (PC1) and its percentage variance are represented by the y-axis, and likewise, the x-axis represents the second principal component (PC2) and its percentage variance. .... 47
- Figure 2.4 **PCA plot of the 1,078 chromosome 2L SNPs in the six *An. gambiae* populations (N=79).** Plot descriptors and legend are the same as for Figure 2.3. .... 48
- Figure 2.5 **PCA plot of the 4,097 SNPs in the six *An. gambiae* populations (N=79) after removal of the chromosome 2L SNPs (n=1,078) suspected of confounding genetic structure.** Plot descriptors and legend are the same as for Figure 2.3. Here, polygonal code was applied to better illustrate geographic rather than karyotypic clustering. .... 49
- Figure 2.6 **ADMIXTURE plot of chromosome 2L SNPs showing probable ancestry fractions (K=2) of the six *An. gambiae* populations (N=79) based on the 2La inversion.** Vertical bars represent individual mosquitoes,

which are stacked from left to right in groups of 16 (except: Bugala (N=13), Bukasa (N=13), and Sserinya (SY) (N=5)) according to their alphabetical sampling site. Populations are listed on the x-axis. Ancestry fractions ( $K=2$ ) are assigned according to maximum likelihood probabilities. Each fraction represents a karyotype of different color: blue = inverted ( $2L_a/a$ ), yellow = standard ( $2L^{+a}/2L^{+a}$ ), and blue/yellow = heterozygote ( $2L_a/+^a$ ). ..... 51

Figure 3.1 **Schematic drawing illustrating different proboscis coloration patterns in the Punctulatus Complex.** Top down: type A (*An. farauti*), B (*An. punctulatus*), and C (*An. koliensis*). Note the polymorphism within types B, and C. From Rozeboom & Knight (1946). ..... 67

Figure 3.2 **Predicted geographic distribution of the *Anopheles farauti* species complex.** Entomological sampling sites (Australia-Yorkeys Knob; Papua New Guinea-Madang; Solomon Islands-Haleta; Vanuatu-Tanna) are detailed in black open-framed rectangles. Adapted from the Malaria Atlas Project (2019)..... 69

Figure 3.3 **Principal components analysis illustrating genetic variance observed in 17,757 SNPs shared between the wild-type *An. farauti* individuals (N=16) across the four geographical sampling sites.** The y-axis represents the first principal component (PC1) and its percentage variance, and likewise, the x-axis represents the second principal component (PC2) and its percentage variance. .... 81

Figure 3.4 **Admixture plot of the 17,757 SNPs illustrating probable ancestry fractions ( $K=4$ ) of wild-type *An. farauti* individuals (N=16).** Vertical bars represent individual mosquitoes as listed on the x-axis. Ancestry fractions are assigned according to maximum likelihood probabilities. .... 82

Figure 3.5 **Admixture plot of the 17,757 SNPs illustrating probable ancestry fractions ( $K=5$ ) of the wild-type *An. farauti* individuals (N=16).** Vertical bars represent individual mosquitoes as listed on the x-axis. Ancestry fractions are assigned according to maximum likelihood probabilities. .... 83

Figure 4.1 **Map of the Solomon Islands archipelago.** Inset top right: the black-framed rectangle illustrates the location of the Solomon Islands within the southwestern Pacific region, and has been magnified to provide detail as shown in the main picture. Ontong Java (circled in red) is a coral atoll with approximately 2,000 inhabitants that is administrated by the Solomon Islands government through Malaita Province. .... 93

Figure 4.2 **Approximate positions of the Intertropical and South Pacific Convergence Zones during the austral warm season (November-April).** Black arrows depict near surface winds. The green circle illustrates the Solomon Islands. Papua New Guinea is described as PNG. Colored areas

represent landmasses (yellow), bands of rainfall (dark blue), and the Western Pacific Warm Pool (WPWP) (red), which is a unique climatic feature with a permanent sea surface temperature > 28°C giving it the capability to modify heating and cooling of the tropical Pacific Ocean. The y-axis depicts latitude in degrees, and the x-axis, likewise, longitude. Adapted from Jaffrés et al. 2018..... 95

Figure 4.3 **Map of the Solomon Islands’ administrative provinces.** Areas are demarcated by grey dotted lines. Honiara (marked by the black star), is also an administrative province but too small to mark as such at this magnification..... 96

Figure 4.4 **Scanning electron micrographs of *Culex quinquefasciatus* head, and olfactory structures.** (a) Maxillary palps are circled in white. (b: arrows) and (c: asterisks) illustrate individual peg sensilla at different magnifications. Scale bars: (a) 250 µm, (b) 100 µm, and (c) 5 µm. From (Syed and Leal 2007).103

Figure 4.5 **Visualization of odor plumes using TiCl<sub>4</sub> smoke.** The picture on the left illustrates the different types of plume that can be generated from an odorant stimulus. Voltage outputs on the right measured changes in smoke density as a demonstration of how plume type could influence mosquito flight choice towards an odorant stimulus during upwind flight. From (Geier, Bosch, and Boeckh 1999). .....104

Figure 4.6 **Map of the New Georgia group of islands.** Inset top right: The thick red-framed rectangle illustrates Western Province’s location within the Solomon Islands archipelago. ....106

Figure 4.7 **Monthly rainfall data recorded at Munda, Western Province, Solomon Islands (8.33°S, 157.27°E) from 2000-2013.** Mean rainfall is recorded in mm, and represented by the blue bars. A red seasonality trend line has been fitted to the data to illustrate the two rainy seasons, November-March and July-August, which are marked by the black arrows above the graph. Rainfall data was obtained from the Pacific-Australia Climate Change Science and Adaptation Planning Program (Australian Government Bureau of Meteorology 2018).....108

Figure 4.8 **Satellite image of Jack Harbour village illustrating major entomological and demographic features of the transmission setting.** The yellow outlines highlight swamp areas assumed to be *An. farauti* oviposition sites. The orange area encompasses the main village. In between the large swamp, and main village is a wide expanse of open grassed area (green rectangle) used mostly by the villagers for leisure activities such as soccer. The red-dashed rectangle demarcates the proposed placement of the carbon dioxide baited-traps. Notable structures are described accordingly. Map data: Google, 2018.....111

- Figure 4.9 **Portable mosquito tent used by the human volunteer to protect against potential bite activity during the experiment.** Entry and exit are accessed through the zippered door on the lateral aspect. ....114
- Figure 4.10 **Schematic diagrams of the trap design.** Length and width measurements presented from (a) aerial, and (b) lateral aspects, which also demonstrates entomological rationale. Female mosquitoes sense a CO<sub>2</sub> source (yellow square), and are guided by its plume to a potential blood meal. (1) Upward flight is initiated upon encountering the net material, which ceases upon (2) capture by the eave structures. ....116
- Figure 4.11 **Mosquito collectors assembling a trap.** Visible are the four corner poles used to frame the netting that formed the trap construct, which was held in place at a height of 0.9 m by mason line as indicated by the orange arrow. In the foreground is excess material, which when extended upwards 45° formed the eaves. These were secured in place using Velcro, and mason line for stability. ....117
- Figure 4.12 **Satellite image of Jack Harbour village illustrating re-placement of the experimental traps (red solid rectangles) to maximize *An. farauti* catch numbers.** Residences hosting traps are labeled Site 1, 2, or 3, and were located at elevation in between the two major *An. farauti* oviposition sites (yellow outlined swamps). Map data: Google, 2018. ....122
- Figure 4.13 **Location of the first trap (Site 1) after re-placement to the elevated area between the two suspected *An. farauti* oviposition sites.** The main residence is situated in the background to the left, and cooking area to the right. ....123
- Figure 4.14 **Boxplot summarizing attractiveness of CO<sub>2</sub> source to outdoor-biting populations of *An. farauti* over the 12-night experimental period.** The x-axis identifies the different CO<sub>2</sub> sources (control, yeast, human). Catch numbers are measured per night by the y-axis. Median values (solid black horizontal lines), quartiles (white boxes), variance (dashed black vertical lines), outliers (open circles), and error bars are indicated. ....129
- Figure 4.15 **Bar chart illustrating total number of anopheline mosquitoes captured per hour over the 12-night experimental period.** Mean temperature (°C) (orange line), and humidity (%) (dark blue line) measurements for the period are also presented. Black numbers above the blue bars are the total number of mosquitoes captured in the one hour time period indicated on the x-axis (hours). y-axes represent: left (temperature in °C), and right (humidity in %, and mosquito count). ....132

## TABLES

TABLE 2.1 GENETIC DIFFERENTIATION BETWEEN POPULATIONS AS MEASURED BY MEDIAN WEIR-COCKERHAM WEIGHTED $F_{ST}$ ESTIMATES .....	52
TABLE 2.2 ESTIMATES OF EFFECTIVE POPULATION SIZE ( $N_E$ ) ADOPTING A MINOR ALLELE FREQUENCY FILTER OF 5%.....	56
TABLE 3.1 DISTRIBUTION AND VECTOR STATUS OF THE <i>ANOPHELES LUNGAE</i> COMPLEX AND <i>AN. PUNCTULATUS</i> GROUP * .....	65
TABLE 3.2 GEOGRAPHIC METADATA FOR ENTOMOLOGICAL COLLECTIONS OF <i>ANOPHELES FARAUTI</i> S. L. MOSQUITOES SAMPLED FROM NATURAL POPULATIONS ACROSS ITS ECOLOGICAL DISTRIBUTION.....	72
TABLE 3.3 VARIANT CALL SUMMARY COUNTS OF RAW AND FILTERED DATASETS FOR THE 16 WILD-TYPE <i>ANOPHELES FARAUTI</i> INDIVIDUALS.....	79
TABLE 4.1 HUMAN POPULATION SIZE, DENSITY, DISTRIBUTION, AND MEAN ANNUAL GROWTH RATES FOR THE TEN ADMINISTRATIVE PROVINCES OF THE SOLOMON ISLANDS (2009).....	97
TABLE 4.2 PHENOTYPIC CLASSIFICATION OF GLUCOSE-6-PHOSPHATE DEHYDROGENASE VARIANTS BASED ON ERYTHROCYTE ENZYMATIC ACTIVITY AND ASSOCIATED CLINICAL OUTCOMES .....	102
TABLE 4.3 DAILY CLIMATIC CONDITIONS RECORDED DURING SUGAR-FERMENTED YEAST CO <sub>2</sub> PREPARATION .....	125
TABLE 4.4 OBSERVED CLIMATIC CONDITIONS RECORDED IN JACK HARBOUR DURING THE STUDY PERIOD, 16-29 SEPTEMBER 2014 .....	127
TABLE 4.5 SUMMARY ANALYSIS OF VARIANCE BETWEEN THE MEAN NUMBER OF MOSQUITOES CAPTURED IN THE THREE EXPERIMENTAL CONDITIONS (TRAPBAIT) .....	130
TABLE 4.6 COMPARISON OF GROUP MEANS (\$TRAPBAIT) BY TUKEY'S HONESTLY SIGNIFICANT DIFFERENCE <i>POST-HOC</i> TEST WITH 95% FAMILY-WISE CONFIDENCE LEVEL .....	131

TABLE A.1 GPS COORDINATES FOR AND DISTANCES BETWEEN THE SIX <i>ANOPHELES GAMBIAE</i> ENTOMOLOGICAL SAMPLING SITES IN THE NORTHWESTERN LAKE VICTORIA BASIN.....	141
TABLE B.1 SOFTWARE AND PARAMETERS APPLIED IN THE BIOINFORMATICS ANALYSIS OF 96 WILD-TYPE <i>ANOPHELES GAMBIAE</i> INDIVIDUALS SEQUENCED USING AN <i>ECORI</i> AND <i>MSERI</i> RESTRICTION SITE-ASSOCIATED DIGEST .....	144
TABLE C.1 2LA INVERSION KARYOTYPING RESULTS FOR <i>ANOPHELES GAMBIAE</i> MOSQUITOES SAMPLED FROM SIX NATURAL POPULATIONS IN THE NORTHWESTERN LAKE VICTORIA BASIN.....	148
TABLE D.1 SOFTWARE AND PARAMETERS APPLIED IN THE BIOINFORMATICS ANALYSIS OF 16 WILD-TYPE <i>ANOPHELES FARAUTI</i> INDIVIDUALS SEQUENCED BY A WHOLE GENOME SHOTGUN APPROACH .....	153



## ACKNOWLEDGMENTS

Many people have contributed to this journey. It began way back in 2004 outside a London pub one cold, wet December night with *Michelle Bayley* of Find Your Way Coaching. I thank her for the eighteen months of self-assessment, and reflection, which eventually steered me back to my bliss—scientific research.

I am indebted to *Mary Chambers* at the Oxford University Clinical Research Unit in Ho Chi Minh City, Vietnam, for trusting me with her group at the Hospital for Tropical Diseases in summer 2006. This research was the springboard of my mosquito fieldwork career, and also, I am sure, partly facilitated acceptance of my application to the Master's degree program at the London School of Hygiene & Tropical Medicine (LSHTM). A similar level of gratitude is reserved for *Sharon Cox* (LSHTM), who provided me with the opportunity to manage patient genotyping in the Sickle Cell Study at Muhimbili National Hospital, Dar es Salaam, Tanzania in 2009. I also thank *Brett Lowe* of the Kenya Medical Research Institute-Kilifi for technical support and encouragement during this time.

My advisor, *Frank Collins*, and the other members of my committee, Mike Pfrender, Mike Ferdig, Tom Burkot, and Nicole Achee have been nothing short of superb throughout the duration of these studies. Frank's approach took me outside of my comfort zone (field entomology) and placed me squarely in the center of a

most uncomfortable one (population genetics), but if he hadn't have done so then I never would have been introduced to bioinformatics—an unusual skill for a field entomologist—making me a unique individual within the discipline, and for that I am extremely grateful. I cannot thank *Mike Pfrender* enough for the time and patience that he has afforded me throughout this process but especially during the last academic year in his capacity as co-chair. He has turned me into a better scientist by teaching me the value of critical thinking and careful interpretation as opposed to my prior technique of over-excitement and rush-to-judgment. Another useful skill that I shall carry forward with me in my scientific career. *Mike Ferdig* has been the epitome of cool in complete contrast to my ever-increasing neuroticism. His dry wit and sly humor never failed to brighten my outlook, instantly making me feel like I was in control.

*Tom Burkot's* inclusion on my committee was somewhat of an unexpected surprise due to a departmental issue arising out of Nicole Achee's status as research, rather than teaching, professor. As it turned out, this was ratified and both of them remained as members, which—now with five—might seem greedy (and stressful) but their addition if nothing else brought the total number of “most eminent global malaria vector biologists” on my committee to three (including Frank). The first time I met Tom was in the Solomon Islands conducting the field studies for chapter four. I am extremely grateful for his assistance in setting up the logistics in Jack Harbour but mostly for the incident with Bob Cooper who was performing parity counts. As Tom was clearing up the makeshift lab, Bob squawked loudly “Don't touch those, I'm drying my ovaries!”. It is a phrase I have never ever heard a man say

before, and I doubt I ever will again. Tom looked startled. A range of emotions ensued. I laughed so hard that I cried. Then half an hour later they left. I cried again but the tears were of loneliness. Anyhow, Tom has always provided sound guidance and advice on global malaria policy and recommendations but also specifically on *An. farauti* bionomics. In both areas, he has consistently challenged me to think outside the box, which elicits a sort of “Pavlov’s Dogs” response from me now as in “what would Tom think about this?” but I am thankful that he has pushed me towards this way of thinking because, ultimately, it makes me a better scientist. *Nicole Achee* has consistently been a calming influence throughout her tenure as committee member. She has been a great role model to work alongside, which we did while teaching together in BIOS 30408, and also as someone to learn from through her work with various global organizations. I hope that our paths will continue to cross as I embark upon my career in vector biology.

Over the last six and a half years, *Nora Besansky* has acted as an unofficial sixth member of my committee providing constant guidance and measured advice for success in an advanced scholarly career. She has been a wonderful mentor and role model to follow not only for her professional achievements but also those she has developed in her Lake Tanganyika cichlid aquarium, which she has been kind enough to share with Katie and I after gifting us Pat, one of her *Neolamprologus similis* shell fish. I thank her for many hours of thought-provoking, intelligent conversation that was often interspersed with a wicked sense of humor. *Rebecca Wingert* and *Bridget Lautaru* have both dedicated time and support during the last academic year, which is greatly appreciated, to ensure that I reached this point.

At the University Counseling Center, I owe heartfelt thanks to *Dr. Maureen Lafferty* and *Dr. Suhayl Nasr* for the excellent quality of care that they have provided over the last three years. I am also extremely grateful to have received the same level of commitment towards my wellbeing from *Dr. Eva Ross* who has been enormously influential in the next steps of this journey. Although I met (and sometimes work with) *Neil Lobo* and *Katie Cybulski* at Notre Dame, they really have become two of my most trusted friends in South Bend. I won't go into details here but if they ever read this then they know what I am referring to. There will never be sufficient words or actions to thank them both for the friendship and support they gave me during my darkest hours but I hope they realize how grateful I am for their love and kindness, and know that this is reciprocated. Also, Neil, I have still not forgotten how you tenderly applied topical antibiotics to an area that shall not be mentioned after I contracted a massive *Staph* infection in the Solomon Islands, and couldn't reach the wound unassisted ☺.

None of the research in this dissertation would have been possible without the dedicated efforts from individuals in the field. In the Solomon Islands, *Bob Cooper* (Australian Army Research Institute), *Danyal Odabasi* (Swiss Tropical Institute; University of Basel), my entomology team in Jack Harbour (*Fresher Bron*, *Jiniva Elwin*, *Kezimah Fosal*, *Esther JackNevol*, and *Elwin Josiah*), and the Vector Borne Disease Programme, were instrumental in facilitating the *An. farauti* collections on Kolombangara. Likewise, my collaborators at the Uganda Virus Research Institute in Entebbe (*Jonathan Kayondo*, *Martin Lukindu*, *Fred Ssenfuka*) who ensured that sufficient *An. gambiae* sampling was conducted across the Lake

Victoria basin. In the lab, *Marcia Kern* was extremely kind—and patient—in donating her time, expertise, and experience in assisting me with the *An. gambiae* genomic DNA extractions and 2La karyotyping assays. *Julie Niedbalski*, as always, was there when I needed her. Thank you, both, for your monumental efforts to address whatever strange request I may have asked of you. Additionally, I must also thank *Meredith Doellman*, *Jeff Feder*, and the *Notre Dame Genomics and Bioinformatics Core* who assisted with the RAD library construction. *Matt Sisk* at the University of Notre Dame, Center for Digital Scholarship has been incredibly generous with his time and humor over the last couple of years generating the maps for this dissertation.

In computer world, there were several individuals that took the time to provide me with assistance as I wrangled Python scripting. I thank them for that time and words of encouragement when certain scripts relentlessly returned error messages. These people taught me that even though errors are, ultimately, human, the most experienced programmers are not immune to receiving them, and for that piece of wisdom, I am extremely grateful. Thank you, *Lauren Assour*, *Scott Szakonyi*, and *Caleb Reinking*. I must also thank the developers at VectorBase in particular *Scott Emrich*, *Joe Sarro*, *Dan Lawson*, and *Bob McCallum* for the opportunity to join their group while I explored the merits of analyzing data through the Galaxy platform. *Greg Ragland* at the University of Colorado, Denver wrote the *sampToSam* Perl script that adds read groups to RADseq data. I'm sure I speak on behalf of the many who were conducting RADseq research at Notre Dame during that time when I say this simple little script saved a lot of hours, effort, and sanity for the rest of us.

*Olivia Choudhury* of the Notre Dame Bioinformatics Lab parallelized the *An. farauti* genomic workflows as part of her doctoral studies generating SNP datasets in a rather satisfying few days instead of weeks.

I owe a gigantic debt of thanks, which I am sure I will never be able to repay (there is only so much parasitological and entomological advice I can offer) to my wingman, *Christina Bergey*. Aside from the brief interview she endured for Nora's post-doc position, I first met Christina at Entebbe International Airport as we prepared to embark on an *Anopheles* entomological sampling trip of the Ssesse Islands, Lake Victoria. I would like to think it was my professional influences that swayed her to join Nora's lab rather than start her own at a different institution but I suspect it was probably the shared mal-de-mer, and subsequent emesis on the Entebbe-Nsadzi route that sealed the deal. Either way I am stoked that she chose us. She has taught me so many R programming tips and tricks, and also explained population genetics (my uncomfot zone) in ways that actually make sense. Plus, we found a mutual love in outdoor pursuits, and share-a-thons (not really but she knows what I mean ☺). I hope that our collaborations may long continue and we shall make some awesome discoveries together in the future (scientific as opposed to personal medical susceptibilities). Although I previously thanked *Lauren Assour* in computer world, she really is so much more than a professional Python wrangler. After being my bioinformatics class instructor, I kind of made her my personal one by emailing every so often with programming questions, which she answered and I figured that this was okay so I continued. Out of this "work" relationship, developed a friendship that initially bonded over cats and cocktails but soon advanced into

frivolous whim with a joint attendance session in a Chicago tattoo and piercing studio followed by a runaround and mistaken entry into an Über vehicle, which neither party recognized for several amusing seconds. Lauren, I do not know what I would do without you. It really is that simple.

I have dedicated this dissertation to *my parents and brother* but I also acknowledge them here for their unwavering support, which has been given without question, no matter how strange or dangerous a decision may have appeared i.e. disappearing to Tanzania for three years with no plan and only a back pack.

There really aren't enough words to properly convey how grateful I am to *Shari Hill Sweet* for all of her assistance with my dissertation formatting issues. Thank you from unknown depths for the clarity, support and encouragement that you have shown me during the final few days of my doctoral process.

Finally, the research conducted in this dissertation would not have been possible without funds from the following sources: Bill & Melinda Gates Foundation Grant 45114 (Frank Collins), National Institutes of Health/National Institute for Allergy and Infectious Diseases contract HHSN272201400029C (Frank Collins); Department of Biological Sciences Dissertation Completion Fellowship (Rachel Wiltshire), and Eck Institute for Global Health (Rachel Wiltshire).

## LIST OF ABBREVIATIONS

ACT	Artemisinin-Combination Therapy
AHA	Acute Hemolytic Anemia
AL	Artemether-lumefantrine
AQ	Amodiaquine
AS	Artesunate
bp	Base pairs
CO <sub>2</sub>	Carbon dioxide
CQ	Chloroquine
DDT	Dichloro-diphenyl-trichloroethane
DNA	Deoxyribonucleic acid
F <sub>ST</sub>	Fixation index
GM	Genetically modified
G6PD	Glucose-6-phosphate dehydrogenase
HLC	Human landing catch
IPT	Intermittent Preventive Therapy
IRS	Indoor Residual Spraying
ITN	Insecticide-Treated Net
LLIN	Long-Lasting Insecticidal Net



malERA	Malaria Eradication Research Agenda
MTC	Malaria Transmission Consortium
mtDNA	Mitochondrial DNA
N <sub>e</sub>	Effective population size
nt	Nucleotide
OC	Organochlorine
OP	Organophosphate
PCA	Principal Components Analysis
PCR	Polymerase Chain Reaction
PQ	Primaquine
PQT-VC	Prequalification Team for Vector Control Products
RADseq	Restriction site-associated DNA sequencing
RBM	Roll Back Malaria
rDNA	Ribosomal DNA
RDT	Rapid Diagnostic Test
SMC	Seasonal Malaria Chemoprevention
SNP	Single Nucleotide Polymorphism
SP	Sulfadoxine-pyrimethamine
VOC	Volatile Organic Compound
WHO	World Health Organization
16AnG	16 <i>Anopheles</i> Genomes

## CHAPTER 1:

### MALARIA: A MOST PERSISTENT PARASITIC PROBLEM

#### 1.1 Biology, transmission, disease

Contrary to popular belief—that malaria is transmitted by mosquitoes—malaria is actually the status of being infected with apicomplexan parasites of the genus *Plasmodium*, which are transmitted to humans by the bite of anopheline mosquitoes. It is against both of these organisms that control interventions are targeted.

Malaria is a complex disease. The five *Plasmodium* species (spp.), which infect humans (*Homo sapiens*) have demonstrated extraordinary ability in manipulating their mosquito hosts to maximize transmission i.e. *P. falciparum*-infected *Anopheles gambiae sensu lato* mosquitoes demonstrated larger, and more frequent, blood meals than uninfected, thus allowing the parasite to be transmitted more rapidly among human hosts (Koella, Sørensen, and Anderson 1998). Additionally, there are approximately 41 dominant species of *Anopheles* mosquitoes (Sinka et al. 2012) with the vectorial competence (intrinsic ability to support successful parasite replication and transmission) to transmit *Plasmodium* species, each with discrete bionomic behaviors that allow them to fully exploit their

ecological niche. Transmission is further complicated by the differing epidemiological settings that are generated as a result of endemicity—the quantification of malarial risk (Hay, Smith, and Snow 2008). This measurement is reliant on human-vector interactions, which themselves can be variable and complex i.e. the ability of *Anopheles gambiae* to tolerate aridity through changing chromosomal inversion frequencies in seasonal patterns (Coluzzi et al. 1979) and latitudinal clines (Coluzzi et al. 1979; Cheng et al. 2011). It follows, therefore, that malaria programs must consider, and understand these complexities in the planning and implementation of their control strategies.

## 1.2 A numbers game

Malaria is an ancient malady that has plagued humans for centuries. (P. Russell 1955; Hoeppli 1959; Bruce-Chwatt 1965; Carter and Mendis 2002). Estimates have been made to assess its impact on humanity with six billion deaths being attributed to malaria-related mortalities in the last century alone (Carter and Mendis 2002). Today, malaria remains a major public health problem—especially in sub-Saharan Africa—with the World Health Organization (WHO) estimating that 219 million cases of malaria and 435,000 deaths occurred in 2017 (World Health Organization 2018d). Infants under the age of five years are the most vulnerable individuals, and mortalities in this group (266,000) account for nearly two thirds (61%) of the total malaria deaths (World Health Organization 2018d), supporting the assertion that a child dies of the disease every two minutes (World Health Organization 2017e). However, mortality and morbidity figures have reduced

significantly over the past decade due to dedicated investment in the development, and expansion, of control interventions such as long-lasting insecticidal nets (LLINs), indoor residual spraying (IRS), artemisinin combination therapies (ACTs), and health education, and these hard-won gains must be sustained through the adoption of effective prevention and control management strategies that are appropriate to the malaria situation experienced by each country.

Elimination (defined as reporting zero indigenous cases for three consecutive years) is possible—as demonstrated most recently by Kyrgyzstan (2016), Sri Lanka (2016), and Paraguay (2018)—and as the number of countries reporting fewer than 10,000 malaria cases has increased, the prospects of global elimination appear encouraging. However, it should be noted that none of these countries shoulder the burden like sub-Saharan Africa, who contributed an estimated 49% of global malaria-related mortality in 2017 (World Health Organization 2018d).

### 1.3 Disease management: present day

#### **1.3.1 Asymptomatic infections**

Malaria as a disease can be defined as infection with *Plasmodium* parasites—a condition known as *parasitemia*. Not all infections, however, produce clinical sequelae but these asymptomatic individuals are important in the community since their parasitemia acts as a reservoir of infection that enables malaria vectors to

maintain transmission. It is crucial, therefore, that these carriers are identified, and treated, especially in regions where elimination is the target outcome.

### **1.3.2 Clinical malaria**

Symptomatic infections are managed according to the severity of clinical sequelae. *Uncomplicated malaria* is characterized by classical fever, which cycles in tertian (2-day) or quartan (3-day) patterns according to the *Plasmodium* species that caused the infection. If left untreated, hepato- and splenomegaly develop alongside anemia (due to schizont rupture and erythrocytic destruction) but the natural infection resolves itself after several weeks as long as the individual is not immunocompromised. Response to treatment is usually rapid with symptom resolution occurring within 3-days (N. White 2009). If the infection is caused by *P. vivax* or *P. ovale*, then an opportunity for recrudescence presents itself due to the parasites' ability to remain latent as hypnozoite form in hepatocytes. Relapses are unpredictable and can occur months or even years after initial infection. These are, therefore, another important source of infection to malaria vectors since hypnozoites are parasite reservoirs.

*Complicated, or severe malaria* is almost always caused by *P. falciparum*, which the literature also refers to as falciparum malaria. Major complications of severe malaria include cerebral malaria, pulmonary edema, acute renal failure, severe anemia, hemorrhaging, metabolic acidosis, and hypoglycemia. Any of these symptoms can rapidly develop together or in succession, leaving the patient at risk of death within the space of a few hours (Trampuz et al. 2003). Cerebral malaria is

the most prominent feature of severe falciparum malaria. The parasite's unique ability to express surface antigens, such as *Plasmodium falciparum* Erythrocyte Membrane Protein 1 (*PfEMP1*) on host *P. falciparum*-infected erythrocytes (Baruch et al. 1995), enables cytoadherence in the cerebral microvasculature causing obstruction of blood flow, which likely leads to hypoxia, cerebral edema, and raised intracranial pressure resulting in impaired consciousness, long-term neurological deficits, coma and death. Untreated, cerebral malaria is nearly always fatal.

### **1.3.3 Strategy**

A renewed global commitment to malaria control was adopted at the beginning of the millennium after the previous two decades (1970-1998) had witnessed a deteriorating situation in which the number of annual malaria-related deaths had tripled to an estimated 1 million (World Health Organization 2010). Reasons for the lack of progress in research and control approaches that possibly contributed to these numbers include fragmented efforts that often undermined each other as a result of incoherency between development partners and malaria programs (Nabarro and Tayler 1998). A revised strategy—the Roll Back Malaria (RBM) initiative—was developed with the principal objective of strengthening national health services through the integration of malaria-related activities across the sector (World Health Organization 1993). The global strategy prioritized a reduction in morbidity and mortality rates through the adoption of four key technical components, which sought to:

1. Provide early diagnosis and prompt treatment
2. Plan and implement selective and sustainable preventive measures, including vector control
3. Prevent, or detect and contain epidemics early
4. Strengthen local research capacities to permit and promote regular national assessment of a malaria situation in particular its ecological, social, and economic determinants

Tangible gains were, and continue to be, realized as evidenced by a decline in the estimated numbers of annual malaria cases (262 million-219 million) and deaths (839,000-435,000) between 2000-2017 (World Health Organization 2015c, 2018d). Additionally, intangible successes were quantified with an estimated 663 million clinical cases *averted* through the increased coverage of malaria control interventions (Bhatt et al. 2015) further endorsing the RBM strategy.

### **1.3.4 Prevention**

#### *1.3.4.1 Intermittent preventive therapy (IPT)*

Antimalarial chemoprophylaxis is administered in endemic areas of sub-Saharan Africa with sulfadoxine-pyrimethamine (SP) to suppress erythrocytic stages thus, protecting the most vulnerable individuals from the pathophysiological effects of *P. falciparum* infection. However, in spite of evidence that supports a reduction in the severity of clinical sequelae (Conteh et al. 2010; Kayentao et al. 2013) together with endorsement as a standard anti-malarial health practice (World Health Organization 2017d, 2018b), adoption of IPT-SP for pregnant women

(IPTp) (Yaya et al. 2018) and infants (IPTi) (World Health Organization 2017c) has been disappointing.

#### *1.3.4.2 Seasonal malaria chemoprevention (SMC)*

SMC is a specialized program targeted at infants aged between 3-59 months who are exposed to the high seasonal transmission rates encountered in the ecoclimatic Sahel region of northern Africa. Implementation of therapy with SP + amodiaquine (AQ) commences at the beginning of the transmission season and continues intermittently throughout—up to a maximum of four doses—on the basis that both antimalarial drugs retain sufficient efficacy against *P. falciparum* (World Health Organization 2013). SMC was scaled-up in 2016-2017 reaching an estimated 15.7 million infants from a target population of approximately 29.3 million. Financial constraints was provided as an explanation for lack of coverage (13.6 million) (World Health Organization 2018d).

#### *1.3.4.3 Vector control (adults)*

By far the most extensively implemented and effective preventive strategy—as illustrated by the impact on the reduction of *P. falciparum* infections in Africa (Bhatt et al. 2015)—is control of the mosquito vector. A significant proportion of donor funding has been, and continues to be, allocated to the procurement of insecticides for indoor residual spraying (IRS) operations and the provision of long-lasting insecticidal nets (LLINs), which is indicated by the percentage of budget expenditure e.g. 39% by the Global Fund to Fight AIDS, Tuberculosis and Malaria



(2009) and 59% of the President's Malaria Initiative (2010) (Roll Back Malaria Partnership 2010). Improving coverage and use of these interventions (due to their high efficacy, relatively low cost, operational simplicity, and scalability) has been a key target of malaria control programs over the last 10-15 years, and this strategy has successfully contributed to 80% of the reduction in *P. falciparum* transmission across Africa but it has also intensified, and extended, the rate of pyrethroid resistance now present in malaria vector populations. This is a worrying situation that must be managed judiciously so as not to deteriorate it further since there is only a limited selection of approved insecticide classes currently available for use with IRS and LLINs (World Health Organization 2012; The Global Fund 2018; World Health Organization 2018a). To preserve the use of pyrethroids in LLINs, malaria programs with IRS operations have gradually switched to pirimiphos-methyl formulations (Actellic™, Syngenta) in areas where vectors are susceptible to organophosphate (OP) insecticides (World Health Organization 2012; Oxborough 2016).

#### *1.3.4.4 Supplemental techniques (juvenile stages)*

*Larval control* (microbial toxins, insecticide, predatory fish) and *environmental management* (drainage) target the mosquito's oviposition sites and are recommended as supplementary to LLINs or IRS for implementation where larval habitats are few in number, fixed in location and easily accessible under specific transmission settings appropriate to the local situation in concert with

current interventions (i.e. IRS, LLINs) (Fillinger et al. 2009; World Health Organization 2015a).

#### *1.3.4.5 Communication and behavioral impact*

The RBM Global Strategy aimed to halve malaria-associated mortality by 2010, and again by 2015 (World Health Organization 1993). Achieving those targets required a critical interplay between the malaria program and the individuals that they served. Communication is a key intervention, whose purpose is to effectively disseminate information that increases knowledge, and motivates households to be proactive in their own malaria prevention and treatment thus, reducing the risk of clinical disease. However, strategies must be well-planned and appropriate to local settings, including the involvement of community leaders. They should also be sustainable to ensure the continued acceptance of malaria interventions and services so that communities are playing an active role in protecting their own health. Process indicators (such as attitude to, and correct use of, interventions) should be assessed regularly to determine the effectiveness of the message, and ensure that the strategy has not deviated from the program (Roll Back Malaria Partnership 2008).

#### **1.3.5 Treatment**

It is recommended that all suspected malaria cases be confirmed with a *parasitological diagnosis* (light microscopy or rapid diagnostic test (RDT)) prior to treatment (except infants under 5 years in areas of high, stable transmission who

should be treated as if clinically diagnosed due to the high probability of their fever being malarial), which should be dispensed within 24-hours of the onset of symptoms (World Health Organization 2015b).

A 3-day treatment of *artemisinin-based combination therapies* (ACTs) is the currently recommended practice for uncomplicated *P. falciparum* infections in adults and infants.

Artemisinin is the active principle compound of *Artemisia annua* (Qinghao)—a common herbaceous plant native to temperate China—whose antimalarial properties derive from its antioxidant activity, which rapidly attacks both asexual and sexual parasite stages. Combination therapy includes an artemisinin derivative i.e. artemether (AM), artesunate (AS), dihydroartemisinin (DHA) together with a longer-acting partner drug, whose inclusion is three-fold: (i) it clears any remaining parasites from the blood, (ii) it prevents the development of resistance to artemisinin, which would most certainly occur if it was administered as a monotherapy, and (iii) it provides a degree of prophylaxis due to its longer half-life (Aweeka and German 2008). Currently recommended ACTs for uncomplicated *P. falciparum* infections are: AM + lumefantrine (AL), AS + amodiaquine (AQ), AS + mefloquine (MQ), AS + SP, and DHA + piperaquine (PPQ) (World Health Organization 2015b).

Pregnant women in their first trimester should be treated for 7-days with quinine (QN) + clindamycin due to concerns for the potential teratogenic effects of ACT on organogenesis; however, no adverse effects on the mother or fetus have

been reported during the second or third trimesters, and it is recommended that ACTs should be administered during this period.

Blood stage *P. vivax*, *P. ovale*, *P. malariae*, and *P. knowlesi* infections should be treated as for uncomplicated *P. falciparum* malaria if the species is not known with certainty. Chloroquine (CQ) can also be administered in areas that remain sensitive. Radical cure of *P. vivax* and *P. ovale* hypnozoites (whose latency and unpredictable recrudescence act as a parasite reservoir) is treated with primaquine (PQ). Its administration, however, must be closely supervised for potentially adverse hematological effects especially in individuals that carry genetic variants, which predispose them to acute hemolytic anemia (AHA).

Severe falciparum malaria is treated with parenteral AS for a minimum of 24-hours until oral therapy can be tolerated. Treatment is completed with 3-days of ACT and clinical symptoms management. This recommendation applies to adults, infants, pregnancy in all trimesters, and lactating women.

Individuals categorized as special risk i.e. first trimester of pregnancy, lactating women, infants < 5 kg body weight, patients with human immunodeficiency virus (HIV) co-infections, non-immune travelers, or those demonstrating hyperparasitemia have altered pharmacokinetics that prevent them from receiving optimal doses of recommended antimalarial treatments. The rate of treatment failure in these groups is substantially higher and these individuals should be managed appropriate to their status (World Health Organization 2015b).

## 1.4 Vector control

Previous malaria control programs operated on the basis of reducing morbidity and mortality yet the disease remains a major global public health issue. The last decade has seen a move away from this archetypical strategy with a motivation towards interruption of transmission, elimination, and eventually, eradication (Alonso et al. 2011). In order to achieve this outcome, the basic reproductive rate— $R_0$  (the number of secondary malaria infections arising from the same single infection)—must be reduced to less than one ( $R_0 < 1$ ), and research has shown that this is unlikely to be attained with current vector control methods in regions with the highest malaria burden i.e. sub-Saharan Africa, which experiences intense, perennial transmission (Shaukat, Breman, and McKenzie 2010; The malERA Consultative Group on Vector Control 2011).

The Malaria Eradication Research Agenda (malERA) Consultative Group on Vector Control (2011) identified three development challenges considered critical to achieve interruption of transmission:

1. Preservation, and improvement of the utility of existing insecticide-based interventions
2. Interventions that affect vector species not effectively targeted by current tools
3. Innovative approaches that will permanently reduce the very high vectorial capacities of the dominant malaria vectors in sub-Saharan Africa

Without compromising the progress and achievements gained by the RBM program, the paradigm shift that has been introduced to the global strategy through these development challenges, focuses on improving current vector control

interventions while supplementing them with novel and innovative approaches. It is hoped that both strategies in parallel will have an integrated effect on disease transmission reducing it below the threshold at which it is sustained i.e.  $R_0 < 1$ .

#### **1.4.1 Preservation and improvement of the utility of existing insecticide-based interventions**

Of the three challenges, the first is considered to be the most urgent since resistance mechanisms against the four main insecticide classes licensed for public health use (pyrethroids, carbamates, organophosphates (OPs), and organochlorines (OCs)) are now commonplace as a result of their universal application in vector control operations (Oxborough 2016; Ranson and Lissenden 2016).

The two most frequently identified forms of resistance are *target site insensitivity*, where mutations in neural pathways have evolved that reduce binding of the insecticide, and *metabolic* in which quantitative and qualitative modifications of key enzymes detoxify the insecticide before it can exert a lethal effect upon the vector. *Cuticular* resistance (Balabanidou et al. 2016) and *behavioral* adaptations (Gatton et al. 2013; Carrasco et al. 2019) have also been reported but their impact on malaria transmission rates in response to vector control interventions has not been quantified.

##### *1.4.1.1 Target site insensitivity*

Examples of target site insensitivity are: (i) *kdr* (knock-down resistance), a mutation in the voltage-gated sodium channel gene, which is the target of

pyrethroids and dichloro-diphenyl-trichloroethane (DDT) (Ranson et al. 2000), (ii) AChE (acetylcholinesterase), an enzyme that hydrolyzes the neurotransmitter, acetylcholine, which terminates nerve impulses and is the target site of OPs and carbamates (Ayad and Georgiou 1979), and (iii) Rdl (resistance to dieldrin), a mutation in one of the five  $\gamma$ -butyric acid (GABA) receptor subunits (also a neurotransmitter) that is the target of the cyclodiene insecticides i.e. OCs such as chlordane and dieldrin (Du et al. 2005).

#### *1.4.1.2 Metabolic resistance*

Metabolic resistance is based on enzyme systems that insects apply to detoxify naturally occurring exogenous compounds by over-expression or conformational modification of enzymes that metabolize insecticides. Over-expression of detoxification enzymes is the most common cause of metabolic resistance in mosquitoes and occurs through gene duplication (Bass and Field 2011) or as changes in the regulatory elements (Ingham et al. 2017) or promotor region (Ding et al. 2005) of the gene. Increased enzymatic production allows rapid degradation of the insecticide reducing the effective dose before toxicity is exerted upon the insect. Three major enzyme families have been identified as conferring resistance to insecticides in malaria vectors but are less well characterized than target site mutations due to the complexity of their interactions in metabolic pathways.

Carboxylesterases, which hydrolyze ester bonds, have mostly been associated with resistance to OPs where conformational changes in the enzymes of

resistant *Anopheles* mosquitoes have been identified as notably increasing malathion hydrolysis in comparison to susceptible individuals (Hemingway 1982, 1983).

Cytochrome P450-dependent monooxygenases (P450s) are a diverse superfamily of enzymes, which as hemeproteins can oxidize xenobiotic substrates producing water-soluble molecules for elimination (Mathews, van Holde, and Ahern 2000). P450s have been consistently associated with pyrethroid resistance and several genes from pyrethroid-resistant phenotypes have been functionally characterized as a result of over-expression i.e. *CYP6P3* (Müller et al. 2008), *CYP6P9* (Amenya et al. 2008).

The final group—glutathione S-transferases (GSTs)—are a cytosolic superprotein family, which exist in virtually all living organisms. GSTs are categorized into classes according to the composition of amino acid residues found in the N-terminal domain of their highly conserved G binding sites. Delta and epsilon are the two insect-specific classes, and have extensive representation in the major African malaria vector *Anopheles gambiae*, where 5/8 of the epsilon GSTs were demonstrated to be overexpressed in a DDT-resistant phenotype (Ding et al. 2003).

In addition to these mechanisms, *cross-resistance* can also occur through shared mutations in target sites (Ranson et al. 2000; Essandoh, Yawson, and Weetman 2013) and metabolic detoxification pathways (Edi et al. 2014) thus, the requirement to develop and approve insecticides with differing modes of actions is now vital. Novel insecticidal compounds recently sanctioned for public health use by the WHO Prequalification Team-Vector Control Products (PQT-VC) include



chlorfenapyr (a *pyrrole* that disrupts mitochondrial respiration pathways) (World Health Organization 2017b) and clothianidin (a *neonicotinoid* that acts as a post-synaptic nicotine acetylcholine receptor antagonist) (World Health Organization 2017a) but continued research is a priority.

#### **1.4.2 Interventions that affect vector species not effectively targeted by current tools**

The second challenge requires high quality entomological surveillance data that are specific to the local area as this will inform on the most suitable strategy and intervention with which to interrupt transmission. In the first instance, if there are current control interventions in operation—and transmission rates are high—then the reason for the failure must first be established. If this is determined to be ineffective targeting of the vector then local bionomic studies must commence. Likewise, if there are currently no vector control interventions in place. Longitudinal vector surveys (larval and adult) are fundamental to identifying behavioral patterns and incrimination if there are multiple mosquito species present. Once a vector and its behaviors have been determined an intervention that targets those behaviors can be developed, tested, and implemented (if effective) i.e. barrier screens exploited the exophilic resting behavior of blood-fed and host-seeking *Anopheles* mosquitoes, and although this was developed as a sampling method (T. Burkot et al. 2013), an application of insecticide to the netting material could transform this technique into a lethal intervention.

### **1.4.3 Innovative approaches that will permanently reduce the very high vectorial capacities of the dominant malaria vectors in sub-Saharan Africa**

The third, and final, challenge is the most adventurous. A permanent reduction of vectorial capacity—defined as the “the average number of inoculations with a specified parasite, originating from one case of malaria in unit time, that the population would distribute to man if all the vector females biting the case became infected” (Garrett-Jones 1964)—is necessary for local malaria elimination in sub-Saharan Africa, since residual vectorial capacities of mosquito populations will be capable of perpetuating epidemics if *Plasmodium* parasites are reintroduced to a human population, which has lost partial immunity (The malERA Consultative Group on Vector Control 2011). The majority of research in this area has been focused on genetic control programs—in particular the manipulation of mosquito vectors—that will yield long-term sustained population reduction through either suppression or replacement. However, the required investment in this technology is extensive and the development of other novel approaches must also be encouraged.

### **1.5 Integration of current trial products into future programs**

Vaccine development has been a challenging, yet active, area of research producing a number of potential candidates but only one—RTS,S/AS01 (Mosquirix™) (GSK) (The PATH Malaria Vaccine Initiative 2016)—has reached the required protective efficacy and safety standards in clinical trials. RTS,S/AS01 is a recombinant protein vaccine, which targets the circumsporozoite protein expressed by *P. falciparum* at the pre-erythrocytic stage of infection to induce an immune

response that elicits a protective effect against the development of severe clinical disease.

Between the period 2009-2014, infants in two cohorts—aged 6-12 weeks (n=6,537) and 5-17 months (n=8,922)—received four doses of vaccine as part of a phase 3 clinical trial conducted in seven African countries with differing malaria transmission intensities. In older children, the vaccine prevented approximately 39% of malaria cases and 29% of severe disease (RTS,S Clinical Trials Partnership 2015). After review of these results by independent advisory groups, RTS,S/AS01 was recommended by the WHO as part of a pilot immunization program for young children in selected areas of Ghana, Kenya, and Malawi. Vaccinations are to be assessed over a three-year period (2019-2022) to determine their value as a complementary chemoprophylactic of severe disease in vulnerable groups, which is intended to supplement (not replace) the current core preventive, diagnostic, and treatment measures (World Health Organization 2018c) Beyond this pilot program, however, there is no policy recommendation for the large-scale use of the RTS,S/AS01 vaccine.

#### 1.6 An extended research agenda

The purpose of these dissertation studies is to conduct research that contributes expertise to the extended research agenda (Alonso et al. 2011); in particular, the second and third developmental challenges identified as critical for the maintained interruption of malaria transmission by the malERA Consultative Group on Vector Control (2011).

Chapter two considers the genetic structure and differentiation in natural populations of the malaria vector *Anopheles gambiae* in a micro-geographic setting. It narrows an information gap related to the third challenge once a permanent reduction strategy is developed as it informs on: (a) whether these populations would be appropriate for field testing entomological efficacy of genetically modified (GM) malaria vectors, and (b) understanding how heritable strategies would behave in, and move through, a targeted population.

Chapters three and four present research that can be related to the second challenge (interventions that affect vector species not effectively targeted by current tools) since they address different biological attributes of *Anopheles farauti*, the principal malaria vector in the southwest Pacific. In certain areas of the region, *An. farauti* exhibits exophagic and exophilic behaviors meaning it demonstrates a propensity for outdoors feeding and resting, respectively, thus avoiding contact with a lethal dose of insecticide from existing indoors-based control tools. While the application of IRS and distribution of LLINs has been successful at reducing annual parasite incidences by killing the proportion of populations that enter domiciles, the mosquitoes that continue to feed outdoors maintain residual transmission, necessitating the development of a novel approach for their control.

Vectorial capacity is underpinned by genetic components that contribute to key biological traits, such as behavior and susceptibility to *Plasmodium* infection, which vary between vectors. Understanding how these factors interact at the genomic level can provide greater insight into how they may be manipulated as targets of control. A genome-wide approach in the analysis of population structure

also provides power to identify genetic processes (such as restricted gene flow or an insecticide-resistant genotype) and the genomic regions that regulate them thus, creating an understanding of how these mechanisms will impact upon control interventions. Chapter three generates SNP-based data that lays the foundation for future studies, which can be leveraged to develop innovative approaches towards targeted interventions.

Finally, chapter four explores the credibility of sugar-fermented yeast as an organic source of carbon dioxide (CO<sub>2</sub>) trap attractant to host-seeking and resting *An. farauti* mosquitoes, which may be contributing to residual transmission in Western Province, Solomon Islands.

## 1.7 Going global

The dissertation is concluded in Chapter five with a brief summary of the investigations conducted, an explanation as to how they have contributed expertise to the research agenda and, finally, their applicability to malaria vector control approaches.

## CHAPTER 2:

### DETERMINING THE POPULATION GENETIC STRUCTURE OF *ANOPHELES GAMBIAE* IN THE NORTHWESTERN LAKE VICTORIA BASIN BY REDUCED-REPRESENTATION SEQUENCING

This chapter appears in part in the published manuscript “Reduced-representation sequencing identifies small effective population sizes of *Anopheles gambiae* in the north-western Lake Victoria basin, Uganda” by Wiltshire et al. (2018) with modifications. Its purpose is to introduce the reader to research that contributes expertise towards the third developmental challenge identified by the malERA Consultative Group on Vector Control: innovative approaches that will permanently reduce the very high vectorial capacities of the dominant malaria vectors in sub-Saharan Africa.

Given that the favored vector-targeted interventions—imagociding through LLINs and IRS— are reaching the limits of their effectiveness, advanced techniques that complement current control strategies are being explored for the purpose of attaining  $R_0 < 1$  and the interruption of transmission. There are many technical approaches by which mosquitoes can be genetically modified (GM) but they are all designed to realize one of two outcomes for the target population into which they

will be released: (i) *suppression*, which reduces the number of competent vectors, and (ii) *replacement* where vectorial capacity is diminished. Once successful genetic modifications have been demonstrated in the laboratory environment they must then be translated to field applications through a series of phased testing pathways if they are to be endorsed by national authorities as a malaria control intervention.

This chapter focuses on a genome-wide population genetic assessment of the principal African malaria vector, *An. gambiae* Giles, 1902 (hereafter, *An. gambiae*) using single nucleotide polymorphism (SNP) markers in the northwestern Lake Victoria basin for the purpose of identifying and characterizing natural populations that may be appropriate for small-scale ecologically-confined releases of GM mosquitoes.

Three island populations (Bukasa, Nsadzi, Sserinya) demonstrated low to moderate genetic differentiation with small effective sizes ( $N_e$ ) implying limited migration and susceptibility to genetic drift. In conclusion, this chapter identifies candidate field sites for small-scale evaluation of entomological efficacy of GM *An. gambiae* mosquitoes.

## 2.1 Abstract

Malaria is the leading cause of global pediatric mortality in children below five years of age. The number of fatalities has reduced significantly due to an expansion of control interventions but the development of novel technologies remains essential in the effort to pursue elimination and eradication. Recent attention has been focused on the release of GM mosquitoes into natural vector

populations as a mechanism of interrupting parasite transmission but despite successful *in vivo* laboratory studies, a detailed population genetic assessment—which must first precede any proposed field trial—has yet to be undertaken systematically. In this chapter, the genetic structure of *An. gambiae* populations in the northwestern Lake Victoria basin is explored to assess whether their characteristics are suitable for testing entomological efficacy of GM mosquitoes in a pilot field release. 488 *An. gambiae* mosquitoes were collected from six locations and a subset (N=96) was selected for restriction site-associated DNA sequencing (RADseq). The resulting SNP marker set was analyzed for effective size ( $N_e$ ), connectivity, and population structure (PCA,  $F_{ST}$ ). 5,175 high-quality genome-wide SNPs were identified. A principal components analysis (PCA) of the collinear genomic regions illustrated that individuals clustered in concordance with geographic origin and with some overlap between sites. Genetic differentiation between populations was varied with inter-island comparisons having the highest values (median  $F_{ST}$ : 0.0480-0.0846).  $N_e$  estimates were generally small (124.2-1920.3). Island populations demonstrated low to moderate genetic differentiation and greater structure suggesting some limitation to migration. The smaller estimates of  $N_e$  indicate that an introduced effector transgene will be more susceptible to genetic drift but to ensure that it is driven to fixation (rather than loss) a robust gene drive mechanism will likely be needed. These findings, together with their favorable geography and suitability for frequent monitoring, demonstrate that the Ssese Islands contain several locations where the *An. gambiae* populations merit further evaluation as candidates for a GM mosquito pilot release.



## 2.2 Introduction

### 2.2.1 Global burden

The World Health Organization (WHO) estimated that 219 million cases of malaria occurred globally in 2017 resulting in approximately 435,000 deaths (World Health Organization 2018d). Infants under the age of five years are the most vulnerable individuals since malaria is the leading cause of pediatric mortality in that age group.

### 2.2.2 Insecticide-based vector control tools

Although existing malaria control interventions such as: (i) the increased distribution of LLINs, (ii) consistent applications of IRS, and (iii) improved access to RDTs and ACTs have significantly reduced the number of cases and deaths in the last decade (World Health Organization 2018d), the development of innovative mosquito vector (The malERA Consultative Group on Vector Control 2011) and *Plasmodium* parasite control technologies (The malERA Consultative Group on Drugs 2011; The malERA Consultative Group on Vaccines 2011) will be required to reduce incidence rates below the critical threshold that sustains transmission if the ultimate objective is elimination, a target difficult or impossible to achieve using traditional control tools in regions with *intense malaria transmission* (Shaukat, Breman, and McKenzie 2010; Alonso et al. 2011).

Increasing coverage of the current vector control methods (LLINs and IRS) has contributed greatly to the reduction in morbidity and mortality; however, these interventions are only effective against populations of indoor-feeding and/or

resting adult female mosquitoes, which remain susceptible to exposure with the four major insecticide classes currently implemented in malaria control programs: pyrethroids in LLINs (and limited IRS applications), and carbamates, organophosphates, and organochlorines for IRS in high transmission areas where the efficacy of pyrethroids must be preserved for LLIN use. Over-reliance on these insecticides over many decades has now resulted in resistance, which has become established across the African continent (Oxborough 2016; Ranson and Lissenden 2016) necessitating the development of new classes as a matter of urgency. Much research in this area is being undertaken, and the recent approval of chlorfenapyr (BASF SE) and clothianidin (Sumitomo Chemical Co. Ltd.) by the WHO PQT-VC (World Health Organization 2017b; The Global Fund 2018) is encouraging since both exhibit different modes of action to those currently in use meaning that cross-resistance is unlikely, especially if applied in combination as recommended by the Global Plan for Insecticide Resistance Management (World Health Organization 2012). While the development of new insecticides improves the utility of existing indoors-based vector control tools, this strategy on its own is insufficient to reduce  $R_0 < 1$  because it fails to effectively target the mosquito vectors that do not enter domiciles, which are responsible for maintaining residual transmission.

### **2.2.3 GM mosquitoes as a vector control solution**

Innovative vector control strategies (such as genetic modification (Curtis 1968)) were first considered in response to failure of the Global Malaria Eradication Program (1955-1969) whose early successes were negatively impacted as a result

of insecticide resistance to DDT and the accumulation of residues within the environment (it has been argued, however, that indiscriminate agricultural operations were largely responsible for both) (Busvine and Pal 1969; Busvine 1978). Although challenging, and technically complex to construct, genetically modified (GM) mosquitoes as an alternative form of vector control has increasingly gained attention in recent years in parallel with the advancement of next-generation sequencing efforts of the major *Anopheles* malaria vector genomes (Holt et al. 2002; Lawniczak et al. 2010; Neafsey et al. 2015).

The principal objective of generating GM *Anopheles* mosquitoes is to decrease their vectorial capacity to transmit *Plasmodium* parasites—through population suppression or replacement (Burt 2014)—by rendering them refractory to infection (Christophides 2005; S. Wang and Jacobs-Lorena 2013) and examples of successful *Anopheles* genetic constructs and drive systems have been demonstrated (Bian et al. 2013; Gantz et al. 2015; Hammond et al. 2016). If GM *Anopheles* mosquitoes are to be established as part of a malaria control intervention then these novel achievements must be successfully translated from bench to field. The first step in realizing this strategy is to obtain a detailed understanding of the genetic structure of the natural populations into which the transgenic construct and gene drive system will be introduced as identifying levels of gene flow (genetic exchange), and the effective population size ( $N_e$ ) will be critical in predicting the dispersal and maintenance of a transgene.

#### **2.2.4 *An. gambiae* species complex**

In sub-Saharan Africa, *An. gambiae* is an important mosquito vector of the *Plasmodium* malaria parasite species that infect humans. It exists as one member of a cryptic species complex describing the taxonomic grouping of Afrotropical mosquitoes that were first identified as the species *An. gambiae* Giles, 1902. During the early 1960s, behavioral differences first alerted researchers to the notion that, despite appearing morphologically identical, these different forms were probably multiple species that co-existed sympatrically: a fact confirmed through a series of genetic backcrosses, which yielded universally sterile male hybrids that demonstrated the existence of five mating-types and, thus, a species complex (Davidson 1962, 1964). The complex currently comprises eight members with varying vectorial competencies: *An. amharicus* Hunt, Wilkerson & Coetzee, 2013 (non-vector); *An. arabiensis* Patton, 1905 (vector); *An. bwambae* White, 1985 (minor vector due to limited distribution); *An. coluzzii* Coetzee & Wilkerson, 2013 (vector); *An. gambiae* (vector); *An. melus* Theobald, 1903 (vector); *An. merus* Dönitz, 1902 (vector); *An. quadriannulatus* Theobald, 1911 (non-vector) (VectorBase 2017).

#### **2.2.5 *An. gambiae* population structure across the African continent**

*An. gambiae* population structure across the African continent has been extensively studied and was determined to be unexpectedly shallow (Lehmann et al. 1996; Besansky et al. 1997; Lehmann et al. 1999, 2000). Comparison of allozymes (mean  $F_{ST}$ : 0.036) and microsatellites (mean  $F_{ST}$ : 0.016) revealed extensive inter-population gene flow over a 6,000 km distance (Lehmann et al. 1996) that

contrasted sharply with those across the Kenyan Rift Valley Complex (KRVC) (mean microsatellite  $F_{ST}$ : 0.104; mean mitochondrial DNA (mtDNA)  $F_{ST}$ : 0.176), a much shorter distance of 700 km, which was attributed to the KRVC acting as a physical barrier to gene flow (Lehmann et al. 1999, 2000). Oceanic island studies of *An. gambiae* population structure have also demonstrated varying degrees of differentiation that range from considerable genetic exchange in the Bijagós archipelago of Guinea-Bissau ( $F_{ST}$ : 0-0.019) to restricted gene flow between the Comoros Islands ( $F_{ST}$ : 0.093-0.126) (Marsden et al. 2013). Despite the desirable genetic characteristics observed in *An. gambiae* populations of the Comoros, they are not well suited to the frequent monitoring that transgenic field studies require being nearly 1,000 km offshore. A comparably appropriate alternative site would be a lacustrine setting with multiple islands in a malarious region: Lake Victoria.

#### **2.2.6 Previous Lake Victoria *An. gambiae* population genetic studies**

There have been two previous *An. gambiae* population genetic studies in Lake Victoria. Chen et al. (2004) developed six microsatellites from five island and six mainland populations in western Kenya, which showed low but statistically significant genetic structure (mean  $F_{ST}$ : 0.0010-0.019,  $p < 0.001$ ) that also supported a significant correlation between geographic distance and genetic differentiation (Mantel:  $p < 0.001$ ). Kayondo et al. (2005) examined genetic structure in *An. gambiae* populations in the Ssese Islands, the focus of the present study, using microsatellite markers with temporal sampling that also demonstrated low but statistically significant genetic differentiation (mean  $F_{ST}$ : 0.014-0.105,  $p < 0.05$ ). In

contrast with Chen et al. (2004), however, that study found no support for the isolation-by-distance model (Mantel:  $p = 0.787$ ) and concluded that the Ssesse populations varied as a result of: (i) restricted gene flow (due to separation from the mainland by water), (ii) small effective size, and (iii) temporal instability, which combined had provided these mosquitoes with the opportunity to differentiate genetically.

### **2.2.7 Current objective**

The study aimed to follow up the research of Kayondo et al. (2005) by determining the current genetic structure in the same *An. gambiae* populations using recent advances in next-generation sequencing technologies. SNP markers were selected to capture high-density sequence variation due to their: abundance in the *An. gambiae* genome (VectorBase 2017), lower mutation and genotyping error rates, adaptability to high-throughput assays, utility in generating an informative marker panel applicable to future discovery, and population genetic research enquiries. A restriction site-associated DNA sequencing (RADseq) (Miller et al. 2007; Baird et al. 2008) approach was applied as the most economical high-resolution technique to generate a genome-wide SNP marker set for this important malaria vector in the Lake Victoria region.

## 2.3 Materials and methods

### 2.3.1 Study area

The study area is located in the Ssesse Islands, an archipelago in the northwestern division of the Lake Victoria basin, southern Uganda (Figure 2.1). Each of the 84 islands varies in size i.e. the largest—Bugala—is 105 km<sup>2</sup> while some are merely islets of rock, creating a total land coverage area of 454.8 km<sup>2</sup>. The islands share a general topographical characteristic in that they rise as gentle slopes from lake level (1,220 m above sea level (ASL)) to central flat-topped ridges at a maximum elevation of 1,260 m ASL (Kalangala Town, Bugala) (Thomas 1941; Ssegawa and Nkuutu 2006).

The climate is equatorial. There are two wet seasons: a main one from March-May, and a lesser one in November-December but rainfall occurs monthly (mean 140 mm), which is reflected by the highest recorded annual precipitation rates (2,000 mm +) in Uganda (Thomas 1941; National Environment Management Authority (NEMA) 2005; Ssegawa and Nkuutu 2006). Annual temperatures range from 18.3°C (February) to 27.2°C (August) with relative humidity being lowest in February (68%) but highest in November (> 94%) during the warmer rainy season. These climatic conditions are ideal for the generation and maintenance of *Anopheles* oviposition sites, thus allowing for stable, perennial malaria transmission.



Figure 2.1 **Sampling locations in northwestern Lake Victoria and the southern Uganda peninsular.** Inset top left: Uganda is highlighted in black to illustrate its location within continental Africa. Inset bottom left: Legend lists entomological sampling sites corresponding to numbered black crosses in the main schematic. The black star marks Kampala, the capital city.



The Ssese Islands fall under the administrative jurisdiction of the Kalangala District (Kalangala) local government. The most recent census (2014) lists the population as 54,293: a 56% increase from the previous official figure of 34,800 (2002) (Uganda Bureau of Statistics 2016), which is most likely attributable to the palm oil production and tourism industries that have recently been established on Bugala. Populations tend to be clustered in small communities along the shoreline since fishing is the major economic activity. Kalangala has some of the highest malaria incidence rates in Uganda with 208 cases per 1000 population having laboratory confirmed and/or clinically diagnosed malaria infections (Uganda Ministry of Health 2016). In the most vulnerable group—infants under the age of 5 years—annual prevalence (44%) also indicates one of the highest national transmission rates (Uganda Bureau of Statistics 2016), which is most likely a reflection on the lack of vector control activities in the region. There has never been an organized IRS campaign as part of a government-supported malaria control effort, and the first distribution of LLINs (to pregnant women and infants less than 5 years of age) did not occur until 2009/2010 (National Malaria Control Program 2018). A dedicated National Universal Coverage campaign has since distributed nearly 50 million LLINs nationwide with Kalangala receiving their allocation in November 2017 (National Malaria Control Program 2018). Prior to the mass net distributions, 61% of households in Kalangala were recorded as owning at least one LLIN but usage by all groups was approximately half (household population: 44%, pregnant women: 56%, and infants less than 5 years of age: 50%) (Uganda Bureau of Statistics 2016).

*Anopheles gambiae sensu lato* (*s. l.*) mosquitoes were collected from seven sites: (1) Kansambwe, Nsadzi, (2) Lutoboka, Bugala, (3) Kafuna, Sserinya, (4) Bbosa, Sserinya and (5) Nakibanga, Bukasa from the islands and (6) Lunnyo, Entebbe, and (7) Naama, Wamala from the mainland, reflecting the microsatellite populations analyzed by Kayondo et al. (2005). The villages in the Ssesse Islands are inhabited by human populations that vary in size from hundreds (i.e. Kafuna, Bbosa) to thousands (i.e. Kansambwe, Lutoboka) of individuals. In addition to the continuous fishing traffic that is typically seen at the boat-landing sites, there is notable marine transportation between the mainland and Lutoboka (BL) via an official ferry route, and also the smaller water-taxi type services that frequently traverse the lake i.e. Entebbe-Kansambwe (NZ). Entebbe sits on a southern peninsular extending into Lake Victoria. It differs from the other sampling sites in that it is highly populated (2014 census: 69,430) (Uganda Bureau of Statistics 2016) and urbanized. Naama, located by the shores of the inland Lake Wamala (64 km northwest of Entebbe), is an agricultural village of similar size to Kafuna and Bbosa (SY). The geographic distances between all of the sampling sites are, however, outside of the known flight range of *An. gambiae s. l.* (Gillies 1961) meaning that migration between populations under the mosquitoes' own power, while possible if wind-assisted, is unlikely. Details of global positioning system coordinates listing longitude, latitude, and distances between entomological sampling sites are included in APPENDIX A: .

### **2.3.2 Entomological sampling**

Collections were made at random intervals between July and October 2012. Indoor-resting adult females were collected from houses or common buildings within a 3 km radius of the boat-landing site for each island and the Entebbe locations. Sampling at Wamala was conducted with the same criteria but used Naama village as a center point. Buildings were constructed from a combination of mud and/or wooden walls and thatched/plastic/corrugated sheet metal roofs (Figure 2.2). Specimens were captured between 06:00-10:00 am via battery-powered mechanical aspirators. If insufficient numbers of indoor-resting adult females were collected then aquatic larval sampling was conducted from 5-10 surrounding breeding sites (type varied by location but generally small pools, puddles or abandoned boats), taken back to the laboratory and reared into adults in the water that they were collected in. This water was supplemented with mice feed pellets as required.



Figure 2.2 **Typical housing associated with villages, and small towns in the Ssesse Islands.** Walls are constructed from wooden boards and roofs are corrugated sheet metal (left), which are sometimes covered with a combination of plastic and thatched materials (right). Adult female *Anopheles* mosquitoes can be found resting indoors on walls and objects i.e. bed nets (usually near children) after taking a blood meal. They can also be seen higher up in the eaves of the property where the size of the gaps between the walls and roof are extensive allowing for frequent entry and exit to the exterior.

### 2.3.3 Species identification and preservation

*Anopheles gambiae s. l.* mosquitoes were morphologically identified from other anopheline species based on the taxonomic keys of Gillies & de Meillon (1968). Female specimens were individually preserved in 80% ethanol prior to transportation to the University of Notre Dame (USA). Molecular identification of *An. gambiae* and *Anopheles arabiensis* (the other important malaria vector in the *An. gambiae* complex) was determined by Scott et al. (1993) using legs and/or wings. Any specimens that did not amplify with this method were processed with the polymerase chain reaction (PCR) assays of Koekemoer et al. (2002), which identifies five of the eight formally recognized members of the Funestus Group, and Beebe & Saul (1995) whose technique (although initially developed to discriminate between members of the *An. punctulatus* complex) facilitates molecular identification of many species through variation in the nucleotide sequences of the second internal transcribed spacer (ITS2) region that separates the 5.8S and 28S rDNA subunits (Paskewitz, Wesson, and Collins 1994; Collins and Paskewitz 1996). A minor modification to the rDNA-ITS2 amplification step of Beebe & Saul (1995) was a 1°C increase in the annealing phase to 52°C, and product visualization with both protocols using 1.5% agarose gels stained with SYBR Safe (Life Technologies Corp.). Only specimens that were molecularly identified as *An. gambiae* were processed further i.e. genomically.

#### **2.3.4 Genomic DNA extraction**

Genomic DNA was extracted from individual mosquitoes using a laboratory stock solution of 2% cetyltrimethyl ammonium bromide (CTAB). Each specimen was placed in an Eppendorf tube containing 200 µl of CTAB and electrically homogenized with a sterile conical Teflon pestle. RNA was removed from the homogenate by adding 20 µl RNase A (10 mg/ml) (laboratory stock) and leaving it to incubate at room temperature (RT) for 5 minutes. Proteins were removed with the addition of 20 µl of Proteinase K (20 mg/ml) (Qiagen GmbH, Germany). The solution was briefly vortexed (1-2 seconds) on a low setting (3-4) to encourage maximum digestion and incubated at 56°C for 1 hour. Exoskeleton and other cellular detritus were pelleted by RT centrifugation at 14,000 rpm for 5 minutes. The supernatant was transferred to a Phase Lock Gel tube (5 Prime GmbH, Germany) with 250 µl of UltraPure™ Phenol:Chloroform:Isoamyl alcohol (25:24:1, v/v) (Invitrogen Corp.) for extraction via the standard Phenol:Choloroform method. To ensure thorough mixing, the tube was manually inverted (x10) and the resulting organic layer was removed by centrifugation on maximum speed (14,000 rpm) at RT for 5 minutes. The supernatant was then transferred to a new tube containing 200 µl ice-cold isopropanol, manually inverted x10 again to ensure thorough mixing and centrifuged on maximum speed at 4°C for 15 minutes. The isopropanol was discarded and the DNA pellet rinsed with 200 µl ice-cold 70% ethanol followed by centrifugation on maximum speed at 4°C for 5 minutes. This process was repeated using ice-cold 95% ethanol, which was discarded after a final centrifugation step. Tubes were air-dried for 5 minutes to remove any remaining trace ethanol and the

DNA pellet re-suspended in 50 µl sterile 1x TE buffer (Tris-HCl, EDTA, pH 8.0) (TE) by incubation at 56°C for 5 minutes. DNA was stored at -20°C until further required. Samples were quantified with the QuantiFluor dsDNA System (Promega Corp.) to ensure accuracy. This step was important for the RADseq technique, which required a minimum DNA concentration of 20 ng/µl for restriction enzyme digestion. Sixteen samples of the highest concentration from each of the six locations were selected for analysis to comprise one 96-well plate. 27 samples had DNA concentrations of less than 20 ng/µl. 19 samples (<16.99 ng/µl) were concentrated to 20 ng/µl through 30 minutes of evaporation at 45°C using a heated centrifuge and re-suspended in sterile 1x TE in a heat block for 20 minutes at 55°C. Eight samples (16.99-19.78 ng/µl) were not concentrated on the basis that they were close enough to the recommended minimum concentration and further pipetting may have caused a reduction in DNA quantity. 6 µl of each sample was pipetted into a 96-well plate for RADseq library construction.

### **2.3.5 RADseq library construction and sequencing**

RADseq libraries were prepared as per Parchman et al. (2012), which was modified to incorporate paired-end (PE) chemistry. All samples were digested with *EcoRI* and *MseI* restriction enzymes (NEB, Inc.) and incubated at 37°C for 2 hours, then 65°C for 20 minutes with the heated thermal cycler lid (Eppendorf AG) at 105°C followed by a 4°C hold without the lid. The digested DNA fragments were then ligated to the *EcoRI* and *MseI* adapters with T4 DNA Ligase (NEB, Inc.).

The *EcoRI* adapter sequences consisted of Illumina adapters and primer sequences, a unique 8-10 nucleotide (nt) barcode created by a Python script (Max Planck Institute Bioinformatics Group 2010) that permits identification of the origin of each sequencing read, a protector base to prevent further restriction site cutting, and additional bases to match the sticky ends of the cut sites: (*EcoRI* adapter sequences: 5'-CTCTTTCCCTACACGACGCTCTTCCGATCT + 8-10 nt barcode + C-3', and 3'-TGTGAGAAAGGATGTGCTGCGAGAAGGCTAGA + 8-10 nt barcode + G-5'). The *MseI* adapter sequences were modified from the original protocol to facilitate PE sequencing strategy and also consisted of Illumina adapters and primer sequences, a protector base, and additional sticky end-matching bases: (JT-*MseI*1: 5'-GCAGAAGACGGCATACGAGATCGTGATGTGACTGGAGTTCAGACGTGTGCTCTTCCGATC-3', and JT-*MseI*2: 5'-TAGATCGGAAGAGCACACGTCTGAACTCCAGTCACATCAGATCTCGTATGCCGTCTTCTGCTTG-3'). The DNA plate was incubated in a thermal cycler at 16°C for 2 hours with the heated lid at 20°C followed by a 4°C hold without the lid. Each reaction was then diluted with 90 µl 0.1x TE (pH 7.5) to prevent any further cut reactions, and prepare the product for polymerase chain reaction (PCR) amplification.

Adapter-ligated fragments were amplified using Illumina PCR primers, which were designed to amplify only those DNA sequences with the *EcoRI*- and *MseI*-ligated adapters. Modifying the *MseI* adapter to facilitate PE sequencing necessitated modification of the reverse Illumina PCR primer, accordingly: (Illpcr1 – 5'-AATGATACGGCGACCACCGAGATCTACACTCTTTCCCTACACGACGCTCTTCCGATCT-3'; JT-Illpcr2 – 5'-CAAGCAGAAGACGGCATACGAGATCGTGATGTGACTG-3'). This step



was performed running two separate 20 µl PCR amplification reactions for each adapter-ligated DNA sequence to ameliorate stochastic differences in the resulting reaction products. Both PCR plates were incubated in a thermal cycler using the following profile: 98°C for 30 seconds; 30 cycles of 98°C for 20 seconds; 60°C for 30 seconds; 72°C for 40 seconds, and a final extension at 72°C for 10 minutes with the heated lid at 105°C followed by a 4°C hold without the heated lid.

Reaction products were pooled and purified with Agencourt AMPure XP (Beckman Coulter, Inc.) magnetic beads then size-selected using the automated BluePippin (Sage Science, Inc.) system, which recovered eluted DNA fractions between 400-500 base pairs (bp). Sequencing was accomplished in a single lane run on an Illumina HiSeq 2000 (v.1.5 encoding) machine at the University of California-Davis, Sacramento, USA via the Beijing Genomics Institute.

### **2.3.6 Determination of 2L chromosomal karyotypes**

Molecular karyotyping of the 2La inversion was conducted as per White et al. (2007) with a modified thermal cycler profile as follows: 94°C for 2 minutes; 30 cycles of 94°C for 30 seconds; 58°C for 30 seconds; 72°C for 45 seconds; a final extension at 72°C for 5 minutes, and a 4°C hold. The resulting products were analyzed on 1.5% agarose gels stained with SYBR Safe (Life Technologies Corp.).

### **2.3.7 Bioinformatics processing**

Each read from the Illumina HiSeq 2000 was 100 nt in length beginning with the individual barcode (8-10 nt) ligated to the *EcoRI* end of the amplified fragments plus the single protector base (C or G), and the six bases corresponding to the cut

site (GAATTC) followed by 83-85 potentially variable bases. After quality checking of the sequence data in FastQC v.0.10.1 (Babraham Institute 2011), Illumina sequencing adapters were removed using Trimmomatic v.0.30 (Bolger, Lohse, and Usadel 2014). RAD barcodes were stripped from the reads and replaced by unique identifiers specific to each individual mosquito by a custom Python script, Trimmer (Notre Dame Bioinformatics Lab 2014). Sequence reads were then aligned against the *AgamP4* reference genome (VectorBase 2014) using Burrows-Wheeler Alignment (BWA) v.0.6.2 (Li and Durbin 2009) prior to variant calling and annotation with UnifiedGenotyper in GenomeAnalysisToolKit (GATK) v.3.3.0 (Van der Auwera et al. 2013). High quality SNP calls used in downstream analysis were obtained firstly through the application of the hard-filtering parameters as described in Alternate Protocol 2 of the GATK best practices pipeline (Van der Auwera et al. 2013) followed by high quality streaming of the dataset as per Fontaine et al. (2015) using VCFtools v.0.1.15 (Danecek et al. 2011). A detailed description of the bioinformatics pipeline, including parameters, is listed in APPENDIX B: .

Individuals were pruned from the dataset on the basis of kinship and/or missing data. Familial relationships were assessed by pairwise comparison of kinship coefficients estimated using the [--relatedness] (Yang et al. 2010) and [--relatedness2] (Manichaikul et al. 2010) parameters in VCFtools v.0.1.15 (Danecek et al. 2011). Relationships that identified individuals as full- or half-siblings resulted in their removal from the dataset. Missing genotypes were assessed on an individual

basis using the [--missing-indv] parameter in VCFtools v.0.1.15 (Danecek et al. 2011). Any found to have > 80% missing data were also discarded.

### 2.3.8 Population genetics

Population structure was visualized with two methods: (1) principal components analysis (PCA) using the software packages PLINK v.1.9 (Chang et al. 2015) and R v.3.2.1 (R Core Team 2014), which reduces noise and redundancy thus, maximizing the signal due to variance, and (2) ancestry fractions computed from maximum-likelihood estimates using population allele frequencies and genotype probabilities as parameters of a statistical model in the program ADMIXTURE v.1.23 (Alexander, Novembre, and Lange 2009). The number of ancestral populations ( $K$ ) with which to run the model was chosen by a cross-validation (CV) procedure that identified the lowest error value for which the model had the best predictive accuracy.

Genetic differentiation between populations was quantified by Wright's fixation indices (pairwise  $F_{ST}$ ) (Wright 1978) using Weir-Cockerham weighted multiallelic estimates (Weir and Cockerham 1984) in VCFtools v.0.1.15 (Danecek et al. 2011).  $F_{ST}$  is the *fixation index*, which is used as a measurement of population differentiation due to a reduction in heterozygosity in the *total* population caused by *subpopulation* structure. A value of 0 indicates that the population is in complete panmixis (i.e. no population structure or subdivision), and 1 implies that all of the genetic variation within the population can be explained by its structure (i.e. there is no sharing of genetic material between populations).

Individuals with > 80% missing data were removed to ensure accuracy since simulations have shown that restricting loci to those with complete genotypes results in a near true  $F_{ST}$  distribution (Arnold et al. 2013). To test whether variation was attributable to isolation-by-distance (Wright 1943), a linear regression model of pairwise population differentiations ( $F_{ST}/(1-F_{ST})$ ) against logarithmic transformed geographical distances (Rousset 1997) was created in R v.3.2.1 (R Core Team 2014) using a generalized linear model (GLM) function. Statistical significance between the spatial and genetic sets of distances was measured by the Mantel test with 9,999 permutations (Mantel 1967). Estimates of contemporary effective population size ( $N_e$ ) were obtained using the linkage disequilibrium (LD)-based method LDNe (Waples and Do 2008) of NeEstimator v.2.01 (Do et al. 2014) with a minor allele frequency screen of 5%.

## 2.4 Results

### 2.4.1 Species identification and dataset composition

594 female *An. gambiae* s. l. mosquitoes were morphologically identified from the other anophelines that were captured across all sites. 488 individuals were molecularly identified as *An. gambiae* and one as *An. arabiensis* (from the Entebbe collection site). 100 specimens did not amplify any product in the *An. gambiae* species identification assay despite repeated attempts. These individuals were also processed through the *An. funestus* cocktail PCR to determine whether any of the sampling locations had mixed populations of major malaria vectors; however, no

amplification product (704 bp) was recorded for these specimens. In a final attempt to identify further *An. gambiae* individuals, the rDNA-ITS2 PCR method of Beebe & Saul (1995) was performed. This assay differs from that of Scott *et al.* (1993) as it utilizes the species-specific variation in the ITS2 sequences of the rDNA rather than the non-transcribed IGS regions. No amplification product (426 bp) was observed using tissue (e.g. legs and/or wings) directly without first performing a CTAB extraction, and with this method only one further individual was identified as *An. gambiae*. It was concluded that the unknown specimens could not be identified to species because they were too badly degraded possibly as a result of their physical condition during collection, transportation, and preservation. Five individuals from the Nsadzi collection (sampled as larvae) were not processed due to error. Species identification results are available as a supplementary file through CurateND.

A total of 373,099,980 reads were generated by the Illumina HiSeq 2000 platform. After demultiplexing the raw data of sequencing adapters, barcodes, *EcoRI* and *MseI* restriction cut sites, and protector bases, a total of 172 million reads averaging 1.6 million per mosquito (N=96) were retained for genomic alignment. 103 million forward reads (86.2%), approximately 83 bases in length, successfully mapped to the *AgamP4* reference genome (VectorBase 2014), which were then used in downstream analyses.

Examination of kinship identified a large number of familial relationships between individuals in the Sserinya (SY) and Bugala (BL) populations. Eleven individuals from SY and three individuals from BL were excluded from further analysis on the basis that their full- and half-kinship could confound the data at each

site and in comparison, with others. Three individuals from the Bukasa (BK) population with > 80% missing genotype information were also removed from the dataset (N=79). Kinship coefficient estimates and percentage missing data values per individual are listed in Table S3 of the published supplementary material that supports this research. It can also be downloaded from CurateND as an attachment accompanying this dissertation.

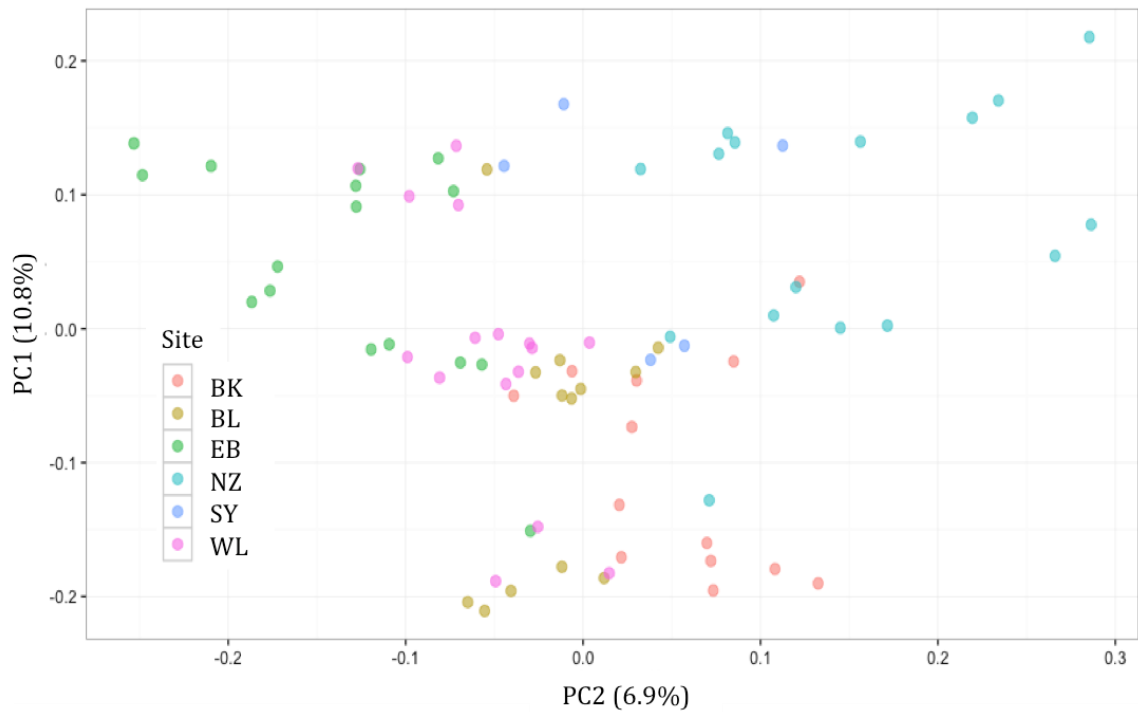
#### **2.4.2 Chromosomal mapping and distribution of SNPs**

After high quality SNP calling, application of hard filters to, and pruning from, the dataset, a total of 5,175 SNPs were identified and mapped to the *AgamP4* chromosomes (VectorBase 2014) as follows: X (n=347), 2L (n=1,078), 2R (n=1,514), 3L (n=936), 3R (n=1,204), and mitochondrial (n=1). 95 SNPs were unable to be assigned to any chromosome (UNKN) but were included in a population genetic analysis when the collinear genome was being explored. The UNKN SNPs are most likely physically located in the highly repetitive pericentromeric regions (Sharakhova et al. 2007), which are challenging genomic positions to assemble and map.

#### **2.4.3 Population structure**

Visualization of population structure by PCA illustrated how the SNPs genetically clustered within and between collection sites. Genome-wide analysis (n=5,175) showed individuals clustering into three discrete groups on the first principal component (PC1) in a non-geographical configuration (Figure 2.3), which was also observed in the chromosome 2L (n=1,078) PCA (Figure 2.4), a pattern

likely driven by polymorphism with respect to the 2La inversion (APPENDIX C: ) (O'Loughlin et al. 2014). When 2L SNPs were removed from the dataset, or when other chromosome arms were analyzed individually, genetic structure showed individuals generally clustering in concordance with their geographic origin (Figure 2.5; Figure S1 of the published supplementary material supporting this research, which can also be downloaded from CurateND as an attachment accompanying this dissertation). Since the 2La inversion is known to confound population genetic structure (O'Loughlin et al. 2014; Neafsey et al. 2015), chromosome 2L SNPs were removed from the dataset (n=4,097) for further downstream analysis.



**Figure 2.3 Principal components analysis (PCA) plot of the 5,175 genome-wide SNPs shared between the six *An. gambiae* populations (N=79).** Each circle represents an individual mosquito. Entomological sampling sites are color-coded according to the legend (Site) as follows: BK (Bukasa); BL (Bugala); EB (Entebbe); NZ (Nsadzi); SY (Sserinya), and WL (Wamala). The first principal component (PC1) and its percentage variance are represented by the y-axis, and likewise, the x-axis represents the second principal component (PC2) and its percentage variance.



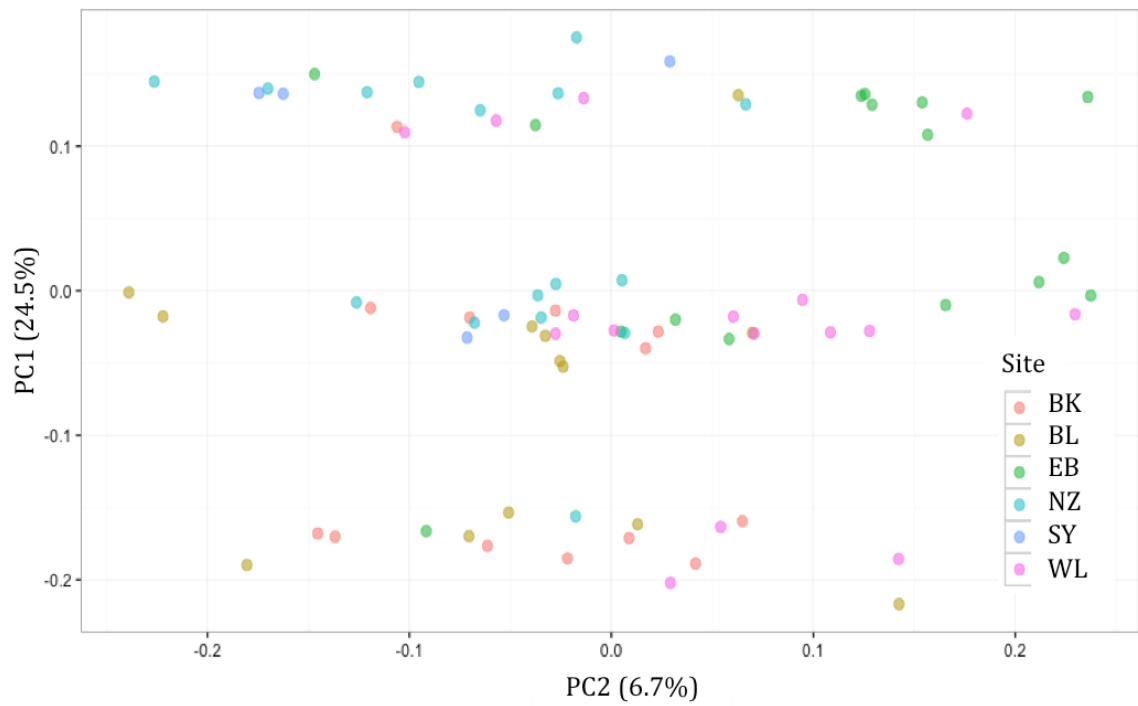


Figure 2.4 **PCA plot of the 1,078 chromosome 2L SNPs in the six *An. gambiae* populations (N=79).** Plot descriptors and legend are the same as for Figure 2.3.

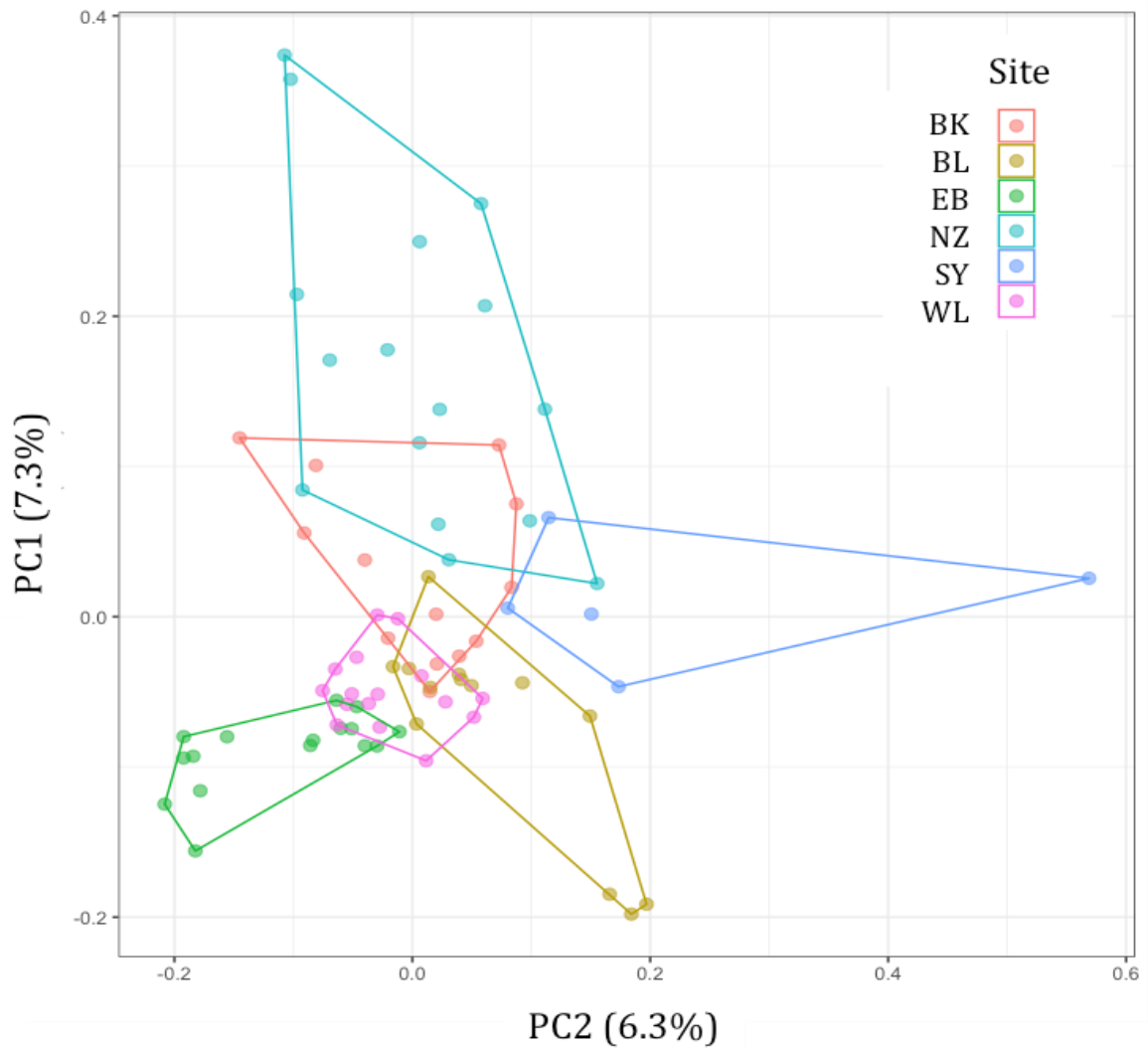


Figure 2.5 **PCA plot of the 4,097 SNPs in the six *An. gambiae* populations (N=79) after removal of the chromosome 2L SNPs (n=1,078) suspected of confounding genetic structure.** Plot descriptors and legend are the same as for Figure 2.3. Here, polygonal code was applied to better illustrate geographic rather than karyotypic clustering.

Population structure was additionally tested by model-based estimation of ancestry using the software program ADMIXTURE v.1.23 (Alexander, Novembre, and Lange 2009). CV testing for each chromosome produced error estimates that indicated the populations shared only one ancestry except for 2L where  $K=2$  was the most likely number of fractions (**Error! Reference source not found.**). This was in line with the expectation that the populations would cluster into their chromosomal inversion arrangements ( $2L^{+a}/^{+a}$ ,  $2L^a/a$ , and  $2L^{+a}/a$ ) rather than geographic locations for this region of the genome (APPENDIX C: ).

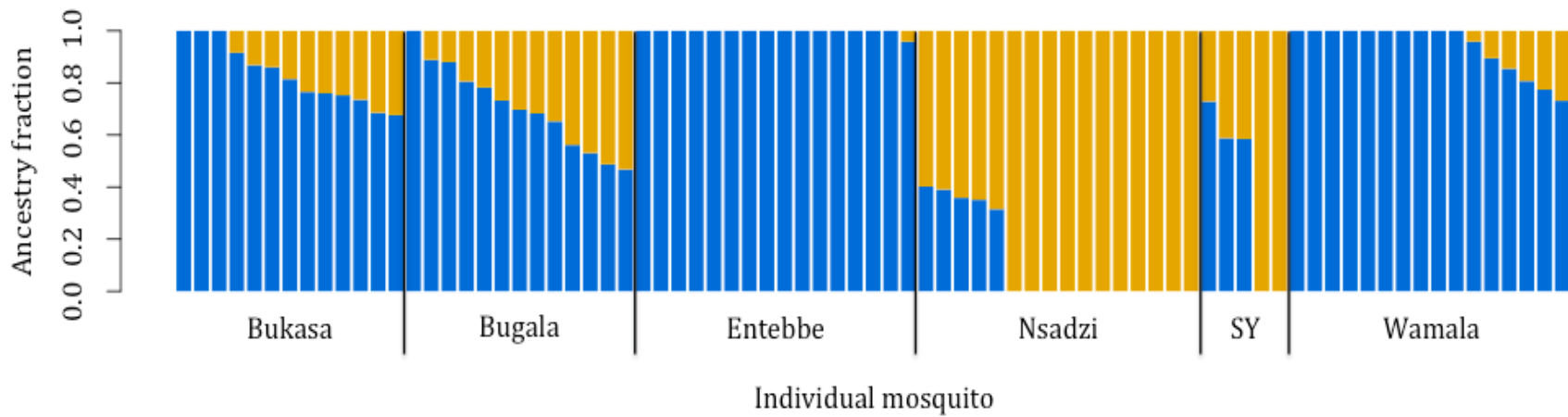


Figure 2.6 **ADMIXTURE plot of chromosome 2L SNPs showing probable ancestry fractions ( $K=2$ ) of the six *An. gambiae* populations ( $N=79$ ) based on the 2La inversion.** Vertical bars represent individual mosquitoes, which are stacked from left to right in groups of 16 (except: Bugala ( $N=13$ ), Bukasa ( $N=13$ ), and Sserinya (SY) ( $N=5$ )) according to their alphabetical sampling site. Populations are listed on the x-axis. Ancestry fractions ( $K=2$ ) are assigned according to maximum likelihood probabilities. Each fraction represents a karyotype of different color: blue = inverted ( $2La/a$ ), yellow = standard ( $2L^{+a}/2L^{+a}$ ), and blue/yellow = heterozygote ( $2La/+^a$ ).

#### 2.4.4 Genetic differentiation

Pairwise  $F_{ST}$  comparisons between the populations were computed for all mapped SNPs according to Weir & Cockerham (1984) weighted estimates. Median values were used in the analysis since a null distribution histogram showed that they were not normally distributed. Moderate amounts of genetic differentiation were observed between most of the populations (median  $F_{ST}$ : 0.0342–0.0903) for the 4,097 SNPs across the collinear genome after removal of the SNPs on chromosome 2L (TABLE 2.1). Generally, inter-island comparisons generated the greatest differences between populations with the strongest signals of genetic differentiation being observed in the comparisons with Sserinya (median  $F_{ST}$  > 0.08).

TABLE 2.1

GENETIC DIFFERENTIATION BETWEEN POPULATIONS AS MEASURED BY MEDIAN  
WEIR-COCKERHAM WEIGHTED  $F_{ST}$  ESTIMATES

<i>Population</i>	<i>Wamala</i> <sup>a</sup>	<i>Bukasa</i>	<i>Bugala</i>	<i>Sserinya</i>	<i>Nsadzi</i>
<i>Entebbe</i> <sup>a</sup>	0.0342	0.0503	0.0457	0.0100	0.0444
<i>Wamala</i> <sup>a</sup>	-	0.0446	0.0412	0.0903	0.0389
<i>Bukasa</i>	-	-	0.0532	0.0800	0.0520
<i>Bugala</i>	-	-	-	0.0826	0.0480
<i>Sserinya</i>	-	-	-	-	0.0846

NOTE:

<sup>a</sup> Denotes mainland population

Geographical distance between populations is often the primary force driving genetic differentiation, therefore, isolation-by-distance as the model explaining the variation between the populations was tested by simple linear regression of median  $F_{ST}/(1-\text{median } F_{ST})$  against a natural logarithm transformation of geographic distance (Rousset 1997). The resulting GLM plot showed no evidence of a correlation between the two variables ( $y = 0.104975 - 0.012853x$ ;  $R^2 = 0.05$ ; Mantel:  $p = 0.2$ ) meaning that geographic distance alone could not explain the variation observed between the populations (Figure 2.7).

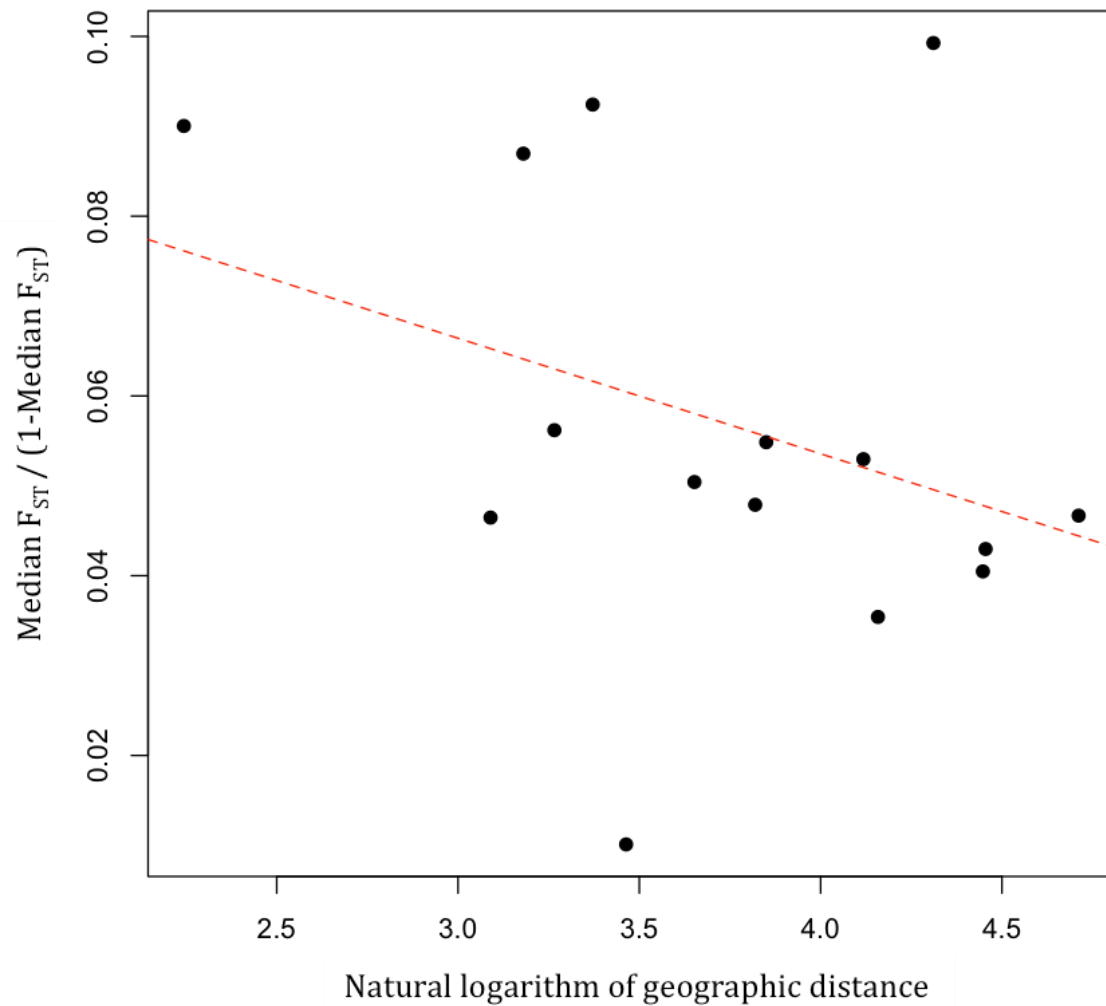


Figure 2.7 **Generalized linear model (GLM) plot of genetic differentiation against log geographic distance based on median  $F_{ST}$ .** The regression equation ( $y=0.104975-0.012853x$ ) describes the statistical relationship between the predictor  $\ln(\text{geographic distance})$  and response ( $\text{Median } F_{ST}/(1-\text{Median } F_{ST})$ ) variables. The line of best fit is represented by the red dotted line.

#### 2.4.5 Effective population size

Estimations of  $N_e$  were obtained using the LDNe (Waples and Do 2008) method in NeEstimator v.2.01 (Do et al. 2014) on the basis of superior performance compared to other single-sample estimators (Do et al. 2014; Gilbert and Whitlock 2015; J. Wang 2016). Generally, smaller estimates of  $N_e$  were observed for all populations (TABLE 2.2) compared to those recorded for other continental populations of *An. gambiae* i.e. 6,689 (Kenya) (Lehmann et al. 1998); 13,200 (Equatorial Guinea) (Athrey et al. 2012); 2 million (East Africa) (O'Loughlin et al. 2014). The largest estimates were seen in Bugala (1,098.3) and Wamala (1,920.3), which were substantially higher than the other populations. The infinity estimates recorded for the Sserinya populations were derived from negative points, which implied that variation was due to sampling error alone and not genetic drift (allelic frequency changes due to random sampling). This was unsurprising given the high levels of kinship observed in both populations as a result of the unrepresentative entomological sampling that necessitated the removal of eleven individuals from the dataset. The coefficient of variation is a measurement of genetic drift specific to the LDNe method of NeEstimator v.2.01 (Do et al. 2014) and is calculated as the inverse of  $N_e$ . The lower estimates of 0.001 observed for Bugala and Wamala indicated that these populations would be more resistant to the effects of genetic drift compared to the higher coefficients of variation recorded for Bukasa (0.005), Entebbe (0.005), and Nsadzi (0.008), which implied vulnerability to allelic dropout or fixation.



TABLE 2.2

ESTIMATES OF EFFECTIVE POPULATION SIZE ( $N_e$ ) ADOPTING A MINOR ALLELE FREQUENCY FILTER OF 5%

<i>Population</i>	<i>Bukasa</i>	<i>Bugala</i>	<i>Entebbe</i> <sup>a</sup>	<i>Nsadzi</i>	<i>Sserinya</i>	<i>Wamala</i> <sup>a</sup>
<i>Sample size</i>	13.0	13.0	16.0	16.0	5.0	16.0
$X_h$ <sup>b</sup>	8.0	9.4	13.8	14.6	4.7	15.0
<i>LD</i> ( $r^2$ ) <sup>c</sup>	0.17	0.14	0.09	0.09	0.37	0.08
<i>Estimated</i> $N_e$	211.7	1,098.3	213.5	124.2	$\infty$	1,920.3
<i>95% CIs</i> <sup>d</sup>	180.6-255.4	637.6-3913.7	195.8-234.6	118.4-130.7	$\infty$	1124.2-6523.8
<i>CV</i> <sup>e</sup>	0.005	0.001	0.005	0.008	$\infty$	0.001

NOTE:

<sup>a</sup> Denotes mainland population<sup>b</sup>  $X_h$  is the *harmonic mean* of the six sample sizes defined as the reciprocal of the arithmetic mean of the reciprocals of the sample sizes. See (Clark-Carter 2005) for details<sup>c</sup>  $r^2$  is an indication of *linkage disequilibrium* (LD)<sup>d</sup> Parametric confidence intervals for estimated  $N_e$ <sup>e</sup> CV is the coefficient of variation, which is calculated by  $1/N_e$

## 2.5 Discussion

### 2.5.1 There are limited malaria vector species in the Ssesse Islands

The majority of the dataset comprised of *An. gambiae* mosquitoes. Only 1 of 481 female anophelines was molecularly identified as *An. arabiensis* and this was from the mainland (Entebbe) population. Kayondo et al. (2005) reported approximately 20% of the second year Bukasa collection (N=47) as *An. arabiensis*, which was attributed to asynchronous entomological sampling. This observation highlights the importance of systematic population sampling to establish changes in (a) species composition, (b) vector abundance, and (c) seasonality, which are some of the factors that can influence the genetic structure and effective size of a population. It is also interesting to note that in this study no *An. arabiensis* or *An. funestus* were collected in the islands inferring a limited malaria vector species distribution here: a key component of a GM mosquito control effort (World Health Organization 2014). In the event that entomological surveillance demonstrates the presence of additional *An. gambiae s. l.* or *An. funestus* vectors, a GM release would still be valuable in monitoring gene drive efficacy since the potential for transference of the genetic construct to sibling species could be assessed.

### 2.5.2 The 2La inversion confounds population genetic structure

Previous studies have demonstrated a strong association between the frequency of the 2La inversion and aridity, which shifts seasonally according to climate (Coluzzi et al. 1979; Touré et al. 1998) but there are no prior published data about its distribution

in the Ssesse Islands. This is, therefore, the first reported 2La karyotypic distribution of *An. gambiae* mosquitoes in the Ssesse Islands.

Principal components analysis (PCA) illustrated that the 2La inversion confounded population genetic structure (Figure 2.3, Figure 2.4) thus, chromosome 2L SNPs were removed from the dataset to disclose the underlying population structure of the collinear genome (Figure 2.5).

### **2.5.3 Moderate but significant genetic differentiation is observed in island populations**

Genetic differentiation among the Ssesse islands was moderate in magnitude (median  $F_{ST}$ : 0.0480–0.0846) but significantly greater than the very low differentiation between *An. gambiae* populations observed across opposite sides of continental Africa (mean  $F_{ST}$ : 0.016) (Lehmann et al. 1996) and comparable in magnitude to populations separated by the KRVC (mean  $F_{ST}$ : 0.104), which acts as a physical barrier to gene flow (Lehmann et al. 1999). It would seem reasonable to suggest that water also acts as a physical barrier to gene flow in locations where it separates populations—just as the KRVC does further inland—since higher differentiation and low amounts of gene flow were identified in oceanic island studies of *An. gambiae* in the Comoros (mean  $F_{ST}$ : 0.199–0.250) (Marsden et al. 2013), and of *An. arabiensis* in Madagascar, Reunion, and Mauritius (mean  $F_{ST}$ : 0.169) (Simard et al. 1999). Kayondo et al. (2005) reported mean  $F_{ST}$  values of 0.014–0.105 in the same *An. gambiae* populations sampled here, which are of a lower magnitude than those in the oceanic island studies but not unexpected given the smaller distances involved that allowed for frequent human-marine transportation

routes or even wind-borne diffusion, which might have passively dispersed mosquitoes. This would suggest that water is not an absolute physical barrier to gene flow in this region as supported here by the evidence indicating limited migration between populations (TABLE 2.1; Figure 2.5) and greater genetic structure that is present in the islands. This was also observed in the PCA plot (Figure 2.5) where individuals clustered in concordance with their geographic origin. The two mainland sites, Entebbe and Wamala, also showed signs of population differentiation (median  $F_{ST}$ : 0.0342) (Figure 2.5 polygons: green-Entebbe; pink-Wamala). Neither of these locations is separated by water but the sampling point in Entebbe—Lunnyo—sits on the edge of a small harbor, which is separate from the rest of the town. There are no other obvious geographical barriers to gene flow between these sites so the explanation as to why the two mainland *An. gambiae* populations appear somewhat differentiated from each other must be due to other unknown factors of demographic, ecological, and/or anthropogenic origin.

An array of molecular markers have been used to explore genetic differentiation in continental populations including microsatellites, mtDNA, allozymes and, more recently, SNPs (O’Loughlin et al. 2014; The *Anopheles gambiae* 1000 Genomes Consortium 2017). That all of these marker systems identify the same pattern of low genetic differentiation on the continent suggests that the heightened differentiation estimated in this study is real, and not attributable to differences in the genetic markers themselves: an assertion strengthened by the previous Ssesse Islands studies of Kayondo et al. (2005) and Lukindu et al. (2018) using microsatellites and mtDNA, respectively.

#### 2.5.4 Island populations have small effective sizes

Higher levels of genetic differentiation could also be explained to some extent by the small estimates of effective population size ( $N_e$ ) that were obtained through the LD method of NeEstimator v.2.01 (Do et al. 2014).  $N_e$  determines how random genetic drift affects the stability of allelic frequencies, which are more variable in small populations i.e. there is a greater probability that an allele will become either fixed ( $f=1.0$ ) or lost ( $f=0.0$ ) in a smaller population since fewer individuals may or may not be carrying  $f$ . It follows, therefore, that  $N_e$  will have a greater influence on genetic variation in these populations.

Kayondo et al. (2005) estimated that the island populations consisted of smaller demes in the hundreds (397-677) compared to the mainland populations that were in the thousands (8,810-8,935). This was anticipated since *An. gambiae* is usually found in close proximity to human habitation (Coluzzi et al. 1979) and the collection sites in the islands are less intensely populated than those on the mainland. The estimates in this study are generally smaller but comparable to those of Kayondo et al. (2005) (< 397) with the exception of Bugala, which has increased to 1,098. Over the last few years, the human population size on Bugala has grown substantially (2002: 34,800 - 2016: 54,293) (Uganda Bureau of Statistics 2016) as a result of: (i) economic development arising out of the expansion of a 10,000 hectares oil palm plantation and related mill facility (Carmody and Taylor 2016), and (ii) tourism. Coupled with increased boat traffic to/from the mainland, human population growth may have also led to the concomitant population growth of *An. gambiae*.

One of the key components of a population genetic analysis is the temporal stability of the population. A limitation of this study is that there is only one time point to estimate  $N_e$ , which can fluctuate in accordance with climatic changes. It is recommended, therefore, that future research should focus on multi-year, and multi-season longitudinal sampling to account for variances in malaria vector abundance (Mukiama and Mwangi 1989; Kabbale et al. 2013) to mitigate any impact on the timing of a proposed GM mosquito release, which may adversely affect effective population sizes.

## 2.6 Conclusion

This is the first genome-wide SNP-based study of *An. gambiae* population connectivity and effective size in the Lake Victoria region. The island populations comprise of a dominant malaria vector species (*An. gambiae*) with low to moderate genetic differentiation and greater structure suggesting some limitation to migration between them. Smaller estimates of effective population size indicate that an introduced effector transgene should be susceptible to genetic drift but to ensure that it is driven to fixation instead of loss the construct would have to be paired with a robust gene drive mechanism.

Taking these findings into consideration, together with their favorable location and suitability for frequent monitoring, the Ssesse Islands contain several candidate *An. gambiae* populations, which merit further evaluation in regard of a potential GM mosquito pilot release.

## CHAPTER 3:

### IDENTIFYING THE GEOGRAPHIC STRUCTURE OF *ANOPHELES FARAUTI* POPULATIONS IN THE SOUTHWEST PACIFIC USING SINGLE NUCLEOTIDE POLYMORPHISM MARKERS

The purpose of this chapter is to present the reader with research, which contributes expertise towards the second developmental challenge identified by the malERA Consultative Group on Vector Control: interventions that affect vector species not effectively targeted by current tools. Prior to the development and distribution of LLINs, IRS with DDT was the primary vector control tool in the southwest Pacific malaria programs of the 1970s. Huge reductions in *Anopheles punctulatus* and *An. koliensis* populations were achieved (because of their high anthropophilic affinity), and also in the *An. farauti* populations that demonstrated (what is considered to be) classic late night indoor-biting behavior. A key entomological outcome, which developed in the Solomon Islands as a result of DDT selection pressure, concerned a proportion of the *An. farauti* population that were observed to feed outdoors and earlier in the evening—behavioral adaptations, most commonly referred to as resistance (Taylor 1975a, 1975b). These mosquitoes, as a consequence of failing to enter domiciles where lethal doses of insecticide had been

applied, failed to come into contact with the control intervention thus, maintaining residual transmission, and it is *these* populations that novel interventions must target in order to *supplement* LLINs and larval source management.

Malaria vector behavior is variable and most likely results from complex genetic and environmental interactions, which develop in response to exogenous cues that differ between species, and populations within species. Vectorial capacity is determined by genetic components that contribute to key biological traits, such as behavior and susceptibility to *Plasmodium* infection, which are also variable between vectors. Understanding how these factors interact at the genomic level can provide greater insight into disease transmission since it permits the detailed analysis of functional characteristics—such as gene turnover in cuticular proteins and their role in insecticide resistance—that can be manipulated as targets for malaria vector control.

A genome-wide approach in the analysis of population structure also provides power to identify genetic processes (such as restricted gene flow or an insecticide-resistant genotype) and the genomic regions that regulate them thus, creating an understanding of how these mechanisms will impact upon control interventions.

Entomological sampling of natural *Anopheles farauti* populations was conducted in four locations (Australia, Papua New Guinea, Solomon Islands, and Vanuatu) across its geographic distribution. Four individuals from each site (N=16) were sequenced with a whole genome shotgun approach. Single nucleotide polymorphism (SNP) markers confirmed geographic population structure and also



identified genetic similarity between the populations in Australia and Vanuatu as well as a lack of diversity within each. In conclusion, the research in this chapter enhances prior *An. farauti* population genetic studies by adopting a SNP-based approach, which permitted high resolution genome-wide analysis that detected a previously unidentified geographic relationship in this important malaria vector.

### 3.1 Systematics of malaria vectors in the southwest Pacific

The anopheline mosquitoes in this region have been previously well described (D. Lee 1987). Members of the *Anopheles punctulatus* group—a complex consisting of 13 species—are the primary vectors responsible for transmitting malaria in the southwestern Pacific region (TABLE 3.1). Although not formally incriminated as a vector, certain members of the *Anopheles lungae* complex have recently been recorded as man-biting species (T. Burkot et al. 2018) and are presented here for completeness.

TABLE 3.1

DISTRIBUTION AND VECTOR STATUS OF THE *ANOPHELES LUNGAE* COMPLEX AND*AN. PUNCTULATUS* GROUP \*

<i>Group/complex-species</i>	<i>Distribution</i>	<i>Vector status</i>
<b><i>An. lungae</i> complex</b>		
<i>An. lungae</i> Belkin & Schlosser, 1944	Solomon Islands	Unresolved <sup>a</sup>
<i>An. nataliae</i> Belkin, 1945	Solomon Islands	Unresolved
<i>An. solomonis</i> Belkin, Knight & Rozeboom, 1945	Solomon Islands	Unresolved <sup>a</sup>
<b><i>An. punctulatus</i> Group</b>		
<i>An. punctulatus</i> Dönitz, 1901	New Guinea Solomon Islands	Primary
<i>An. koliensis</i> Owen, 1945	New Guinea Solomon Islands	Primary
<i>An. clowi</i> Rozeboom & Knight, 1946	New Guinea	Non-vector
<i>An. rennellensis</i> Taylor & Maffi, 1991	Solomon Islands	Non-vector
<i>An. sp. near punctulatus</i> Foley, Cooper & Bryan, 1995	New Guinea	Non-vector
<b><i>An. farauti</i> complex</b>		
<i>An. farauti</i> Laveran, 1902	Australia New Guinea Solomon Islands Vanuatu	Primary
<i>An. hinesorum</i> Schmidt, 2001	Australia New Guinea Solomon Islands	Secondary
<i>An. torresiensis</i> Schmidt, 2001	Australia	Unresolved
<i>An. farauti</i> 4	New Guinea	Secondary
<i>An. farauti</i> 5	New Guinea	Non-vector
<i>An. farauti</i> 6	New Guinea	Secondary
<i>An. irenicus</i> Schmidt, 2003	Solomon Islands	Non-vector
<i>An. farauti</i> 8	New Guinea	Secondary

NOTE:

\* Adapted from Beebe et al. (2013)

<sup>a</sup> *An. lungae* and *An. solomonis* were identified as biting humans in Western Province, Solomon Islands but no *Plasmodium* DNA was detected in heads or thoraces as an indicator of malaria sporozoite transmission (T. Burkot et al. 2018)

*Anopheles punctulatus* was the first species to be described from the Madang area of Papua New Guinea (Dönitz 1901) followed by *An. farauti* in Vanuatu (Laveran 1902). Further species were not identified until the extensive entomological studies carried out during World War II recognized *An. koliensis* and *An. clowi* as closely related members in what became known as the Punctulatus Complex (Rozeboom and Knight 1946).

The distinguishing morphological feature used to identify a specimen to species was the unique coloration and markings of the proboscis, which Rozeboom & Knight (1946) designated as *type A* (all black-scaled labium with a pale apical ring at the basal tip: *An. farauti*), *type B* (apical half of the labium extensively or completely covered in white scales: *An. punctulatus*), and *type C* (apical half of the labium has a ventral patch of white scales: *An. koliensis*) (Figure 3.1); however, polymorphism within types B and C generated confusion in the literature and this form of identification was eventually considered unreliable.

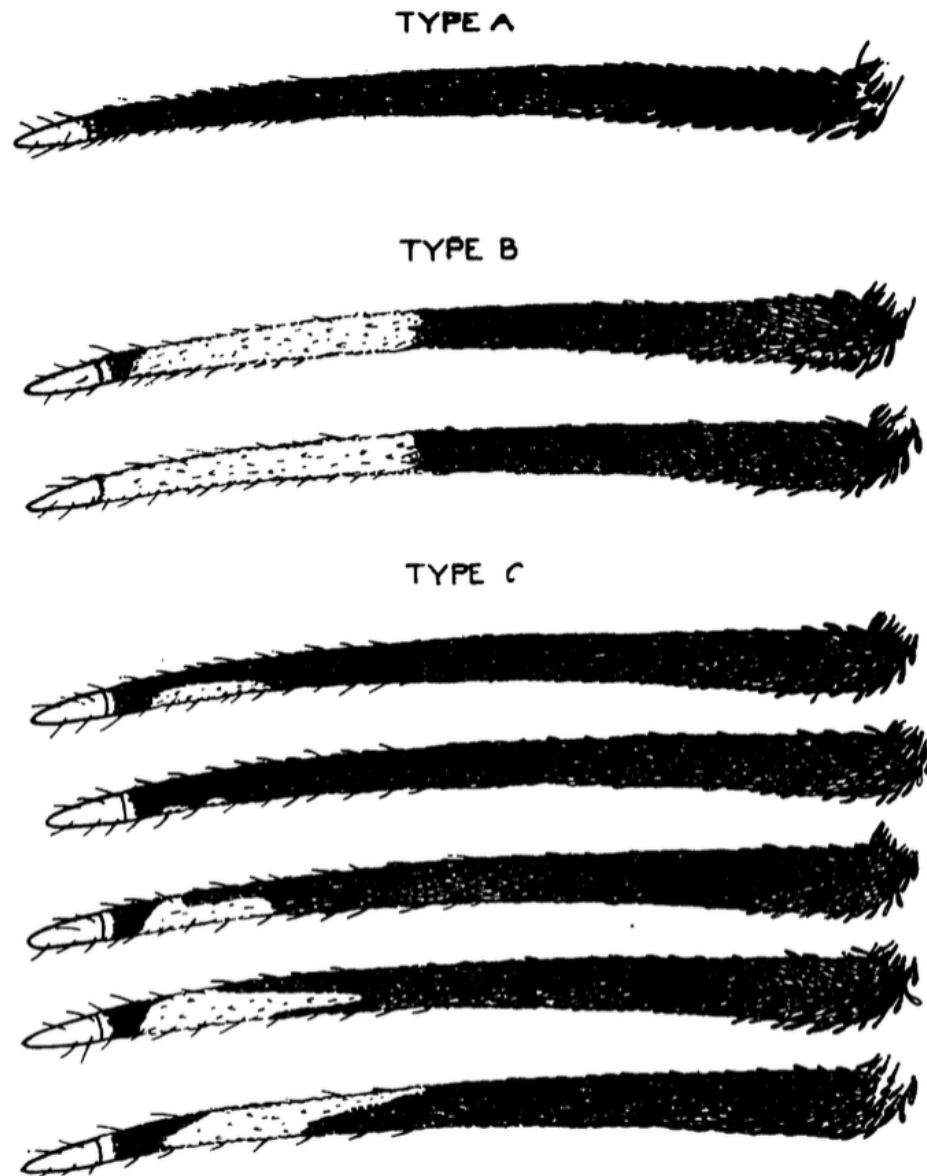


Figure 3.1 **Schematic drawing illustrating different proboscis coloration patterns in the Punctulatus Complex.** Top down: type A (*An. farauti*), B (*An. punctulatus*), and C (*An. koliensis*). Note the polymorphism within types B, and C. From Rozeboom & Knight (1946).

Cytogenetics became a more informative method of species identification after the discovery of two further species (*An. hinesorum* and *An. torresiensis*) due to post-mating barriers (Bryan and Coluzzi 1971; Mahon and Miethke 1982; Mahon 1983). This was followed by advanced molecular techniques, which used allozyme markers that determined the species *An. farauti* 4, 5, and 6 in Papua New Guinea (Foley et al. 1993), *An. irenicus* in the Solomon Islands (Foley, Meek, and Bryan 1994), and *An. sp. near punctulatus* (an uncommon non-vector ) from the Papua New Guinea highlands (Foley, Cooper, and Bryan 1995).

The *An. farauti* species complex shares an extensive distribution, which is demarcated in southeastern Indonesia—by the hypothetical zoogeographical boundaries (separating Asia and Australia) first described by Wallace (1863) then Weber (1902)—and in the southwest Pacific Ocean as far south as Vanuatu, which is the naturally occurring limit of *Anopheles* mosquito vectors (Buxton Line 170°E, 20°S). Figure 3.2 shows the predicted geographic distribution of the complex as estimated with likelihood probabilities derived from Boosted Regression Tree modelling (Sinka et al. 2011) Likelihood of occurrence is indicated by either high (red) or low (blue) probability of locating a species within the complex.

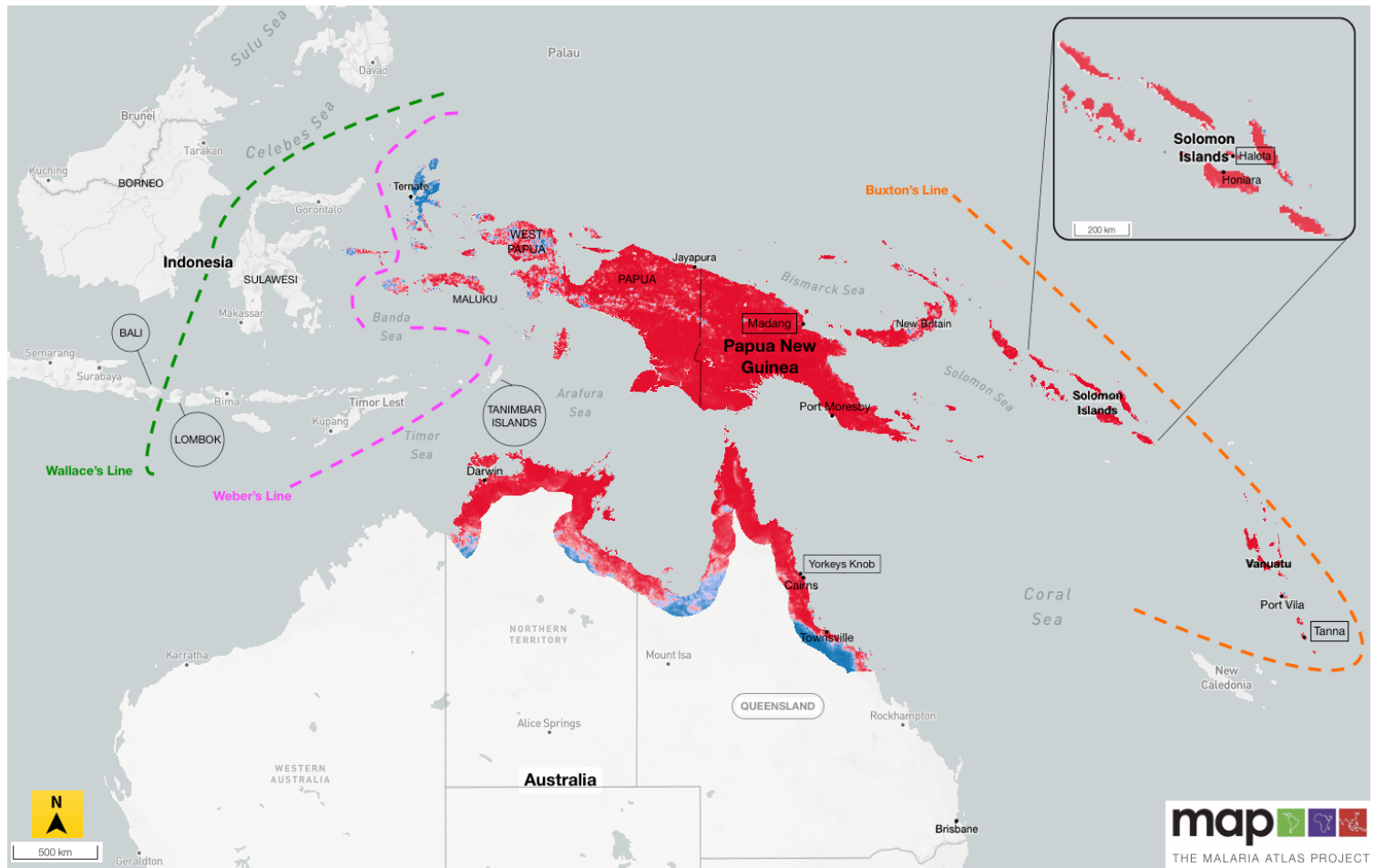


Figure 3.2 **Predicted geographic distribution of the *Anopheles farauti* species complex.** Entomological sampling sites (Australia-Yorkeys Knob; Papua New Guinea-Madang; Solomon Islands-Haleta; Vanuatu-Tanna) are detailed in black open-framed rectangles. Adapted from the Malaria Atlas Project (2019).

The southwest Pacific experiences the second highest rates of malaria transmission globally following those of sub-Saharan Africa (World Health Organization 2018d) despite the application of IRS operations and distribution of LLINs by the Pacific Community (SPC) member states. This is most likely due to the primary malaria vector *An. farauti* having the widest geographic range of any species in the complex due to its preference for coastal habitats resulting from a physiological ability to tolerate increased levels of salinity. Previous population genetic studies have revealed distinct geographically structured populations using single locus markers i.e. mtDNA cytochrome oxidase I (COI) and the ribosomal intragenic transcribed spacer regions (ITS1 and ITS2) (Beebe et al. 2000; Bower et al. 2008; Ambrose et al. 2012, 2014) demonstrating sufficient sensitivity to detect barriers to gene flow.

### 3.2 Current objective

Motivation for this study developed from the research objectives of the *Anopheles* 16 Genomes (*An16G*) Project (Neafsey et al. 2013). Sequencing of *An. gambiae*—the major sub-Saharan malaria vector (Holt et al. 2002)—generated a genome-wide resource, with the capacity to support an array of genetic studies that facilitated a greater understanding of the biological traits they underpinned (Touré, Oduola, and Morel 2004). For many years following the *An. gambiae* sequencing effort (mainly as a result of the high costs associated with reagent chemistry and the availability of high-throughput sequencers), the equivalent genomic resources were not available for other important malaria vectors and research was limited to the

application of specific genes or genetic marker sets such as microsatellites and mtDNA. The majority of biological processes are coordinated by complex multigene interactions as opposed to a single gene. It follows, therefore, that single locus or minimal markers will most probably be insufficient to elucidate the genetic mechanisms that underlie these processes.

Of the approximately 535 *Anopheles* species identified to date, only 41 are considered to be malaria vectors of major importance (Sinka et al. 2012; The Malaria Atlas Project 2019) and, within that subset, their ability to transmit malaria (vector competence) differs leaving researchers to consider the biological basis of such variability. The An16G Project was conceived to generate genomic and transcriptomic resources of 16 anopheline species that would facilitate a comparative framework to enable greater comprehension of vectorial capacity, its biological components (i.e. physiology, molecular architecture, and behavior), and how these could possibly be manipulated with the aim of reducing malaria transmission. The current objective of this study was to apply the newly available *An. farauti* genomic resources to wild-type individuals from natural populations across its ecological range, and identify any patterns of genetic structure.

### 3.3 Materials and methods

#### 3.3.1 Entomological sampling

*Anopheles farauti* s. l. mosquitoes were sampled from natural populations in 2012 as part of the An16G Project (Neafsey et al. 2013, 2015) by oral aspiration



(barrier screen and human landing catch (HLC) collections) or as dead specimens (CDC light trap) (TABLE 3.2).

TABLE 3.2

GEOGRAPHIC METADATA FOR ENTOMOLOGICAL COLLECTIONS OF *ANOPHELES FARAUTI* S. L. MOSQUITOES SAMPLED FROM NATURAL POPULATIONS ACROSS ITS ECOLOGICAL DISTRIBUTION

<i>Country</i>	<i>Location</i>	<i>Collection method</i>	<i>GPS co-ordinates</i>
Papua New Guinea	Madang, (Mirap)	Barrier screen <sup>a</sup>	-4°45'10"S, 145°40'0"E <sup>b</sup>
Australia	Queensland, (Yorkeys Knob)	CDC light trap	-16°49'24"S, 145°42'57"E
Solomon Islands	Central, (Haleta)	HLC <sup>c</sup>	-9°5'56"S, 160°6'56"E
Vanuatu	Tanna, (Uiak)	HLC <sup>c</sup>	-19°26'3"S, 169°13'38"E <sup>d</sup>

NOTE:

<sup>a</sup> (T. Burkot et al. 2013)

<sup>b</sup> Converted from decimal degrees (-4.7527, 145.6667) (National Geodetic Survey 2018)

<sup>c</sup> Human landing catch

<sup>d</sup> Converted from decimal degrees (-19.434193, 169.227215) (National Geodetic Survey 2018)

### 3.3.2 Species identification and preservation

*An. farauti* specimens were morphologically identified to species complex under field conditions using the taxonomic keys of Belkin (1962a, 1962b). After preservation on silica beads (Papa New Guinea) or in ethanol (Australia, Solomon Islands, Vanuatu) specimens were transferred to the laboratory where a single leg

was dissected as source material for molecular species identification by the ITS2 assay of Beebe and Saul (1995). Confirmed *An. farauti* specimens were transported to the University of Notre Dame, USA where a subsequent ITS2 assay was conducted to validate initial results. Subsets from each collection were then selected for genomic sequencing at the Broad Institute of MIT and Harvard, USA. All shipments were completed as whole animals on silica beads.

### 3.3.3 Whole genome library construction and sequencing

Sequences were generated using an Illumina whole genome shotgun approach, which is detailed in the supplementary pages of Neafsey et al (2015). A brief summary follows, however, for introduction and ease of reference.

#### 3.3.3.1 Reference genome

The *An. farauti* reference genome (*AfarF1* assembly) was generated using individuals of the FAR1 strain (*AfarF1*) (source: Papua New Guinea, isofemale, subcolony, Malaria Research and Reference Reagent Resource Center (MR4)) from three sequencing libraries with different insert sizes on an Illumina MiSeq platform in a single run:

1. 180 bp 'fragment': created from a single individual to minimize heterozygosity
2. 1.5 kb 'jump': generated from the same individual as for the fragment library and created to circumvent difficult to align regions such as repetitive DNA
3. 38 kb 'Fosill': (Fosmid-scale Illumina-compatible jump library) constructed of high molecular weight DNA from several hundred individuals to improve scaffolding

### 3.3.3.2 Samples from natural populations

Four individuals were randomly selected from each natural population (*wild-type*) (N=16) and each sequenced for four runs on an Illumina HiSeq2000 platform with v.1.9 encoding.

## 3.3.4 Bioinformatics processing

### 3.3.4.1 *AfarF1* assembly

*AfarF1* genomic sequences were assembled using the ALLPATHS-LG algorithm (Gnerre et al. 2011). Specific parameters applied to the *AfarF1* reads were “HAPLOIDIFY=True” to further minimize heterozygosity and “ReadFilterByKmerFreq=0.6”, a *K*-mer (sequence of *K* consecutive bases) normalization tool that down-samples high coverage data. Scaffolding gaps and errors were identified post-assembly by Pilon (Broad Institute 2012b) and a manual analysis for quality was conducted with GAEMR (Broad Institute 2012a). Assembly contigs (set of contiguous, overlapping reads) were screened for sequence contamination in the National Center for Biotechnology Information (NCBI) mitochondrial and nucleotide databases. Any significant alignments to mitochondrial, host, and/or bacterial sequences were identified and removed. *AfarF1* assembly data is available in the NCBI-Sequence Read Archive under experiment ID SRX349764 (2013) and the VectorBase repository (2013).

#### 3.3.4.2 Wild-type sequence data

Read length and sequence quality of wild-type data were assessed in FastQC v.0.10.1 (Babraham Institute 2011). Illumina sequencing adapters and sequences that did not meet quality parameters were removed with Trimmomatic v.0.30 (Bolger, Lohse, and Usadel 2014). Reads were then mapped to the *Afar*F1 assembly (VectorBase 2013) using Burrows-Wheeler Alignment (BWA) v.0.7.5 (Li and Durbin 2009) and pre-processed with SAMtools v.0.1.18 (Li et al. 2009) and Picard v.1.92 (2013) to ensure compatibility with downstream software. HaplotypeCaller in GenomeAnalysisToolkit (GATK) v.2.5.2 (Van der Auwera et al. 2013) was selected for variant discovery and annotation on the basis of technical sophistication and superiority at recognizing insertion-deletion events (indels).

High quality SNP calls used in population structure analysis were generated by merging the 16 individual datasets and applying filtering parameters to remove variants that did not meet quality criteria, were indels, or were multiallelic SNPs. Using VCFtools v.0.1.15 (Danecek et al. 2011), higher quality filters were applied to the dataset so that only SNPs with a minimum genotype quality of 30, minor allele frequency of 5%, and that were present in every individual were retained. A detailed description of the bioinformatics pipeline, including parameters, is listed in APPENDIX D: .

#### 3.3.5 Population structure

Principal components analysis (PCA) was selected to visualize population structure because of its capacity to transform complex multidimensional data onto a

reduced 2-dimensional feature, which represents the data with the most information about its distribution. Principal components (PCs) are also known as eigenvectors and determine the direction of the reduced dimensional space. Eigenvalues, which explain the variance of the data, determine the magnitude. The merged dataset was first pre-processed through PLINK v.1.9 (Chang et al. 2015) to produce a data format suitable for eigenanalysis in R v.3.2.1 (R Core Team 2014).

Genomic ancestry was computed from maximum-likelihood estimates of population allele frequencies and genotype probabilities in ADMIXTURE v.1.23 (Alexander, Novembre, and Lange 2009). A cross-validation (CV) procedure—that identified the lowest error value for which the model had the best predictive accuracy—was applied prior to selecting the most probable number of ancestral populations ( $K$ ) with which to run the statistical model.

### 3.4 Results

#### 3.4.1 Species identification

All specimens identified in the field as anopheline mosquitoes were molecularly confirmed as *An. farauti* using the described rDNA markers in Beebe and Saul (1995) with the exception of those collected in Tanna. Since *An. farauti* is the sole anopheline species existing in Vanuatu—as identified by previous entomological and genetic surveys (Foley, Meek, and Bryan 1994; Beebe and Saul 1995; Beebe et al. 2000; Beebe and Cooper 2002; Reiff et al. 2007)—field

identification through proboscis morphology (Rozeboom and Knight 1946) was sufficient to confirm specimens to species.

### **3.4.2 *An. farauti* reference genome assembly**

The new *AfarF1* genome was 180,984,331 bp in size, assembled into 550 scaffolds with an N50 of 1,196,527 bp (VectorBase 2013). In comparison to other species sequenced as part of the *An16G* project, this assembly was neither the most contiguous (*An. albimanus*: 204 scaffolds, N50=18,068,499 bp) nor fragmented (*An. maculatus*: 47,797 scaffolds, N50=3,841 bp) (Neafsey et al. 2015).

### **3.4.3 Wild-type *An. farauti* samples**

The 16 wild-type samples were mapped to the newly assembled *AfarF1* reference genome prior to variant calling, then merged globally to generate a final high quality SNP dataset (n=17,757). Summary variant counts for each individual and the global dataset are presented in

TABLE 3.3. Prior to any quality assessment (unfiltered), the Australian samples appear to have the greatest number of variants per individual averaging ~2.3 million, followed by Papua New Guinea (~2.0 million), and the Solomon Islands (~1.5 million). In contrast, Vanuatu samples demonstrated almost 50% less variants per individual (~1.25 million) than Australia. It should be noted here, however, that variants identified by HaplotypeCaller (GATK) do not immediately translate into polymorphisms with biological rationalization. Variant calls can be true SNPs but they can also be structural variants, multiallelic SNPs, insertion-

deletions (indels), and even incorrect base calls, although pre-processing of the data prior to application of HaplotypeCaller should have removed most of the latter.

When first combined the global dataset contained 7,329,580 variants but after the application of filtering parameters that identify lower quality, and redundant (for this study) variants, almost 47% of these were removed generating a new quality assessed SNP dataset (n=3,440,489). SNPs that appeared only once in the dataset (n=883,857) and, therefore, were not representative of a population were removed using the [--maf 0.05] function in VCFtools v.1.15 (Danecek et al. 2011). A final genotype filter (sharedVariantsAllSamples.py) to include only those SNPs present in every individual was applied to ensure completeness (Wiltshire 2019).

TABLE 3.3

VARIANT CALL SUMMARY COUNTS OF RAW AND FILTERED DATASETS FOR THE  
16 WILD-TYPE *ANOPHELES FARAUTI* INDIVIDUALS

<i>Individual</i>	<i>Population</i>	<i>Unfiltered</i>	<i>Quality filtered</i>	<i>+maf005</i> *	<i>+noMissing</i> †
<i>HALETA</i>	Solomon Islands	1,526,693			
<i>HALETA2</i>	Solomon Islands	1,490,287			
<i>HALETA3</i>	Solomon Islands	1,515,197			
<i>HALETA4</i>	Solomon Islands	1,528,798			
<i>MADANG1</i>	PNG ‡	2,056,435			
<i>MADANG3</i>	PNG ‡	2,074,106			
<i>MADANG4</i>	PNG ‡	2,094,455			
<i>MADANG6</i>	PNG ‡	1,987,973			
<i>QLD2</i>	Australia	2,328,473			
<i>QLD3</i>	Australia	2,345,038			
<i>QLD5</i>	Australia	2,347,530			
<i>QLD6</i>	Australia	2,363,182			
<i>TANNA2</i>	Vanuatu	1,552,574			
<i>TANNA3</i>	Vanuatu	1,242,834			
<i>TANNA4</i>	Vanuatu	1,253,126			
<i>TANNA6</i>	Vanuatu	1,253,295			
<i>ALL</i>	Global	7,329,580	3,440,489	2,556,632	17,757

NOTE:

\* *maf005* is the minor allele frequency filter and refers to the parameter in VCFtools v.0.1.15 (Danecek et al. 2011), which removes rare and infrequently occurring SNPs (< 5% of a population) from the dataset

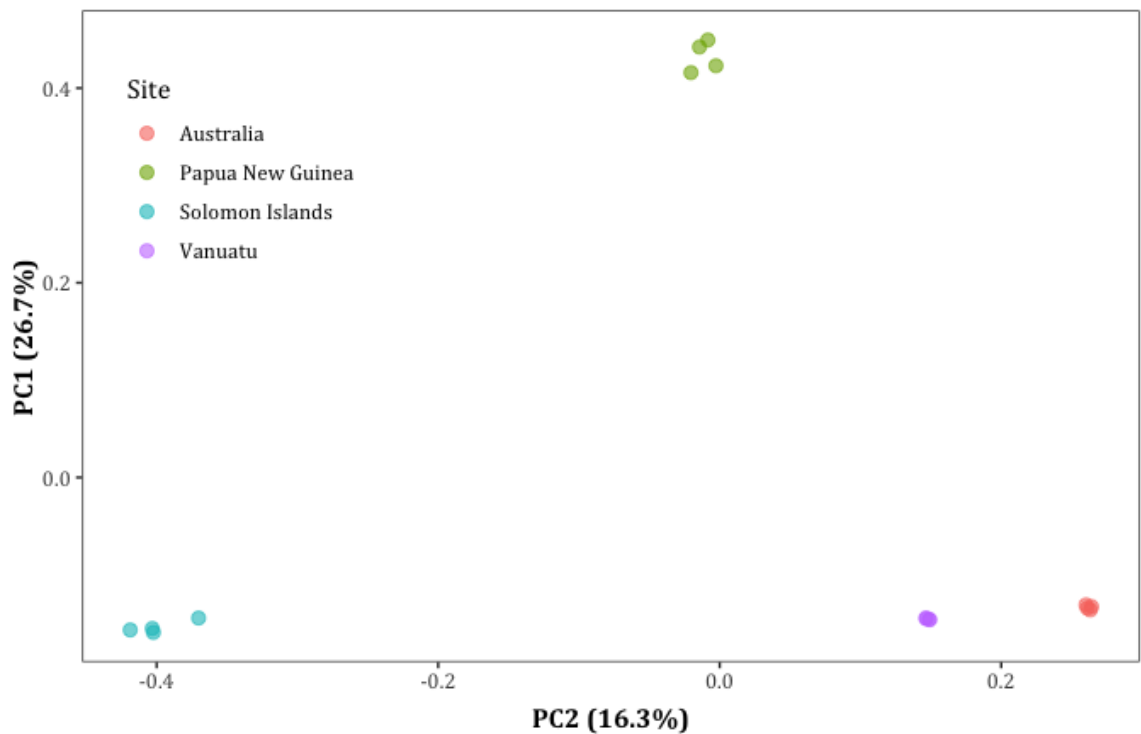
† *noMissing* refers to the filter that removes SNPs, which are not present in every individual of the dataset using the custom Python script sharedVariantsAllSamples.py (Wiltshire 2019)

‡ PNG is an abbreviation of Papua New Guinea



#### **3.4.4 Population structure**

Principal components analysis (PCA) demonstrated that each population (N=4) clustered in concordance with its geographic location (Figure 3.3) as would be expected in sampling sites separated by distances of thousands of kilometers where the presumed barrier to gene flow is an ocean gap. There was a distinct separation between each population with no overlap suggesting that they are genetically isolated. However, the Australia (pink) and Vanuatu (purple) populations are closer together in magnitude implying genetic similarity.



**Figure 3.3 Principal components analysis illustrating genetic variance observed in 17,757 SNPs shared between the wild-type *An. farauti* individuals (N=16) across the four geographical sampling sites.** The y-axis represents the first principal component (PC1) and its percentage variance, and likewise, the x-axis represents the second principal component (PC2) and its percentage variance.

Separation between individuals can be seen as discrete discs in each of the Papua New Guinea (green) and Solomon Islands (blue) groups implying greater variation within these populations than those of Australia and Vanuatu, which in direct contrast appear as single samples with similar eigenvalues.

Population stratification was additionally tested by model-based estimation of global ancestry using the software program ADMIXTURE v.1.23 (Alexander,

Novembre, and Lange 2009). On the assumption that the individuals originated from the same geographical populations in which they were collected ( $K=4$ ), ADMIXTURE assigned four ancestry fractions to the dataset but the resulting distribution was not based on geographical location (Figure 3.4). Instead, individuals from Australia and Vanuatu (black) were observed as one population, Papua New Guinea (sand) as another, and the Solomon Islands (brick and mauve) as two populations. This analysis appears to corroborate that of the PCA, therefore, in which the genetic similarity between the Australia and Vanuatu populations is supported.

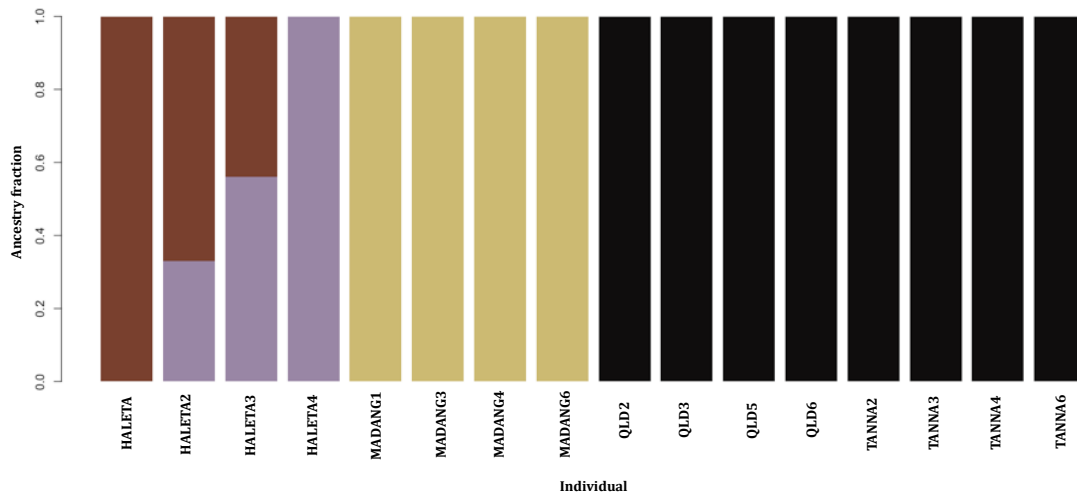


Figure 3.4 **Admixture plot of the 17,757 SNPs illustrating probable ancestry fractions ( $K=4$ ) of wild-type *An. farauti* individuals ( $N=16$ ).** Vertical bars represent individual mosquitoes as listed on the x-axis. Ancestry fractions are assigned according to maximum likelihood probabilities.

Applying ( $K=5$ ) to the model re-stratified the populations geographically (Figure 3.5). Australia (brick) and Vanuatu (grey) demonstrated clear separation as

did Papua New Guinea. The Solomon Islands, however, remained divided into two populations (mauve and black) but with different assigned ancestry fractions to those when  $K=4$ . It should be noted that the palette applied to the admixture plot by R (R Core Team 2014) is arbitrary and there is no qualitative significance to color allocation of populations.

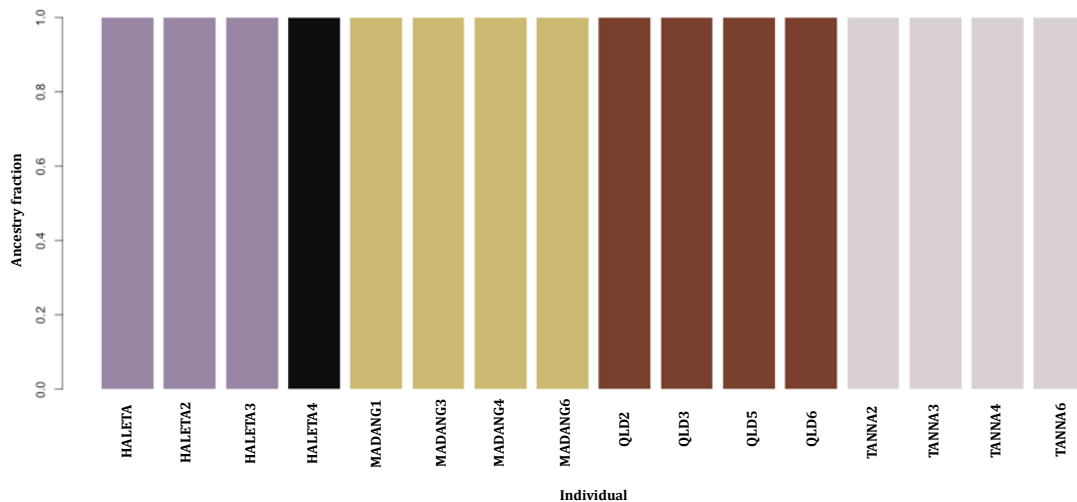


Figure 3.5 **Admixture plot of the 17,757 SNPs illustrating probable ancestry fractions ( $K=5$ ) of the wild-type *An. farauti* individuals ( $N=16$ ).** Vertical bars represent individual mosquitoes as listed on the x-axis. Ancestry fractions are assigned according to maximum likelihood probabilities.

### 3.5 Discussion

The purpose of this study was to identify genetic structure in *An. farauti* populations across its geographic distribution using a genome-wide approach. SNP-based analyses confirmed geographic structuring between populations, and detected a genetic relationship between Australian and Vanuatu populations.

Greater structure identified within the Solomon Islands population implied some limitation to migration. These findings are discussed below.

### **3.5.1 Australia and Vanuatu populations are genetically similar**

PCA (Figure 3.3) and global ancestry assignment (Figure 3.4) inferred that the *An. farauti* populations in Australia and Vanuatu were genetically more similar to each other than to either of the Papua New Guinea and Solomon Islands populations. This relatedness was unexpected since it is likely that physical barriers to gene flow exist between the two sampling sites (Australia: Yorkeys Knob—Vanuatu: Tanna) namely, geographic distance (2,503 km) and the ocean channel across which it extends. It is, therefore, challenging to provide an explanation as to where the similarities arose.

Recent assessments of *An. farauti* population structure in the southwestern Pacific have used mitochondrial DNA (mtDNA) cytochrome oxidase I (COI), and nuclear (ribosomal protein S9 (rpS9) and microsatellite) markers to explain genetic relatedness (Ambrose et al. 2012, 2014) but neither found evidence for an Australia-Vanuatu relationship. Ambrose et al. (2012) identified four distinct geographical groups of *An. farauti* populations based on mtDNA COI haplotypes, one of which being the Solomon Islands together with Vanuatu; however, there were no available ribosomal protein S9 (rpS9) markers from the Queensland (Australia) or Vanuatu samples to enable a nuclear DNA comparison with the SNP data in this study. Further to their 2012 study, Ambrose et al. (2014) again generated a mtDNA COI haplotype network for *An. farauti* populations in the Solomon Islands and

Vanuatu, which identified Tanna (Vanuatu) as being more genetically similar to the northern archipelago populations in Choiseul and Bougainville in addition to the eastern population in Santa Cruz. In contrast, microsatellite analyses with PCA and STRUCTURE (Pritchard, Stephens, and Donnelly 2000) identified Tanna as an unconnected population that was genetically distinct from any of the others sampled across the Solomon Islands archipelago. This finding indirectly lends support to the results of this study with the inference that *nuclear* markers have sufficiently differentiated the Solomon Islands and Vanuatu populations to exclude the suggestion of a southern migration when exploring the mechanism by which *An. farauti* populated the islands of Vanuatu.

*An. farauti* is highly anthropophagic (T. Russell, Beebe, et al. 2016; T. Russell, Burkot, et al. 2016b) in the absence of alternative hosts and it would be reasonable to propose that its existence in such a remote island group is due to human population expansion especially since there is a lack of indigenous land mammals save for a few species of bats (Flannery 1995). The general consensus amongst anthropologists is that migration and colonization of Remote Oceania during the mid-Holocene occurred in the direction of island Southeast Asia, New Guinea, and the Bismark archipelago approximately 3,400 years ago (Anderson and O'Connor 2008; Kayser 2010) mediated, most likely, by voyaging canoe (Anderson and O'Connor 2008), a vehicle more than capable of transporting juvenile stages between landmasses. Although mechanical dispersal is more commonly observed in *Aedes* species (Failloux et al. 1997; Fonzi et al. 2015) than *Anopheles*, this method of distribution provides a mechanism that rationalizes how *An. farauti* migration

across the southwest Pacific could have been achieved. However, on the assumption that certain mosquitoes and humans share a close evolutionary relationship (B. White, Collins, and Besansky 2011) due to the former's blood meal preferences, the proposed human migration route into Remote Oceania does not explain the genetic similarity between the Australia and Vanuatu populations identified in this study. Long distance wind-borne dispersal may be another mechanism by which migration occurs but whose components have yet to be explored in detail.

### **3.5.2 Australia and Vanuatu individuals demonstrate little genomic variation in the generated SNP dataset**

PCA (Figure 3.3) clearly demonstrated that the four individuals in each of the Australia (red) and Vanuatu (purple) populations lacked diversity across the 17,757 SNPs. Each cluster appeared as a single disc indicating that the eigenvalues, which provide a measure of variance quantitatively, were virtually identical for PC1, and PC2. It could be that there is more variation in the lower scoring principal components (i.e. PC3-PC20) for these individuals, or if PCA was run again with only these two populations then variation might be differently distributed due to less noise. To further explore population structure of these individuals, an upstream analysis of each genome could test for relatedness (heterozygosity/runs of homozygosity) in the first instance, which would provide an indication of whether they were truly representative of the population or a sampling bias. Since entomological collections were HLC (Tanna) and CDC light trap (Yorkeys Knob), specimens were adults meaning they were less likely to be immediately related

although, of course, it is entirely possible that they could be. This is in comparison to larval collections where the dipping technique groups larvae, which have developed from eggs most probably laid by a single female and, hence, the resulting adults are related as was the case for the majority of individuals in the Sserinya population discussed in Chapter two. To mitigate sampling bias, additional entomological sampling should be conducted.

### **3.5.3 Greater structure is observed in the Solomon Islands population**

Individuals in the Solomon Islands and Papua New Guinea showed greater variation within each population (Figure 3.3). Global ancestry analysis corroborated PCA and identified further structure in the Solomon Islands as illustrated by the assignment of ancestry fractions in both the  $K=4$  (Figure 3.4) and  $K=5$  (Figure 3.5) models: the latter separating individuals into two discrete populations. A previous MTC entomological survey identified a simple transmission setting at the Haleta sampling location, which is localized since mosquitoes traversed between their oviposition site (a lagoon located at the base of steep-walled volcano) and the village positioned near the shoreline (where they feed) (Neil Lobo, personal communication). It would appear, therefore, that there is limited opportunity for the introduction of genetic material through migration from neighboring populations (if they exist) but this could be supplied via mechanical dispersal across the travel routes between Haleta-Tualagi (the regional hub) in the Nggela Islands and Haleta-Guadalcanal, which have become more frequent.



### 3.6 Conclusion

This is the first genomic assessment of *An. farauti* population structure across the southwest Pacific. Application of a SNP-based approach detected a previously unidentified genetic relationship between populations in Queensland (Australia) and Vanuatu, and also a lack of diversity between individuals within those populations suggesting isolation. Greater structure was identified within the Solomon Islands and Papua New Guinea populations, with the Solomon Islands individuals being separated into two discrete populations by ancestry fraction assignment. In conclusion, the research in this chapter has enhanced previous *An. farauti* population genetic studies with SNP-based approaches that permit high resolution genome-wide analysis.

## CHAPTER 4:

### FIELD EVALUATION OF SUGAR-FERMENTED YEAST AS AN ORGANIC SOURCE OF CARBON DIOXIDE TO ATTRACT *ANOPHELES FARAUTI* MOSQUITOES IN WESTERN PROVINCE, SOLOMON ISLANDS

The purpose of this chapter is to present the reader with research, which contributes expertise towards the second developmental challenge identified by the malERA Consultative Group on Vector Control: interventions that affect vector species not *effectively* targeted by current tools.

In Western Province, Solomon Islands, a proportion of *An. farauti* mosquitoes demonstrate early evening outdoor-feeding behaviors meaning that they avoid entering domiciles where existing indoors-based control tools i.e. LLINs (since IRS operations ceased in 2015) are concentrated thus, avoiding contact with lethal doses of insecticide. These mosquitoes are, therefore, most probably responsible for sustaining residual malaria transmission, which necessitates the development of an outdoors-based intervention that targets the vector through its host-seeking and/or resting behavior.

Carbon dioxide (CO<sub>2</sub>) has often been added to surveillance traps to enhance catch numbers in its role as a chemosensory cue, which the mosquito uses to direct

it towards a source of protein for vitellogenesis. Obtaining an artificial source of CO<sub>2</sub> in the tropics is challenging due to logistical and physical constraints; however, examples of sugar-fermented yeast as an organic source of CO<sub>2</sub> have been successfully demonstrated.

This chapter sought to explore the credibility of sugar-fermented yeast as an organic source of carbon dioxide (CO<sub>2</sub>) trap attractant to host-seeking and resting *An. farauti* mosquitoes in comparison with human-generated CO<sub>2</sub> and a control over a 12-night period in Jack Harbour, Western Province. 653 *An. farauti* mosquitoes were collected in total with the human-generated CO<sub>2</sub> attracting the greatest numbers (n=349) followed by sugar-fermented yeast (n=210), and control (n=94).

In conclusion, the research in this chapter illustrated that sugar-fermented yeast as an organic source of CO<sub>2</sub> was attractive to *An. farauti* mosquitoes in Jack Harbour but improvements to the source design i.e. CO<sub>2</sub> plume composition and delivery could further enhance its appeal.

#### 4.1 Malaria in the Solomon Islands archipelago

The study presented in this chapter was nested within a larger project, The Malaria Transmission Consortium (MTC): Improved Methods to Design and Evaluate Malaria Control Programs, (Bill and Melinda Gates Foundation, grant no. 45114). A long-term goal of the MTC was to develop effective vector control interventions that specifically target the exophagic behavior of malaria-transmitting mosquitoes, which *complement* the existing indoors-based tools—LLINs, insecticide-treated nets (ITNs), and IRS—that are the recommended national, and

international policy worldwide (Solomon Islands National Malaria Control Program 2007; World Health Organization 2017e). The term exophagic is derived from the Ancient Greek suffixes -ἔξω (-éxō) meaning “outer, external”, and -φαγία (-phagía) “the consumption of” or “feeding on”; therefore, exophagic mosquitoes demonstrate a propensity for feeding *outdoors*, a behavior not targeted by LLINs, ITNs, or IRS, meaning that the MTC’s objective was also in alignment with that of the malERA Consultative Group on Vector Control (2011)’s second development challenge.

The application of LLINs and IRS via the Solomon Islands’ control programs of the 1970s was effective in reducing the abundance of *An. punctulatus*, and *An. koliensis* to the point where their contribution to malaria transmission was negligible. During this period, however, *An. farauti* had shifted its peak biting activity from nocturnal and endophagic to crepuscular and exophagic thus, establishing itself as the principle malaria vector of the Solomon Islands (Avery 1973; Paik and Avery 1973; Taylor 1975a, 1975b). The previous long-term reliance on LLINs and IRS as primary vector control tools has had two major effects on malaria transmission in the Solomon Islands: (1) they have selected for a population of exophagic *An. farauti* through pressure on endophagic mosquitoes, which has probably contributed to stable malaria transmission, and (2) their effectiveness in further vector control is now limited since they are unable to reduce *An. farauti* populations below the threshold required to sustain malaria parasite transmission. Therefore, an alternative form of control must be developed that targets their exophagic and exophilic tendencies.

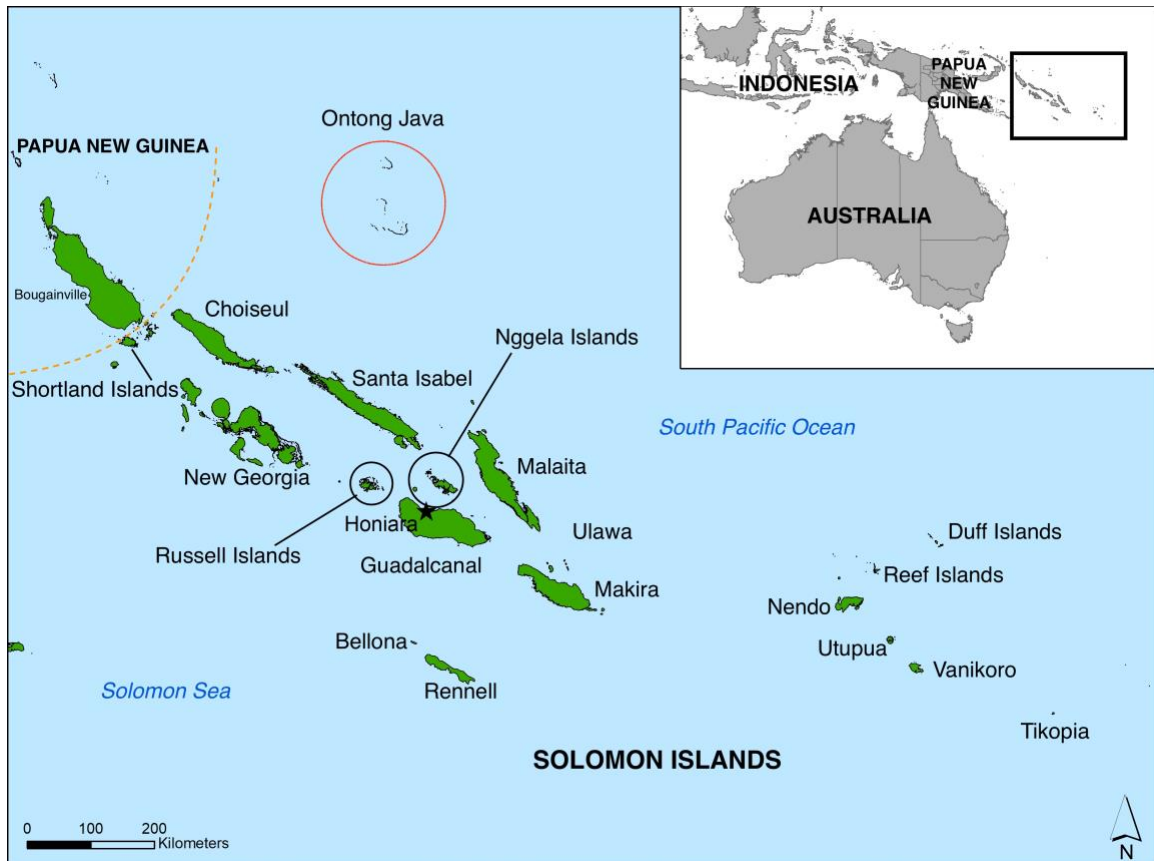
#### **4.1.1 Site selection**

The Solomon Islands was selected as an experimental site by the MTC on the basis of: (1) its commitment to strong working relationships with development partners such as the Japan International Cooperation Agency (JICA), the Pacific Community (SPC), and the Australian government (Australian Agency for International Development (AusAID); Department of Foreign Affairs and Trade (DFAT)), who also acted as a key donor providing major economic support with the Global Fund to Fight AIDS, Tuberculosis, and Malaria (Global Fund) (C. Burkot and Gilbert 2017); (2) cumulatively lower malaria transmission rates (National Vector Borne Disease Control Programme 2013) resulting from the increased distribution of LLINs, and IRS campaigns facilitated by the renewed international assistance, and, (3) preliminary research already conducted in Haleta, Central Province (Figure 3.2) that identified a simple transmission setting whereby a single vector species (*An. farauti*) would traverse between its oviposition site at the base of a volcano, and the village where it blood fed thus, limiting its local distribution.

#### **4.1.2 Physical geography**

The Solomon Islands is a scattered archipelago in the southwest Pacific Ocean, which stretches over 1,700 km between 5-12°S and 152-163°E from the eastern-most point of Papua New Guinea to the northern-most boundary of Vanuatu. The six largest islands (Choiseul, Santa Isabel, New Georgia, Guadalcanal, Malaita, and Makira) form a central double-chain but nearly 1,000 islands comprise

a total land area of 30,407 km<sup>2</sup> located within a greater oceanic region of approximately 1,500,000 km<sup>2</sup> (Cook, McMeniman, and O'Neill 2008) (Figure 4.1).

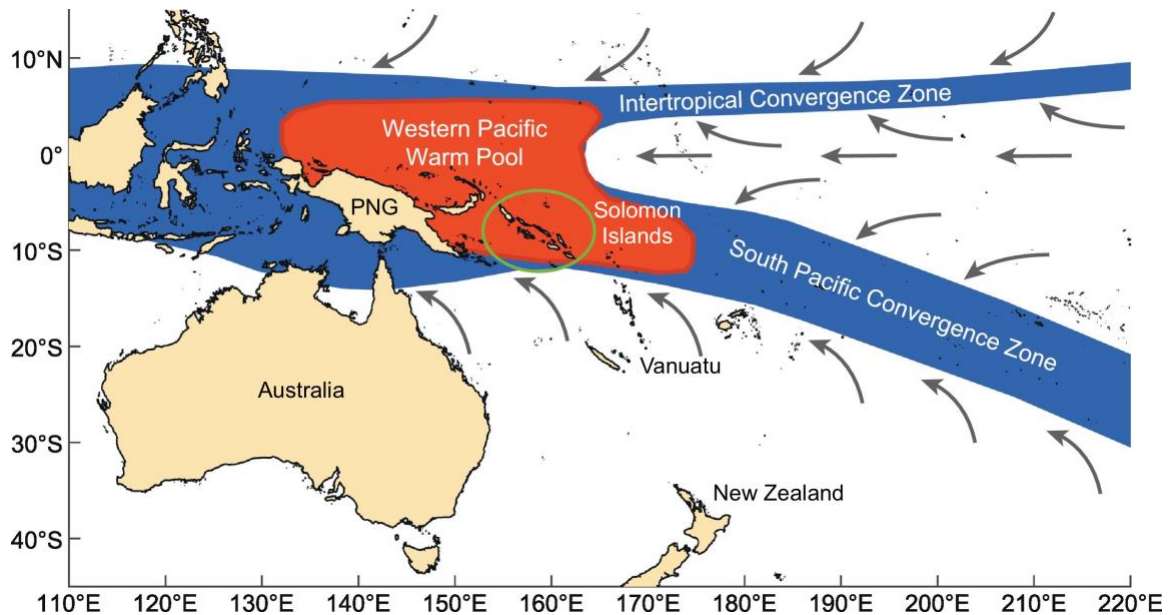


**Figure 4.1 Map of the Solomon Islands archipelago.** Inset top right: the black-framed rectangle illustrates the location of the Solomon Islands within the southwestern Pacific region, and has been magnified to provide detail as shown in the main picture. Ontong Java (circled in red) is a coral atoll with approximately 2,000 inhabitants that is administrated by the Solomon Islands government through Malaita Province.

#### 4.1.3 Climate

The climate of the Solomon Islands is relatively stable with tropical features that rarely fluctuate. The mean daily temperature is 27°C (range: 23-31°C), and humidity is high (74-92%) (Rural Development Division 2001; Solomon Islands

National Statistical Office 2009; Bennett et al. 2014). There are two distinct seasons; however, the austral warm (wet), and cool (dry), which are driven primarily by the changing motions of convective, and baroclinic activities in the two atmospheric circulation features positioned in the region: the Intertropical Convergence Zone (ITCZ), and the South Pacific Convergence Zone (SPCZ) (Rural Development Division 2001; Solomon Islands National Statistical Office 2009; Bennett et al. 2014). The SPCZ is oriented in a northwest to southeast direction across the southwest Pacific landmasses (Figure 4.2). It contains a vast convective cloud band (Vincent 1994) that channels latent heat, and ocean surface moisture to higher latitudes, which influences the low pressure systems, and high precipitation rates that control much of the climate patterns observed in many of the South Pacific islands. The warm season (November-April) occurs when the physical characteristics of the SPCZ are fully developed, and it reaches its spatial maximum (140°W) (Lorrey et al. 2012). This leads to increased convective activity, which generates stormy, west-to-northwesterly monsoonal winds, and heavy rainfall that continue until the cool season begins (May-October) when the SPCZ loses convective strength (as it contracts towards Papua New Guinea), and southeasterly trade winds prevail creating stable airstreams with reduced precipitation (Lorrey et al. 2012).



**Figure 4.2 Approximate positions of the Intertropical and South Pacific Convergence Zones during the austral warm season (November-April).**

Black arrows depict near surface winds. The green circle illustrates the Solomon Islands. Papua New Guinea is described as PNG. Colored areas represent landmasses (yellow), bands of rainfall (dark blue), and the Western Pacific Warm Pool (WPWP) (red), which is a unique climatic feature with a permanent sea surface temperature  $> 28^{\circ}\text{C}$  giving it the capability to modify heating and cooling of the tropical Pacific Ocean. The y-axis depicts latitude in degrees, and the x-axis, likewise, longitude. Adapted from Jaffrés et al. 2018.

#### 4.1.4 Sociodemography

Politically, the Solomon Islands is a sovereign state having gained independence from British administration in 1978. The most recent census data (2009) recorded the total population as 515,870 (1999: 409,042)—an estimated annual growth rate of 2.3% (1999: 2.8%)—living in mostly rural residences (80.2%) (1999: 84.4%) on the periphery of islands (TABLE 4.1) (Solomon Islands National Statistical Office 2009). Governance is decentralized at the national level, and administered locally in nine provinces: (1) Central; (2) Choiseul; (3) Guadalcanal; (4) Isabel; (5) Makira-Ulawa; (6) Malaita; (7) Rennell and Bellona; (8)



Temotu; (9) Western, plus a tenth administrative area, (10) Honiara town council (Figure 4.3;). Local governments are responsible for municipal, and provincial services while executive authority remains with the national government holding office who oversees state legislation on key development matters such as education, health, and law etc.



Figure 4.3 **Map of the Solomon Islands' administrative provinces.** Areas are demarcated by grey dotted lines. Honiara (marked by the black star), is also an administrative province but too small to mark as such at this magnification.

TABLE 4.1

HUMAN POPULATION SIZE, DENSITY, DISTRIBUTION, AND MEAN ANNUAL GROWTH RATES FOR THE TEN ADMINISTRATIVE PROVINCES OF THE SOLOMON ISLANDS (2009)

<i>Province</i>	<i>Total population</i>	<i>Land area (km<sup>2</sup>)</i>	<i>Population density *</i>	<i>Population distribution (%)</i>	<i>Mean growth rate (% p.a.) §</i>
<i>Solomon Islands</i>	515,870	30,407	17	100	2.3
<i>Central</i>	26,051	615	42	5	1.9
<i>Choiseul</i>	26,379	3,837	7	5	2.8
<i>Guadalcanal</i>	93,613	5,336	18	18	4.4
<i>Isabel</i>	26,158	4,136	6	5	2.5
<i>Makira-Ulawa</i>	40,419	3,188	13	8	2.6
<i>Malaita</i>	137,596	4,225	33	27	1.2
<i>Rennell-Bellona</i>	3,041	671	5	1	2.5
<i>Temotu</i>	21,362	868	25	4	1.2
<i>Western</i>	76,649	7,509	10	15	2.0
<i>Honiara</i>	64,602	22	25	12	2.7

SOURCE: (Solomon Islands National Statistical Office 2009)

\* Population density is measured as individuals/km<sup>2</sup>

§ Mean growth rate refers to the *average* percentage population increase per annum (p.a.) across the 10-year period in between national census surveys 1999-2009

The economy is based largely on subsistence agriculture supplemented by cash cropping (e.g. cocoa; palm oil; copra; betel nut; and vegetables), fishing, forestry, and mining. The reliance on unsustainable logging together with international aid, however, leave little prospect for inclusive economic growth threatening the government's ability to maintain the above average level of healthcare that it has been delivering to its population, and, ultimately, attempts to eliminate malaria.

Malaria control in the Solomon Islands is implemented by the National Vector Borne Disease Programme (NVBDP), which was integrated into the Ministry of Health and Medical Services in 1992 (National Vector Borne Disease Control Programme 2013). Key intervention strategies are targeted against the (i) *vector* through the distribution of LLINs/ITNs (IRS having been discontinued in 2015 due to limited resources and technical challenges) (National Vector Borne Disease Control Programme 2016), and (ii) *parasite* through the administration of Coartem® (Novartis, Switzerland)—an artemether-lumefantrine (AL) antimalarial drug—prescribed as the first line course of treatment for uncomplicated and severe *P. falciparum* infections. Parenteral artesunate (AS) is also administered as a first line treatment for severe malaria while 7-day quinine (QN) monotherapy is a second line. *P. falciparum* infections in pregnancy are treated with a weekly dose (5 mg/kg body weight) of chloroquine (CQ). AL is also the recommended course of treatment for vivax malaria together with a 14-day radical cure of primaquine (PQ) (National Vector Borne Disease Control Programme 2013, 2016; World Health Organization 2018d).

#### 4.1.5 *Plasmodium* species prevalence

The four major *Plasmodium* species that cause pathogenicity in humans—*P. falciparum*, *P. vivax*, *P. ovale*, and *P. malariae*—are present in the Solomon Islands but *P. knowlesi* has never been reported, most probably because the country is beyond the range of the parasite's reservoir host (macaque monkeys) (K.-S. Lee et al. 2011), and its competent vectors, *An. latens* (Vythilingam et al. 2006), and *An. cracens* (Vythilingam et al. 2008).

As *P. vivax* (2017: 30,169) has now surpassed *P. falciparum* (2017: 15,400) as the prevalent malaria species in the Solomon Islands (World Health Organization 2018d), a strategic change is required (due to differences between the parasites' cell biology) in the approach to control taken by the NVBDP, and the Ministry of Health and Medical Services as they attempt to advance towards elimination. While it is important to maintain adoption and practice of the currently recommended malaria control policies (National Vector Borne Disease Control Programme 2013; World Health Organization 2018d), it is crucial to supplement these activities with active case detection and surveillance of the *P. vivax* hypnozoite reservoir. This is because of its ability to remain quiescent within host hepatocytes then spontaneously recrudesce presenting a new parasite population for ingestion by *Anopheles* mosquitoes thus, maintaining *Plasmodium* transmission. Radical treatment for *P. vivax* infections is the 8-aminoquinolone compound, primaquine (PQ). However, this must be prescribed by a physician since the consequences of indiscriminate administration can be catastrophic due to the presence of a genetic variant in the population, which can cause acute hemolytic anemia (AHA) if the individual is a

carrier of the variant, and their PQ dosing schedule is not managed correctly. At present, national services for providing radical vivax malaria treatment are limited due to the lack of an accessible, cost-effective, rapid diagnostic test, which identifies variant carriers prior to treatment (National Vector Borne Disease Control Programme 2016).

#### **4.1.6 Glucose-6-phosphate dehydrogenase deficiency**

As a result of *P. vivax* becoming the dominant malaria species in the Solomon Islands, and also its radical treatment acting as an exogenous trigger of AHA in certain individuals, it is important to introduce a common hemoglobinopathy that has the capacity to seriously interfere with attempts to eliminate malaria.

Hemoglobinopathies are a congenital group of blood disorders resulting from genetic mutations in globin genes, which affect hemoglobin (Hb) production by one of two mechanisms: (1) downregulation of protein expression during globin chain synthesis, leading to a quantitative reduction in Hb that usually presents clinically as anemia (*thalassemias*), and (2) more serious qualitative reductions that arise from malformations in the polypeptide subunits of the Hb tetramer itself (*structural variants*) since they affect molecular function according to the chemical, and physical properties of the nucleic acid modification.

Glucose-6-phosphate dehydrogenase deficiency is a structural variant hemoglobinopathy caused by polymorphic point mutations in the glucose-6-phosphate dehydrogenase (G6PD) gene. G6PD is a critical enzyme in the erythrocytic environment that catalyzes the first, and rate-limiting, step in the

pentose phosphate pathway (PPP), an alternative metabolic process for glucose catabolism. As G6PD oxidizes glucose-6-phosphate to 6-phosphogluconolactone, the reduced form of nicotinamide adenine dinucleotide phosphate (NADP<sup>+</sup>)—*NADPH* + H<sup>+</sup>—is generated. The primary function of NADPH is to protect the cell from oxidative stress through the reduction of glutathione (GSH), which becomes oxidized (GSSG) when glutathione peroxidase utilizes it as a substrate while reducing peroxides to water. NADPH is then required by glutathione reductase to replenish functional cellular levels of GSH through the reduction of GSSG (Mathews, van Holde, and Ahern 2000). Since erythrocytes lack mitochondria, the PPP is their only source of NADPH for reductive biosynthesis, and a deficiency of G6PD levels or functionality will result in catastrophic rupture (*hemolytic crisis*) due to the increasing concentrations of free radicals, and peroxides accumulating in the cytosol.

G6PD-deficient (G6PDd) individuals present a wide range of clinical symptoms according to the level of enzymatic activity that the variant allele confers on the protein (TABLE 4.2). Hemolyzing triggers include certain foods, chemicals, infection, and pharmacological compounds such as the antimalarial drugs, primaquine (PQ) and chloroquine (CQ) whose antioxidant activities can induce AHA. G6PDd individuals exposed to PQ (and to a lesser extent CQ, which is a weak antioxidant) are, therefore, at greater risk of provoking AHA depending on their G6PD variant classification. Since hemolytic severity is unpredictable, it is essential that an individual suspected of carrying a G6PDd variant be screened prior to PQ or CQ administration.

TABLE 4.2

PHENOTYPIC CLASSIFICATION OF GLUCOSE-6-PHOSPHATE DEHYDROGENASE  
VARIANTS BASED ON ERYTHROCYTE ENZYMATIC ACTIVITY AND ASSOCIATED  
CLINICAL OUTCOMES

<i>Phenotype</i>	<i>Enzyme activity (%)</i>	<i>Type of hemolytic anemia</i>
I—Chronic	0	Non-spherocytic
II—Severe	1-10	Intermittent
III—Moderate	10-60	Stressor-induced
IV—Normal	60-150	None
V—Increased	>150	None

SOURCE: (WHO Working Group 1989)

More than 222 G6PD variants have been identified globally (Minucci et al. 2012; Capoluongo 2018) with the Solomon Islands estimated as having one of the highest national allele frequencies (22.3%) (Howes et al. 2012).

#### 4.2 Mosquito sensory physiology

Mosquitoes use a range of mechanical, and sensory stimuli including visual cues, moisture, heat, carbon dioxide (CO<sub>2</sub>), and volatile organic compounds (VOCs) to locate a source of protein for vitellogenesis. Olfaction is the primary sense perception, which facilitates host-seeking behavior in response to the detection of low levels of CO<sub>2</sub> (4-5%) (Gillies 1980) in exhaled human breath. Advanced electrophysiological studies incorporating neuronal recordings established that olfactory receptor neurons (ORNs) (positioned on the peg sensilla of the maxillary

palps (Figure 4.4)) modulated highly-specialized excitatory responses when stimulated by increasing concentrations of CO<sub>2</sub> (Kellogg 1970; Grant et al. 1995).

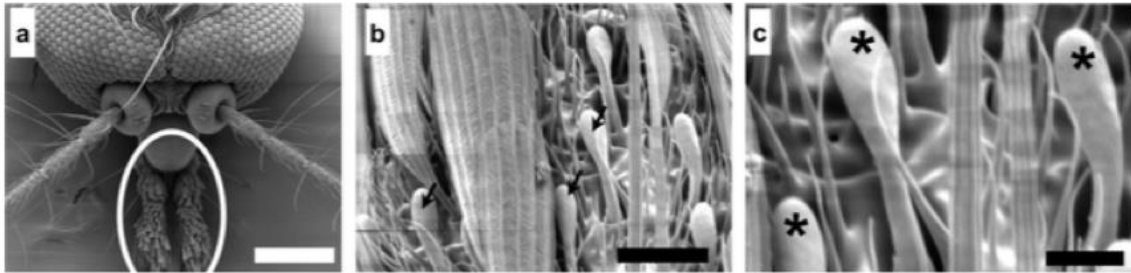


Figure 4.4 **Scanning electron micrographs of *Culex quinquefasciatus* head, and olfactory structures.** (a) Maxillary palps are circled in white. (b: arrows) and (c: asterisks) illustrate individual peg sensilla at different magnifications. Scale bars: (a) 250  $\mu\text{m}$ , (b) 100  $\mu\text{m}$ , and (c) 5  $\mu\text{m}$ . From (Syed and Leal 2007).

Reeves (1951) first demonstrated the attractiveness of CO<sub>2</sub> to female mosquitoes under field conditions. As a result of its influence on the host-seeking response in mosquitoes, CO<sub>2</sub> has often been added to traps to enhance their attractiveness to, and increase catch sizes during surveillance and/or sampling programs. Several studies have compared the attractiveness of CO<sub>2</sub> derived from different sources such as sugar (Saitoh et al. 2004; Smallgange et al. 2010), and molasses (Mweresa et al. 2014). While results vary between the attractiveness of industrial CO<sub>2</sub>, dry ice, and varying combinations of sugar-fermented yeast and molasses-fermented yeast, all demonstrated that CO<sub>2</sub>-enhanced traps collected greater numbers of mosquitoes than traps that did not contain CO<sub>2</sub>.



### 4.3 CO<sub>2</sub> plume structure and dynamics

The spatial, and temporal distributions of the CO<sub>2</sub> plume are critical when generating a credible attractant to mosquitoes. Studies exploring how the concentration, and fine-scale plume structure of host odors (including CO<sub>2</sub>) influenced upwind flight of *Aedes aegypti* (Geier, Bosch, and Boeckh 1999), and *An. gambiae* (Dekker, Takken, and Cardé 2001) showed a distinct preference by both species for the increasing fluctuations in concentration that occur in turbulent, and filamentous odor plumes.

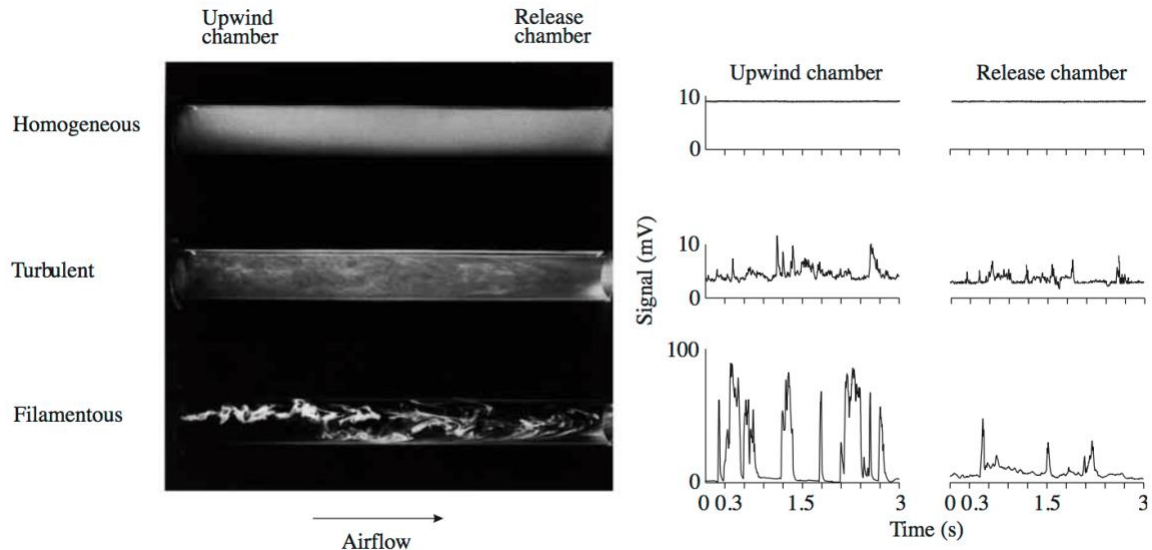


Figure 4.5 **Visualization of odor plumes using TiCl<sub>4</sub> smoke.** The picture on the left illustrates the different types of plume that can be generated from an odorant stimulus. Voltage outputs on the right measured changes in smoke density as a demonstration of how plume type could influence mosquito flight choice towards an odorant stimulus during upwind flight. From (Geier, Bosch, and Boeckh 1999).

### 4.4 Current objective

The objective of this study was to provide preliminary data for the MTC's long-term goal of developing effective vector control interventions—which

specifically target the exophagic behavior of malaria vectors—by evaluating sugar-fermented yeast as an organic source of CO<sub>2</sub> to attract *An. farauti* mosquitoes under field conditions in Western Province, Solomon Islands. Although the study did not target exophagic behavior directly (since it did not measure biting density), it did so indirectly by exploiting the host-seeking route, which mosquitoes took between their assumed oviposition/resting sites and the domiciles where human hosts were active (indoors and outdoors).

#### 4.5 Materials and methods: study site

##### 4.5.1 Physical geography

The study was performed in the village of Jack Harbour (8°03'35"S, 157°11'45"E) on Kolombangara in the New Georgia island group, Western Province (Figure 4.6) (T. Russell, Burkot, et al. 2016a). Alternate names for Jack Harbour that can be found in the literature are Bambari Harbour, Bamberi Harbour, and Mbambare Harbour. GPS coordinates were originally recorded in decimal degrees as -8.059792, 157.195782 (T. Russell, Burkot, et al. 2016a) and converted into degrees-minutes-seconds using the NGS Coordinate Conversion and Transformation Tool (National Geodetic Survey 2018).



Figure 4.6 **Map of the New Georgia group of islands.** Inset top right: The thick red-framed rectangle illustrates Western Province's location within the Solomon Islands archipelago.

Geologically, Kolombangara is a dormant Pleistocene stratovolcano (Smithsonian Institution 2018), Mount Veve, measuring 15 km<sup>2</sup> in diameter, which gently increases in elevation from the shoreline across a flat coastal plain to a height of 700 m ASL at the base of the central cone where the topology abruptly steepens to 1,760 m ASL at the crater rim (Katovai 2016; Smithsonian Institution 2018). Much of the island's low-lying landscape (0-400 m ASL), which was once the domain of tropical rainforests, has now been fragmented by human-mediated activities such as logging and agricultural practices i.e. coconut plantations (Katovai, Burley, and Mayfield 2012). It is in these areas around the circumference of Mount Veve where the majority of the island's inhabitants reside in intermittently located settlements.

#### **4.5.2 Climate**

Climate on Kolombangara is similar to that of the Solomon Islands generally with annual mean temperatures of 26.8-28.0°C recorded monthly (Munda: 8.33°S, 157.27°E) between 2000-2013 (Australian Government Bureau of Meteorology 2018). Kolombangara translates as “Water King” in the local Kolei language, which is evident by the mean annual rainfall of 2,900-4,250 mm/year (Aldrick 1993; Rural Development Division 2001; Australian Government Bureau of Meteorology 2018) that is distributed bimodally in two rainy seasons from November-March, and July-August (Figure 4.7) (Wairiu and Lal 2003).

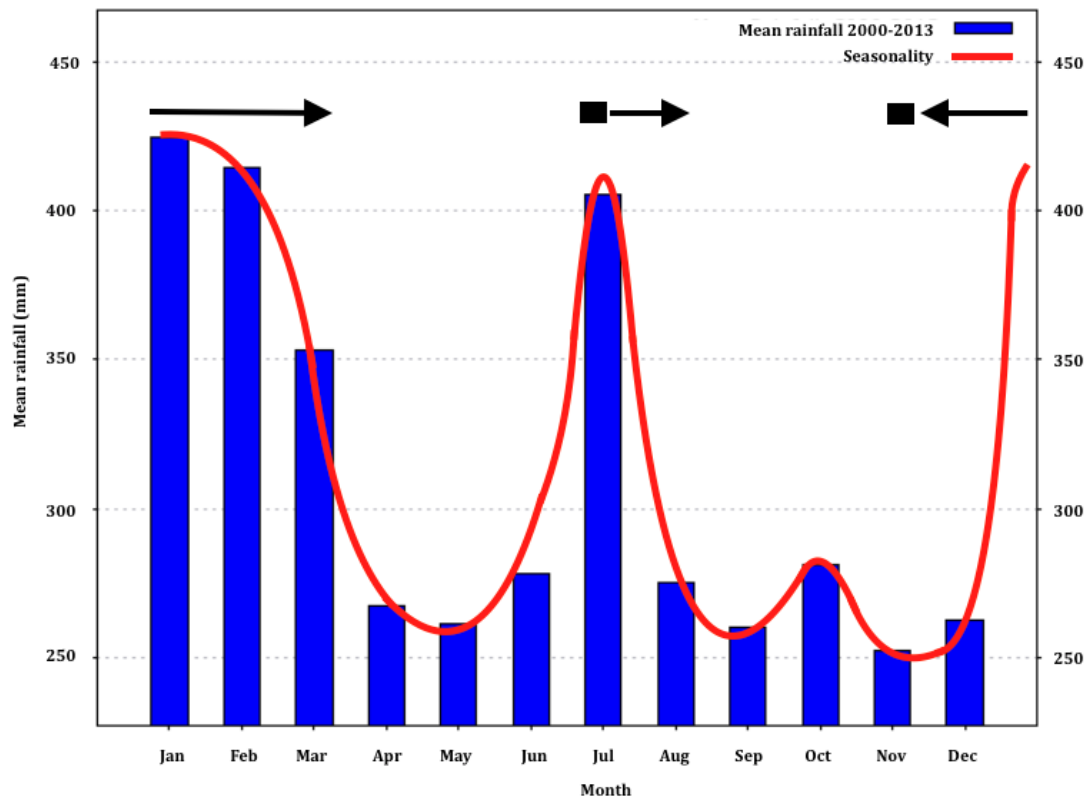


Figure 4.7 **Monthly rainfall data recorded at Munda, Western Province, Solomon Islands (8.33°S, 157.27°E) from 2000-2013.** Mean rainfall is recorded in mm, and represented by the blue bars. A red seasonality trend line has been fitted to the data to illustrate the two rainy seasons, November-March and July-August, which are marked by the black arrows above the graph. Rainfall data was obtained from the Pacific-Australia Climate Change Science and Adaptation Planning Program (Australian Government Bureau of Meteorology 2018).

The magnitude by which the SPCZ influences seasonal climatic activity over the region cannot be understated. Mean precipitation rates generated during the warm (rainy) season in the period 2000-2013 were recorded at 150-300 mm/month, which drastically reduced to 20-145 mm/month once the SPCZ contracted signaling the start of the cool (dry) season (Australian Government

Bureau of Meteorology 2018). Transitions in climatic activity also influence seasonal patterns in malaria transmission. The peak biting season (April-June) identified for *An. farauti* populations in Jack Harbour (T. Burkot et al. 2018) immediately followed retraction of the SPCZ (March) from its physical and spatial maximum. It is suggested that a reduction in convective activity allowed transient rainfall-dependent oviposition sites to form (such as vehicle tracks, footprints, pig wallows, and ground pools), which increased vector abundance due to the absence of flow intensity that usually arrests development through the destruction of eggs and immature stages via flooding.

#### **4.5.3 Sociodemography**

Western Province is the largest of the ten provinces extending across a fragmented landmass with an area of 7,509 km<sup>2</sup> (Figure 4.6). Its primary municipality is Gizo Town, which acts as a regional hub for major health, trading, tourist, and transport services. The total population of Western Province is 76,649 (TABLE 4.1) living in 13,762 households (Solomon Islands National Statistical Office 2009).

Distribution of LLINs is the primary vector control intervention across the Solomon Islands. By the end of 2016, Western Province had received 88% (n=56,411) of its planned allocation covering 15,663 households (National Vector Borne Disease Control Programme 2016). Despite long-term education and communication efforts (National Vector Borne Disease Control Programme 2013, 2016), LLIN/ITN use is still incomplete. In a recent entomological survey conducted

in Western Province, only 68% of study residents self-reported sleeping under LLINs/ITNs (2018).

#### **4.5.4 Malaria vectors**

MTC entomological surveys conducted between March 2014 and August 2016 captured four anopheline species—*An. farauti*, *An. hinesorum*, *An. lungae*, and *An. solomonis*—across 11 sites in Western Province with *An. farauti* identified as the exclusive spatio-temporal species in Jack Harbour. The two surveys conducted prior to this study (June and September 2014), and the seven subsequent surveys that followed (December 2014, February 2015, August 2015, December 2015, January 2016, May 2016, and August 2016) also supported this finding (T. Burkot et al. 2018). Further evidence documenting *An. farauti* as the sole anopheline species in Jack Harbour was obtained from larval habitat surveys (T. Russell, Burkot, et al. 2016b). Neither *An. punctulatus* nor *An. koliensis* were identified from the collections conducted during the MTC entomological surveys of Western Province.

#### **4.5.5 Transmission setting**

The transmission setting was similar to that identified by a previous MTC study conducted in Haleta, Central Province (Figure 3.2, Figure 4.3) whereby mosquitoes simply traversed between their oviposition and blood-feeding sites while seeking hosts. Previously observed exophagic behavior in the Jack Harbour *An. farauti* populations (Neil Lobo, personal communication) identified two swamps as probable oviposition sites (Figure 4.8). It was theorized that mosquitoes would traverse an open grassed area (green rectangle) in between the large swamp

(yellow outline) and main village (orange outline) as part of their host-seeking route. To maximize catch numbers this was the proposed trap placement.

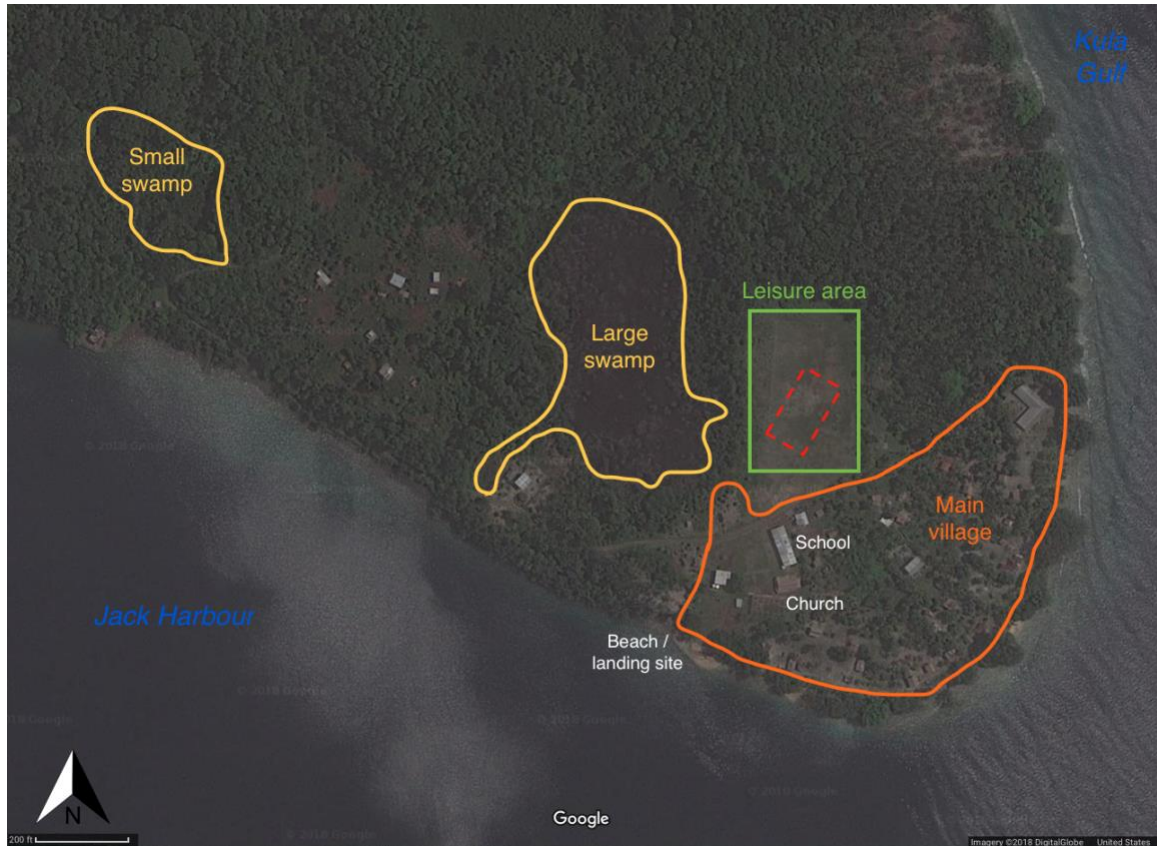


Figure 4.8 **Satellite image of Jack Harbour village illustrating major entomological and demographic features of the transmission setting.** The yellow outlines highlight swamp areas assumed to be *An. farauti* oviposition sites. The orange area encompasses the main village. In between the large swamp, and main village is a wide expanse of open grassed area (green rectangle) used mostly by the villagers for leisure activities such as soccer. The red-dashed rectangle demarcates the proposed placement of the carbon dioxide baited-traps. Notable structures are described accordingly. Map data: Google, 2018.



## 4.6 Materials and methods: trap structure

### 4.6.1 CO<sub>2</sub> sources

The selection of CO<sub>2</sub> source for the non-human bait was limited by the location of the study site. As part of the Solomon Islands archipelago, Western Province is accessible by air (daily flights to the capital city, Honiara, located 380 km east-southeast on the island of Guadalcanal) and water (weekly passenger boat service between Gizo and Honiara via Marovo) (Figure 4.6) (Visit Solomons 2018). Jack Harbour can only be reached by private boats running outboard motors. It would be unfeasible and cost-prohibitive to transport, recharge, and maintain the artificial sources of CO<sub>2</sub> (i.e. dry ice, pressurized gas cylinders or propane), which are routinely applied as trap attractants in usual experimental circumstances.

Previous reports evaluating the efficacy of organically produced carbon dioxide-baited traps using sugar (Smallgange et al. 2010; Meyer Steiger, Ritchie, and Laurance 2014) or molasses (Mweresa et al. 2014) as a yeast-fermenting substrate were reviewed for baseline information to select the most appropriate values for a reference treatment. The mean output (ml/minute) that most closely mimicked human CO<sub>2</sub> release (250 ml/minute) (Mboera and Takken 1997) was produced by Smallgange et al. (2010) who combined 17.5 grams (g) dry yeast with 500 g of refined household sugar and 2.5 liters (L) of water. This solution generated a mean CO<sub>2</sub> production rate of 242.3 ml/minute at 22-25°C, which endured for a collection period beyond that required for this study. CO<sub>2</sub> flow ceased 51 hours post-mixing with catches recorded up to 48 hours although rates decreased over time.

CO<sub>2</sub> was generated by one of two sources: (1) combining 500 g of sugar (refined white table sugar sourced locally in Gizo) + 17.5 g of *Saccharomyces cerevisiae* (Fleischmann's Instant Dry Yeast, Canada) + 2.5 L of water (sourced locally from the river spring, which the Jack Harbour residents used for drinking), or (2) a human volunteer who was protected from potential bite activity within the trap by an untreated portable mosquito tent (Figure 4.9). It was important for the net material to be untreated to eliminate the possibility of an excito-repellent response occurring, which might have confounded the study. Excito-repellency is defined as a kinetic motion elicited in response to a chemical that causes insects to disperse away from the area of the chemical more rapidly than if the area did not contain the chemical (Dethier, Browne, and Smith 1960), and its effects in malaria vectors have previously been demonstrated (Kawada et al. 2014).

#### *4.6.1.1 Control*

Traps that acted as a control were devoid of any CO<sub>2</sub> source i.e. Figure 4.9 without the portable mosquito tent that protected the human volunteer.



Figure 4.9 **Portable mosquito tent used by the human volunteer to protect against potential bite activity during the experiment.** Entry and exit are accessed through the zippered door on the lateral aspect.

#### 4.6.2 Trap design and assembly

A 1.8 m (length) x 1.8 m (width) x 0.9 m (height) trap was constructed from commercially available untreated tulle netting in a square configuration (Figure 4.10). Each corner was wrapped around a wooden pole fixed into the ground and secured with binder clips. At the top of the trap around the entire perimeter, an excess piece of netting measuring 15 cm in height was elevated to a fixed 45° angle to create an eave that captured the mosquitoes as they flew upwards. By creating eaves on each lateral aspect of the trap, the cardinal direction of the mosquitoes'

flight path could be identified, which would potentially identify the type of behavior that was being stimulated by the CO<sub>2</sub> cue i.e. host-seeking or resting after a blood meal since the transmission setting was anticipated to be a simple traversal between the known positions of the oviposition and feeding sites (Figure 4.8). The ends of the eaves were sealed with the same netting to: (1) enable identification of cardinal flight direction, and (2) prevent mosquitoes from leaving prior to collection although they do not appear to initiate downward flight once inside an eave.

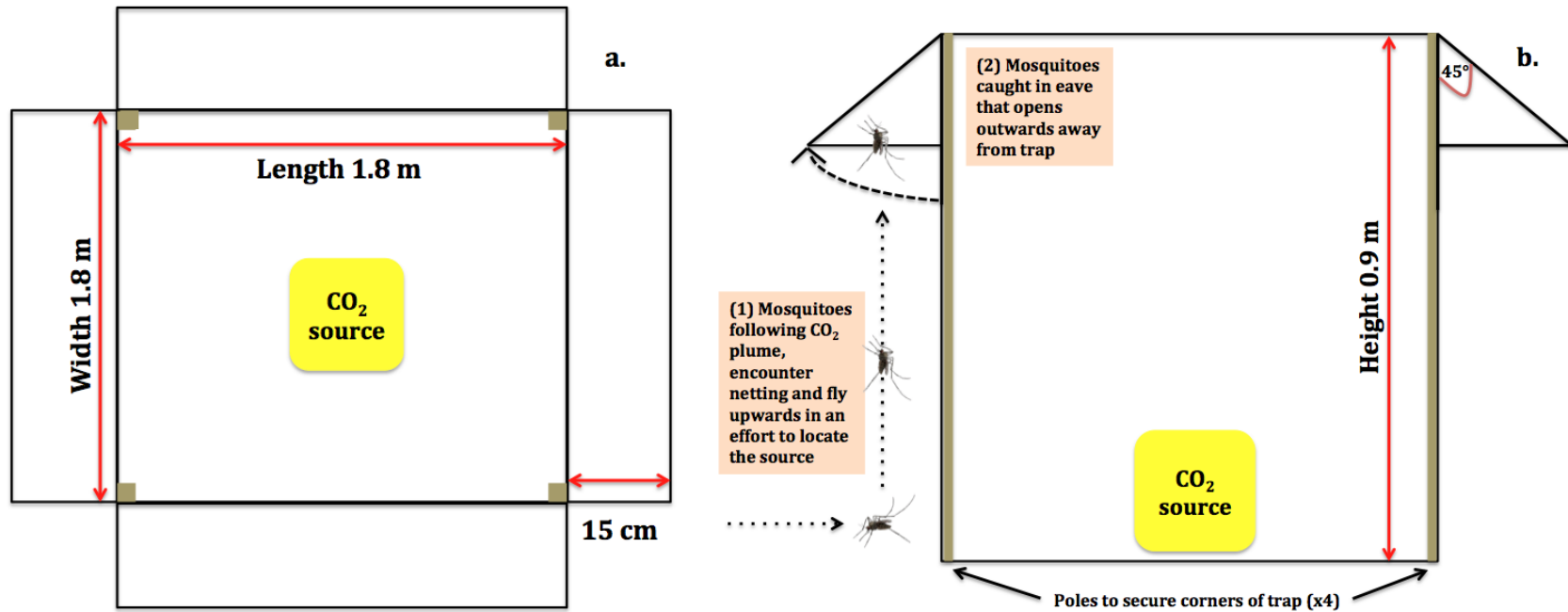


Figure 4.10 **Schematic diagrams of the trap design.** Length and width measurements presented from (a) aerial, and (b) lateral aspects, which also demonstrates entomological rationale. Female mosquitoes sense a CO<sub>2</sub> source (yellow square), and are guided by its plume to a potential blood meal. (1) Upward flight is initiated upon encountering the net material, which ceases upon (2) capture by the eave structures.

The top of the trap remained open where mosquitoes were free to enter and leave (Figure 4.11).



Figure 4.11 **Mosquito collectors assembling a trap.** Visible are the four corner poles used to frame the netting that formed the trap construct, which was held in place at a height of 0.9 m by mason line as indicated by the orange arrow. In the foreground is excess material, which when extended upwards 45° formed the eaves. These were secured in place using Velcro, and mason line for stability.

## 4.7 Materials and methods: study design

### 4.7.1 Experimental model

Since a single treatment (CO<sub>2</sub> bait) was being tested, a 3x3 Latin square experimental design was appropriate and allocated according to the blocking factors: (1) trap position, and (2) number of nights. Baits were rotated between trap positions each evening to control for serial bias. An experimental replicate was complete when each treatment (CO<sub>2</sub> bait) had been measured at each trap position i.e. at the end of three nights. As the 3x3 Latin square produced only nine experimental units, the degrees of freedom were too few for error (because the number of treatments were small); therefore, each replicate was repeated four times to increase statistical power. This equated to twelve nights of collection for the study period.

### 4.7.2 Statistical analysis

The software package R v. 3.2.1 (R Core Team 2014) was used to analyze the dataset statistically. One-way analysis of variance (ANOVA) was selected as the most robust statistical test of significance with which to analyze the dataset as it allowed the simultaneous comparison of multiple group means i.e. (1) control, (2) sugar-fermented yeast-CO<sub>2</sub>, and (3) human-generated CO<sub>2</sub> with which to identify any effect. Tukey's Honestly Significant Difference (HSD) *post-hoc* test was applied to detect the group(s) difference(s) contributing to that effect.



### 4.7.3 Entomological sampling

Three collectors were recruited from a pool of village residents once the study protocol, and associated risks were explained. Each received basic training on the sampling technique of oral aspiration and cup collection followed by mosquito recognition (to discriminate from other entomological Orders i.e. Lepidoptera), and assignment to a particular baited-trap for the duration of the experiment. Collectors were instructed not to approach traps (except for aspirating) so as not to bias the study by attracting mosquitoes to themselves and hence, the traps. Residents, and other individuals not involved in the study, who gather outdoors to eat and socialize (as is the norm from approximately 17:00-21:00 hours in Jack Harbour) were also asked to remain clear of the traps, and the immediate area for the same reason.

Collections began at 18:00 hours and ended at 00:00 hours to match the *An. farauti* outdoor biting profile identified by previous MTC entomological surveys (T. Burkot et al. 2018). At the beginning of the hour, a paper cup (with a waxed interior covered by netting, which was secured by an elastic band) that acted as a collecting vessel was placed at each cardinal direction (N, E, S, and W) around the trap, and also inside. Collectors inspected eaves, sides, and interior of the trap every fifteen minutes aspirating mosquitoes, and placing them in the appropriate collection cup. At the end of each hour, the cups were removed, and a new set positioned as before. Cups were identified by bait (color-coded: (1) control-blue, (2) yeast-yellow, and (3) human-red), and collection times (in hours: 6-7pm, 7-8pm, 8-9pm, 9-10pm, 10-11pm, 11-12pm) to ensure accuracy during catch counts.



#### **4.7.4 Species identification and preservation**

*An. farauti* mosquitoes were sexed, and morphologically identified to species based on the taxonomic keys of Belkin (1962a, 1962b), and Rozeboom and Knight (1946). External abdominal appearance i.e. unfed, blood-fed (full or partial), and sugar-fed was recorded by visual examination. Specimens were preserved on silica gel beads (Delta Enterprises, Inc., USA) in individual 1.5 ml Eppendorf tubes prior to transportation to the University of Notre Dame (USA) for further characterization.

#### **4.7.5 Climatic data**

Temperature, and humidity were measured on an hourly basis throughout the duration of the experiment using a hygrometer. These measurements were also recorded at the time of preparing the sugar-fermented yeast CO<sub>2</sub> solution on experimental days 4-12 inclusive. Rainfall, wind, and moonlight were not measured scientifically but were observed and documented throughout the experiment.

#### **4.7.6 Ethics**

Ethical approval was not required by the University of Notre Dame Institutional Review Board as no human subjects participated in the study. As it was nested under the MTC, however, their ethical approval from the National Health Research & Ethics Committee (2011-05-02, HRE02/16) provided the necessary permissions to satisfy the requirements of the Solomon Islands government.

## 4.8 Results

### 4.8.1 Trap placement

According to observations noted during the MTC entomological surveys conducted prior to this study (Tom Burkot and Bob Cooper, personal communication), greater numbers of *An. farauti* mosquitoes were consistently sampled in HLCs near residences located on the elevated area in between the two swamps than those in the main village. It was decided, therefore, to reposition the traps to this location (Figure 4.12, sites 1-3: red solid rectangles) in order to maximize catch numbers.

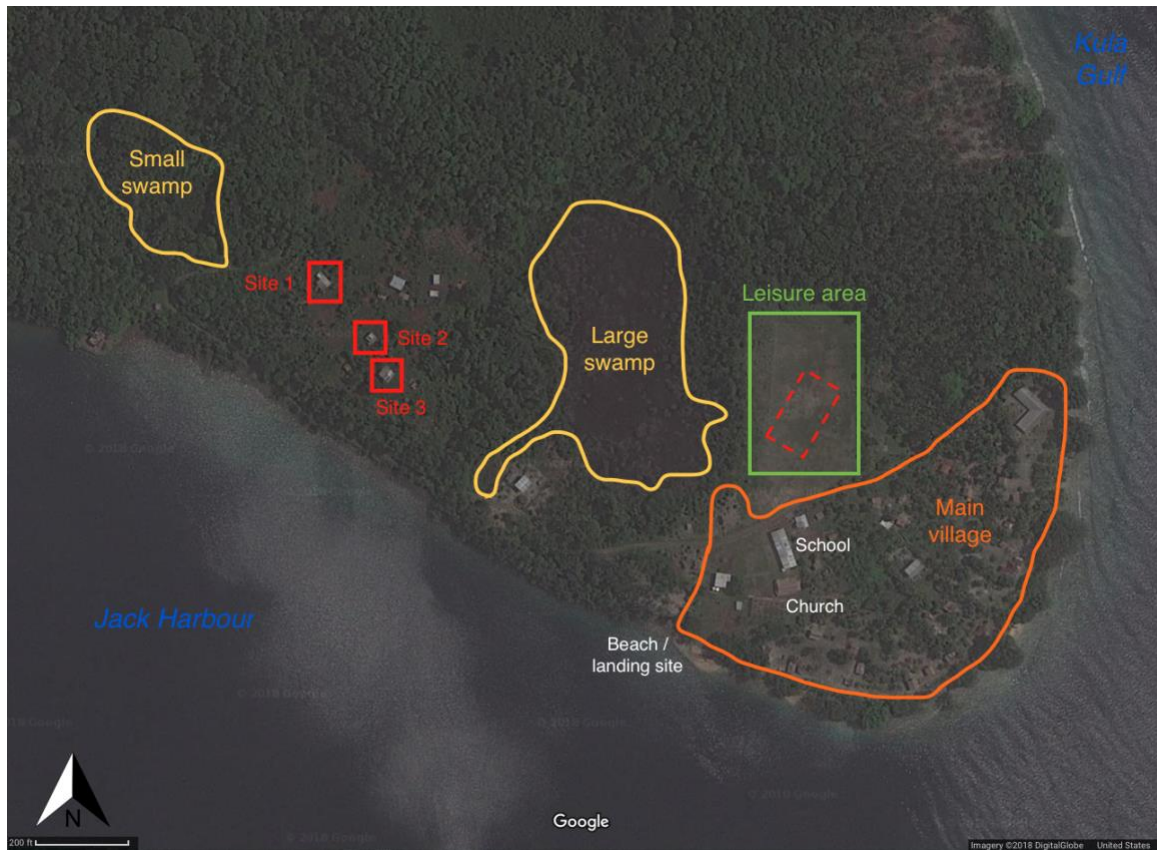


Figure 4.12 **Satellite image of Jack Harbour village illustrating re-placement of the experimental traps (red solid rectangles) to maximize *An. farauti* catch numbers.** Residences hosting traps are labeled Site 1, 2, or 3, and were located at elevation in between the two major *An. farauti* oviposition sites (yellow outlined swamps). Map data: Google, 2018.

Site 1 (Figure 4.13) consisted of a main residence that represents a typical rural structure, constructed of timber planks constructed on a stilt base with a corrugated iron roof. Stilts allow for a shaded area in which to avoid the tropical climate. Windows are styled as wooden or metal shutters, which are mostly propped open but closed during rainstorms, or covered with material. The structure to the right is used as a cooking area. The site illustrates that a typical rural

residence allows for unimpeded transmission of malaria by populations of outdoor-biting *An. farauti* mosquitoes, which take full advantage of their hosts' early evening social activities in these areas. Sites 2 and 3 were similar structures constructed of the same materials thus, providing almost identical experimental conditions at each trap. As observed from the image, traps were placed approximately 10 m from the main residence to ensure that host-seeking/resting behavior was being captured.



**Figure 4.13 Location of the first trap (Site 1) after re-placement to the elevated area between the two suspected *An. farauti* oviposition sites.** The main residence is situated in the background to the left, and cooking area to the right.

#### 4.8.2 Sugar-fermented yeast CO<sub>2</sub> preparation

Initially (days 1-3 inclusive), the sugar-fermented yeast CO<sub>2</sub> source was prepared later in the afternoon (TABLE 4.3) to optimize fermentation conditions in which to produce sufficient CO<sub>2</sub> plumes that coincided with *An. farauti* peak biting activity. However, the study that the sugar-yeast mixture, and its proportions were based on (Smallgange et al. 2010) had been previously conducted with *An. gambiae* mosquitoes in Kenya, which demonstrate a later peak biting profile (20:30 hours) than the *An. farauti* populations in Jack Harbour (19:00-20:00 hours) (T. Burkot et al. 2018). To correct for the biting profile differences between the two species, the fermentation process was started earlier in the day (14:00 hours) for the duration of the experiment (days 4-12 inclusive). Laboratory gloves were worn while preparing the mixture to eliminate the possibility of erroneously contaminating the experiment with volatile organic compounds (VOCs). Confirmation of CO<sub>2</sub> production was visual through the production of bubbles.

TABLE 4.3

DAILY CLIMATIC CONDITIONS RECORDED DURING SUGAR-FERMENTED YEAST  
CO<sub>2</sub> PREPARATION

<i>Date</i> (Sept 2014)	<i>Mixture start time *</i>	<i>Temperature (°C)</i>	<i>Humidity (%)</i>	<i>Experiment start time *</i>	<i>Fermentation time elapsed (hours)</i>
<sup>§</sup> Rep 1 16	- <sup>a</sup>	-	-	18:00	-
17	- <sup>a</sup>	-	-	18:09	-
18	- <sup>a</sup>	-	-	18:03	-
Rep 2 20	14:00	33.8	70.0	18:30	4.5
21	14:05	28.4	83.3	18:00	3.9
22	14:04	28.3	90.0	18:00	3.9
Rep 3 23	14:00	30.1	89.4	18:00	4.0
24	14:00	26.8	97.3	18:00	4.0
25	14:00	29.5	88.3	18:14	4.2
Rep 4 27	14:00	27.2	86.8	18:43	4.7
28	14:00	25.0	86.9	18:41	4.7
29	14:00	30.4	84.6	18:00	4.0

NOTE:

\*Time is recorded using the 24-hour clock

<sup>§</sup>Replication numbers (1-4) in the Latin Square experimental design. Each replication lasts 3 days<sup>a</sup>Fermentation times and climatic conditions were not recorded until the second experimental replicate

#### 4.8.3 Climatic data

There were a number of climatic conditions experienced during the 12-night experiment, which occurred over three periods: (1) Tuesday 16-Thursday 18 September 2014, (2) Saturday 20-Thursday 25 September 2014, and (3) Saturday 2-Monday 29 September 2014. These intervals were required to allow for two non-collecting days due to the weekly Sabbath (Friday 18:00 hours-Saturday 18:00

hours) of the Seventh Day Adventists' faith, which is the primary religious denomination practiced in Jack Harbour.

Temperature, and humidity were recorded with a hygrometer, and are presented in TABLE 4.4. Additionally, observations on rainfall, wind, and moonlight are also included although none were measured scientifically.

TABLE 4.4

OBSERVED CLIMATIC CONDITIONS RECORDED IN JACK HARBOUR DURING THE  
STUDY PERIOD, 16-29 SEPTEMBER 2014

<i>Date (Sep 2014)</i>	<i>Temperature (°C) *</i>	<i>Humidity (%) *</i>	<i>Rainfall (hrs.) <sup>a</sup></i>	<i>Wind <sup>b</sup></i>	<i>Moonlight <sup>c</sup></i>
<sup>d</sup> Rep 1 16	26.1	98.1	0.5	-	Present <sup>e</sup>
17	27.1	95.6	0.0	-	Present <sup>f</sup>
18	25.2	99.4	6.4	-	Present <sup>f</sup>
Rep 2 20	28.0	90.4	0.0	Breeze S -> N	Present <sup>f</sup>
21	26.1	96.9	2.3	Breeze N -> S	Present <sup>f</sup>
22	26.7	95.5	6.0	Strong N -> S	Present <sup>f</sup>
Rep 3 23	25.3	96.6	2.6	Strong N -> S	Absent <sup>f</sup>
24	24.0	100.0	17.5	-	Absent <sup>g</sup>
25	21.8	99.9	6.0	-	Absent <sup>h</sup>
Rep 4 27	21.9	98.3	0.0	-	Present <sup>h</sup>
28	19.2	100.0	4.8	Strong All directions	Present <sup>h</sup>
29	26.8	97.1	0.6	Strong E -> W	Present <sup>h</sup>

## NOTE:

\* Temperature and humidity were recorded hourly from 18:00-00:00 hours but mean figures over the collecting period are listed here

<sup>a</sup> Rainfall was recorded in minutes over the course of the day, which was usually an 18-hour period, but is presented in hours for ease of reference

<sup>b</sup> Observations on wind patterns during the experiment are presented as either: (-) wind not observed during this collection period, or a description of the observation containing the strength, and cardinal directions (N-North; S-South; E-East; W-West) indicated by (->)

<sup>c</sup> Moonlight is recorded as either present or absent. Different phases of the moon allow for differing illumination based on its orbital position around the earth. This, and how it could affect mosquito flight activity, is further considered in the discussion section of this chapter

<sup>d</sup> This is the replication number (1-4) in the Latin Square study design. Each replication lasts 3 days

<sup>e</sup> The moon entered its third quarter phase on September 16 (50.0% illumination)

<sup>f</sup> Between September 17-23, the moon was in its waning crescent phase (0.1-49.9% illumination)

<sup>g</sup> On September 24, a new moon appeared although it was not visible from earth (0.0% illumination)

<sup>h</sup> Between September 25-29, the waxing crescent moon (0.1-49.9% illumination) was building up to its first quarter phase (50.0% illumination)



#### 4.8.4 Morphological species identification

*An. farauti* adult female mosquitoes were morphologically identified to sex and species in accordance with the proboscis banding, and color patterns described by Rozeboom and Knight (1946). All specimens demonstrated a Type A proboscis (Figure 3.1). Molecular identification was not considered necessary on the basis of recent evidence identifying *An. farauti* as the exclusive species in Jack Harbour (see Section 4.5.4 Malaria vectors).

#### 4.8.5 Attractiveness of CO<sub>2</sub> sources

653 anopheline mosquitoes were captured by the three baits (control, sugar-fermented yeast, human) across the 12-night experimental period. The human-generated CO<sub>2</sub> bait caught a mean count of 29.1 mosquitoes (median=26.5) on any given night, followed by sugar-fermented yeast (17.5; median=13.0) and control (7.8; median=7.5).

Figure 4.14 illustrates the distribution of the data. Initial observations, which could be stated about any given evening were:

1. The human-generated CO<sub>2</sub> bait attracted more mosquitoes than the control as there was a between-group difference
2. There was no group difference between the sugar-fermented yeast CO<sub>2</sub>, and human-generated CO<sub>2</sub> baits as there was intersection
3. There was no group difference between the control and sugar-fermented yeast CO<sub>2</sub> baits as there was intersection

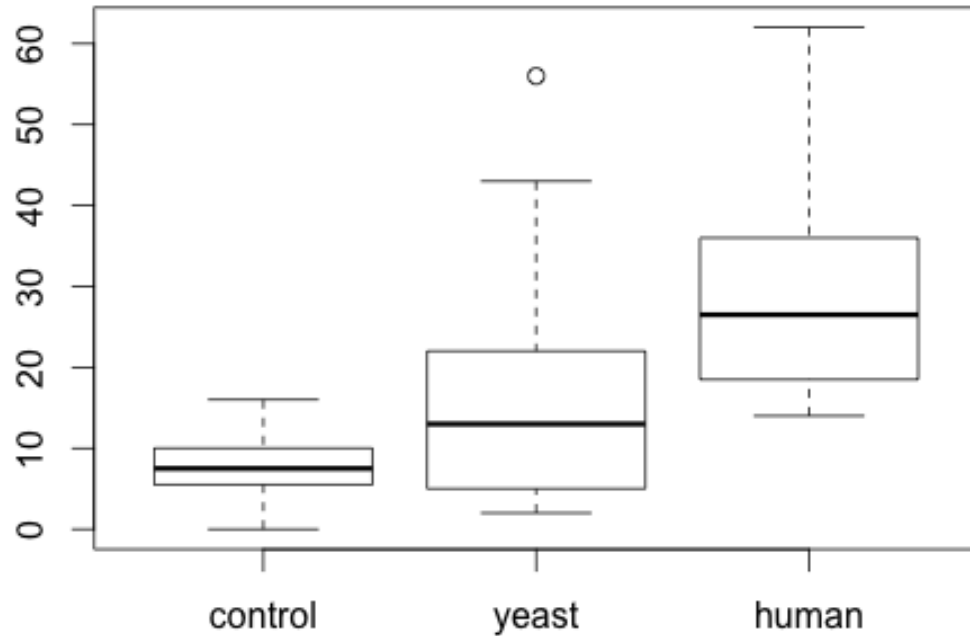


Figure 4.14 **Boxplot summarizing attractiveness of CO<sub>2</sub> source to outdoor-biting populations of *An. farauti* over the 12-night experimental period.** The x-axis identifies the different CO<sub>2</sub> sources (control, yeast, human). Catch numbers are measured per night by the y-axis. Median values (solid black horizontal lines), quartiles (white boxes), variance (dashed black vertical lines), outliers (open circles), and error bars are indicated.

87 of the 656 mosquitoes captured were graded as blood-fed (13.3%). These were collected per CO<sub>2</sub> attractant in the following proportions: control (19.5%), sugar-fermented yeast (35.6%), and human-generated (44.8%). When analyzed as a percentage of the total catch per hour, blood-fed females comprised between 10.0% (20:00-21:00 hours) and 14.9% (21:00-22:00 hours) across all three trap baits.

The analysis of variance (ANOVA) model identified a statistically significant effect between the trap baits ( $F=7.92$ ,  $df=2$ ,  $p<0.002$ ) (TABLE 4.5), and Tukey's HSD *post-hoc* test detected that this was due to differences between the mean number of mosquitoes captured by the control and human baits ( $diff=21.25$ ,  $p<0.001$ ) (TABLE 4.6).

TABLE 4.5

SUMMARY ANALYSIS OF VARIANCE BETWEEN THE MEAN NUMBER OF  
MOSQUITOES CAPTURED IN THE THREE EXPERIMENTAL CONDITIONS  
(TRAPBAIT)

<i>Summary(anova)</i>	<i>df</i>	<i>Sum Sq</i>	<i>Mean Sq</i>	<i>F value</i>	<i>Pr(&gt;F)</i>	<i>Significance</i>
<i>TrapBait</i>	2	2717	1358.4	7.92	0.00155	**
<i>Residuals</i>	33	5660	171.5	-	-	-

NOTE: \*\* indicates statistical significance

TABLE 4.6

COMPARISON OF GROUP MEANS (\$TRAPBAIT) BY TUKEY'S HONESTLY  
SIGNIFICANT DIFFERENCE *POST-HOC* TEST WITH 95% FAMILY-WISE CONFIDENCE  
LEVEL

<i>\$TrapBait</i>	<i>diff</i>	<i>lwr</i>	<i>upr</i>	<i>p adj</i>
<i>human-control</i>	21.25	8.13	34.37	0.001
<i>yeast-control</i>	9.67	-3.45	22.79	0.183
<i>yeast-human</i>	-11.58	-24.70	1.54	0.092

Statistical assumptions of normality (Shapiro-Wilk:  $p > 0.005$ ), and variance (Levene:  $F(2,33) = 2.89$ , *ns*) about the distribution of the data were not violated.

#### 4.8.6 Host-seeking/resting patterns

The number of mosquitoes captured per hour was analyzed to identify whether the *An. farauti* populations in the location between the two swamps exhibited any time-related host-seeking/resting patterns. Figure 4.15 illustrates that although catches began as early as 18:00 hours ( $n=46$ ), the majority of mosquitoes were captured between 21:00-23:00 hours ( $n=308$ ).

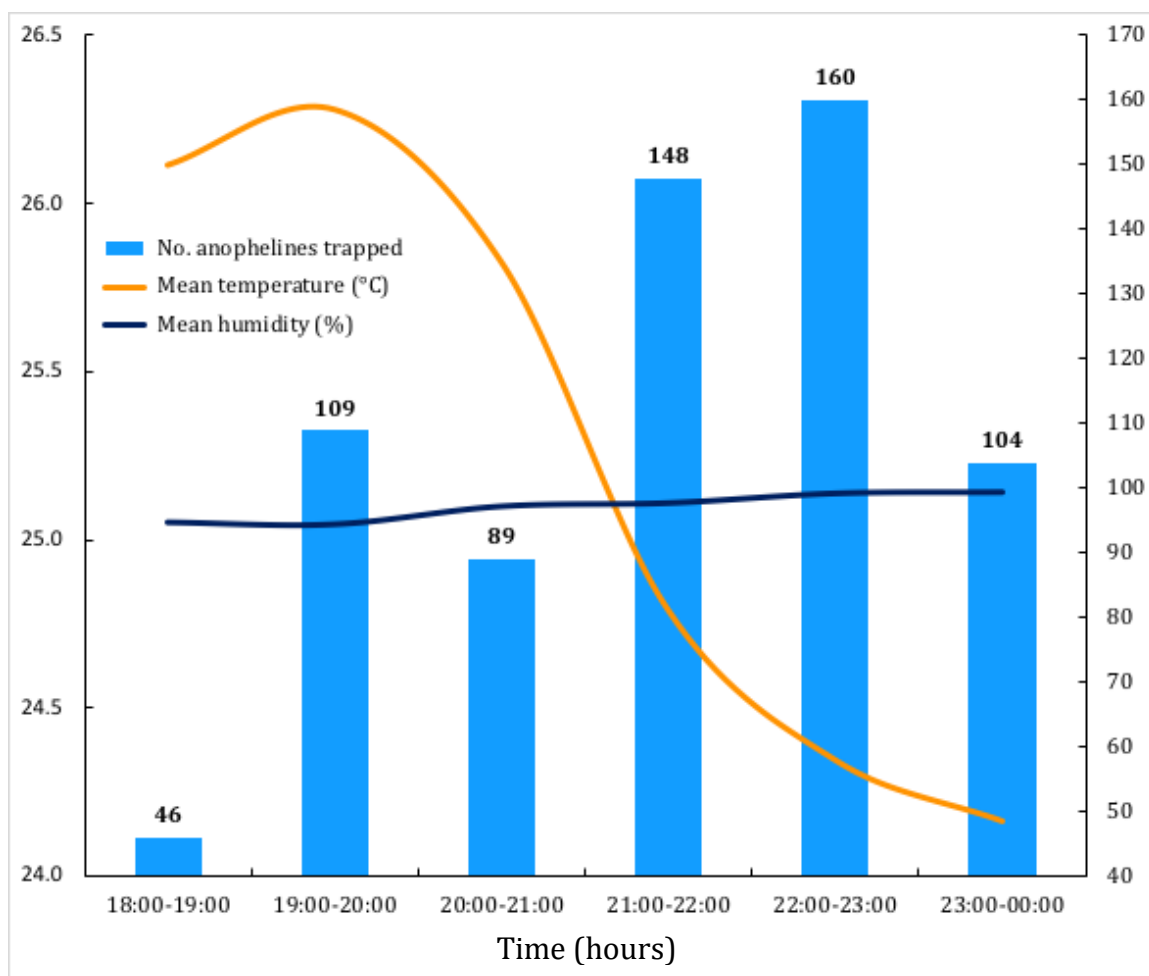


Figure 4.15 **Bar chart illustrating total number of anopheline mosquitoes captured per hour over the 12-night experimental period.** Mean temperature (°C) (orange line), and humidity (%) (dark blue line) measurements for the period are also presented. Black numbers above the blue bars are the total number of mosquitoes captured in the one hour time period indicated on the x-axis (hours). y-axes represent: left (temperature in °C), and right (humidity in %, and mosquito count).

#### 4.9 Discussion

The purpose of this study was to evaluate the effectiveness of sugar-fermented yeast as a source of organic CO<sub>2</sub> to attract outdoor-biting and/or resting populations of *An. farauti* in Jack Harbour, Western Province. Mosquitoes were

captured during a 12-night experimental procedure whereby three sources (control, sugar-fermented yeast CO<sub>2</sub>, human-generated CO<sub>2</sub>) were used as trap baits. A total of 653 *An. farauti* mosquitoes were caught with the human bait attracting the greatest numbers (n=349) followed by sugar-fermented yeast (n=210), and control (n=94). Statistical analysis found no evidence to support sugar-fermented yeast as an organic source of CO<sub>2</sub> attracting greater numbers of *An. farauti* than human-generated CO<sub>2</sub> overall; however, there was overlap, which inferred that on any given evening there was no between-group difference. In conclusion, this study illustrated that sugar-fermented yeast as an organic source of CO<sub>2</sub> was attractive to populations of *An. farauti* in Jack Harbour, Western Province, and improvements to the source design i.e. plume composition and delivery, could further enhance its attraction.

#### **4.9.1 Human-generated CO<sub>2</sub> attracted greater numbers of *An. farauti***

The human-generated CO<sub>2</sub> bait attracted significantly more mosquitoes than the control and the sugar-fermented yeast CO<sub>2</sub> but the latter was not statistically significant. The biological difference between these two sources is the human, who also produces VOCs that the mosquito detects to determine whether or not it will take a blood meal from that source. It does not sense VOCs, however, until after it has used CO<sub>2</sub> to direct it towards the source of the VOCs. Therefore, the sugar-fermented yeast CO<sub>2</sub> source could be improved by enhancing the plume dynamics to make it more attractive to *An. farauti*. For example, the mixing vessel in this study comprised a 5 L carboy with a wide open neck that dispensed the CO<sub>2</sub>, which was probably not conducive to creating the turbulent, filamentous plume that the

mosquitoes required as a sensing tool. Additionally, the carboy was placed on the ground in the center of the trap. As CO<sub>2</sub> is denser than air, it should have been placed higher (1 m above the ground) to allow the CO<sub>2</sub> to tumble to the ground, also creating a more turbulent plume (Zainulabeuddin Syed, personal communication).

Ideally, the physiological and behavioral thresholds of detection of CO<sub>2</sub> sensitivity for *An. farauti* should first be determined in order to generate a source that closely mimics the naturally produced CO<sub>2</sub> exhaled by a human. Following on from that it is necessary to study the composition, and structure of CO<sub>2</sub> plumes with particular reference to *An. farauti* using wind tunnels, stimulus chambers and olfactometers in order to establish the most desirable configuration for field evaluation.

#### **4.9.2 Peak host-seeking/resting behavior occurred between 21:00-23:00 hours**

The majority of mosquitoes were captured between 21:00-23:00 hours across the 12-night experiment suggesting peak host-seeking/resting behavior occurred during this period. The MTC entomological survey conducted immediately prior to this study (9-16 September 2014) identified peak biting activity of *An. farauti* to occur between 19:00-20:00 hours as measured by mean biting density (bites per person per hour) (T. Burkot et al. 2018). Captures occurring after peak biting could be explained as resting behavior—the net material that the traps were constructed from would certainly be sufficient to act as a resting site for mosquitoes returning to oviposit (T. Burkot et al. 2013)—especially since *An. farauti* is not

known to exhibit endophily after a blood meal but this was not tested for specifically. The number of blood-fed females captured during this period (n=44; 50.6%) does not support either behavior. In fact, if blood-fed females were resting why were greater numbers captured in the traps baited with CO<sub>2</sub> sources than the control trap when, presumably, their physiological need to obtain protein had been met?

#### 4.10 Conclusion

This study illustrated that sugar-fermented yeast as an organic source of CO<sub>2</sub> was attractive to *An. farauti* mosquitoes in Jack Harbour, Western Province and improvements to the source design i.e. CO<sub>2</sub> plume composition and delivery, could further enhance its appeal.



## CHAPTER 5:

### CONCLUSIONS

Chapters two, three, and four presented research that sought to explore how the genetic and bionomic behaviors of two important anopheline mosquitoes might respond to alternative vector control approaches, which the malERA Consultative Group on Vector Control (2011) had identified as critical developmental challenges in the effort to interrupt sustained malaria transmission. A brief summary of each chapter, together with comments that consider how the application of this research can be functional in the wider context of malaria vector control, concludes this dissertation.

#### 5.1 Research summary

##### **5.1.1 Micro-population structure of *An. gambiae* in the Ugandan Lake Victoria basin**

Natural populations of *An. gambiae* in the northwestern Lake Victoria basin were entomologically sampled and sequenced by a reduced-representation technique to enhance previously completed research as the genomic approach permitted resolution of genetic structure (eigenanalysis) at the nucleotide level. A

dataset of 5,175 SNPs was generated. Traditional population genetic measurements such as differentiation ( $F_{ST}$ ) and effective size ( $N_e$ ) also supported eigenanalysis stratification (PCA) based on geographic location. In conclusion, three island populations (Bukasa, Nsadzi, Sserinya) demonstrated greater structure, low to moderate genetic differentiation, and small effective sizes that combined, were indicative of limited migration with susceptibility to genetic drift. These sites would, therefore, be appropriate candidates for the evaluation of small-scale ecologically-confined entomological efficacy of GM *An. gambiae* releases.

#### **5.1.2 Macro-population structure of *An. farauti* in the southwest Pacific region**

Natural populations of *An. farauti* were entomologically sampled from four geographically distant locations in the southwest Pacific (Australia, Papua New Guinea, Solomon Islands, Vanuatu) and whole genome sequenced via a shotgun approach. A dataset of 17,757 SNPs was generated. Population stratification analysis confirmed the four geographic populations genetically (PCA) and also identified similarity between the individuals in Australia and Vanuatu (PCA, ADMIXTURE). Variation was observed within populations from Papua New Guinea and the Solomon Islands, while Australia and Vanuatu were less diverse (PCA). Individuals in the Solomon Islands demonstrated greater structure with the ancestry assignment model (ADMIXTURE) separating them into two populations. This research enhanced preceding *An. farauti* population genetic studies with a SNP-based approach, which permitted high resolution genome-wide analysis that detected a previously unknown genetic relationship.

### **5.1.3 Sugar-fermented yeast as a source of CO<sub>2</sub> to attract *An. farauti* mosquitoes in Western Province, Solomon Islands**

653 *An. farauti* mosquitoes were collected over a 12-night period in an evaluation of effectiveness of sugar-fermented yeast as an organic source of CO<sub>2</sub> to attract outdoor-biting and/or resting populations in Jack Harbour, Western Province. Human-generated CO<sub>2</sub> attracted the greatest numbers (n=349) followed by sugar-fermented yeast (n=210), and control (n=94). There was no statistical evidence to support sugar-fermented yeast as an organic source of CO<sub>2</sub> being more attractive to *An. farauti* mosquitoes than human-generated CO<sub>2</sub> overall (ANOVA:  $F=7.92$ ,  $df=2$ ,  $p<0.002$ ); however, intersection in the distribution of the data inferred there was no between-group difference on any given evening. This finding could be exploited to improve the source design i.e. CO<sub>2</sub> plume composition and delivery, which would further enhance its attraction.

### **5.2 Contribution that extends the research agenda**

SNP-based assessment of genetic structure in *An. gambiae* (Chapter two) and *An. farauti* (Chapter three) populations advances previously conducted research as these markers permit high resolution analysis across the genome. Although single locus markers i.e. mtDNA can adequately identify structure, multiple loci have greater power in detecting cryptic relationships that would otherwise remain latent, and genetic mechanisms, which underpin the key biological traits contributing to vectorial capacity.

Evaluation of sugar-fermented yeast as an organic source of CO<sub>2</sub> trap attractant to *An. farauti* mosquitoes has generated further insight into the bionomic behavior of these populations in Jack Harbour, Western Province, Solomon Islands.

### 5.3 Application to malaria vector control approaches

The *An. gambiae* research presented in Chapter two is part of a suite of publications (Kayondo et al. 2005; Lukindu et al. 2018; Wiltshire et al. 2018; Bergey et al. 2019) that, combined, lend support to sites in the Ssesse Islands, Lake Victoria being candidates for small-scale, ecologically confined field testing of GM mosquitoes. This is the next step in the pathway, which moves *Anopheles* GM mosquito research forward from the recent technical advances in the lab (Gantz et al. 2015; Hammond et al. 2016; Kyrou et al. 2018) to the field (World Health Organization 2014; James et al. 2018) in an effort to make GM vector control real.

Chapter four generated preliminary data, which suggested that sugar-fermented yeast as a source of organic CO<sub>2</sub> could be an effective attractant to *An. farauti* mosquitoes. This could be added to traps for use in either: (i) surveillance, or (ii) as an intervention if it incorporates insecticide into its design.

APPENDIX A:

GLOBAL POSITIONING SYSTEM (GPS) COORDINATES FOR AND DISTANCES  
BETWEEN THE SIX *ANOPHELES GAMBIAE* ENTOMOLOGICAL SAMPLING SITES IN  
THE NORTHWESTERN LAKE VICTORIA BASIN

TABLE A.1

GPS COORDINATES FOR AND DISTANCES BETWEEN THE SIX *ANOPHELES GAMBIAE* ENTOMOLOGICAL SAMPLING SITES IN  
THE NORTHWESTERN LAKE VICTORIA BASIN

<i>From</i> <sup>a</sup>	<i>Longitude</i>		<i>Latitude</i>		<i>To</i> <sup>a</sup>	<i>Longitude</i>		<i>Latitude</i>		<i>Distance</i> (km) <sup>‡</sup>
	<i>Decimal degrees<sup>§</sup></i>	<i>Degrees, minutes, seconds</i>	<i>Decimal degrees<sup>§</sup></i>	<i>Degrees, minutes, seconds</i>		<i>Decimal degrees<sup>§</sup></i>	<i>Degrees, minutes, seconds</i>	<i>Decimal degrees<sup>§</sup></i>	<i>Degrees, minutes, seconds</i>	
<i>EB</i> <sup>*</sup>	32.4604	32°27'37"E	0.0676	0°04'03"N	<i>WL</i> <sup>*</sup>	31.9971	31°59'50"E	0.4096	0°24'35"N	64.0
					<i>BL</i>	32.2939	32°17'38"E	-0.3090	0°18'32"S	45.6
					<i>BK</i>	32.4483	32°26'54"E	-0.4878	0°29'16"S	61.4
					<i>SY</i>	32.3641	32°21'51"E	-0.2523	0°15'08"S	31.9
					<i>NZ</i>	32.5856	32°35'08"E	-0.0860	0°05'10"S	22.0
<i>WL</i> <sup>*</sup>	31.9971	31°59'50"E	0.4096	0°24'34"N	<i>BL</i>	32.2939	32°17'38"E	-0.3090	0°18'32"S	86.1
					<i>BK</i>	32.4483	32°26'54"E	-0.4879	0°29'16"S	111.2
					<i>SY</i>	32.3641	32°21'51"E	-0.2523	0°15'08"S	74.5
					<i>NZ</i>	32.5857	32°35'08"E	-0.0860	0°05'10"S	85.4
<i>BL</i>	32.2939	32°17'38"E	-0.3090	0°18'32"S	<i>BK</i>	32.4483	32°26'54"E	-0.4879	0°29'16"S	26.2
					<i>SY</i>	32.3641	32°22'18"E	-0.2523	0°15'08"S	9.4
					<i>NZ</i>	32.5857	32°35'08"E	-0.0860	0°05'10"S	38.6
<i>BK</i>	32.4483	32°26'54"E	-0.4879	0°29'16"S	<i>SY</i>	32.3641	32°21'51"E	-0.2523	0°15'08"S	24.1
					<i>NZ</i>	32.5857	32°35'08"E	-0.0860	0°05'10"S	47.0
<i>SY</i>	32.3641	32°21'51"E	-0.2523	0°15'08"S	<i>NZ</i>	32.5857	32°35'08"E	-0.0860	0°05'10"S	29.1
<i>SYK</i>	32.3717	32°22'18"E	-0.2638	0°15'50"S						
<i>SYB</i>	32.3564	32°21'23"E	-0.2407	0°14'27"S						

TABLE A.1 (CONTINUED)

NOTE:

<sup>a</sup> Towns and villages are abbreviated as follows: EB (Entebbe), WL (Lake Wamala), BL (Bugala), BK (Bukasa), SY (Sserinya), SYK (Sserinya-Kafuna), SYB (Sserinya-Bbosa), and NZ (Nsadzi). Also refer to Figure 2.1 for further details

\* Denotes mainland site

<sup>§</sup> GPS coordinates were recorded in decimal degrees and converted into degrees, minutes, and seconds at <https://www.ngs.noaa.gov/NCAT/> (National Geodetic Survey 2018)

<sup>#</sup> Distances calculated at: <http://boulter.com/gps/distance/>

## APPENDIX B:

### BIOINFORMATICS PIPELINE DETAILING SOFTWARE PROGRAMS AND PARAMETERS IMPLEMENTED IN THE ANALYSIS OF THE *ANOPHELES GAMBIAE* RADSEQ DATASET



TABLE B.1

SOFTWARE AND PARAMETERS APPLIED IN THE BIOINFORMATICS ANALYSIS OF 96 WILD-TYPE *ANOPHELES GAMBIAE* INDIVIDUALS SEQUENCED USING AN *ECORI* AND *MSERI* RESTRICTION SITE-ASSOCIATED DIGEST

<i>Software</i>	<i>Version</i>	<i>Function</i>	<i>Parameters</i>	<i>Reference</i>
<i>FastQC</i>	<i>0.10.1</i>	Sequencing quality check		(Babraham Institute 2011)
<i>Trimmomatic</i>	<i>0.30</i>	Remove Illumina sequencing adapters	ILLUMINACLIP:TRUSEQ3-PE.FA:2:30:10 LEADING:5 TRAILING:5 SLIDINGWINDOW:4:15 MINLEN:50	(Bolger, Lohse, and Usadel 2014)
<i>Trimmer.py</i>	<i>Custom</i>	RAD barcode, cut site, protector base removal; addition of unique IDs matching barcodes to individuals		(Notre Dame Bioinformatics Lab 2014)
<i>Burrows-Wheeler Alignment (BWA)</i>	<i>0.6.2</i>	Alignment of reads to <i>AgamP4</i> reference genome	-t 12 -q 5 -l 32 -k 3 -n 9 -o 1	(Li and Durbin 2009)
<i>sampToSam.pl</i>	<i>Custom</i>	Addition of read groups (unique to sample IDs) to		(Ragland 2017)

TABLE B.1 (CONTINUED)

<i>Software</i>	<i>Version</i>	<i>Function</i>	<i>Parameters</i>	<i>Reference</i>
		SAM files; change overall quality score for GATK compatibility		
<i>Picard</i>	<i>1.119</i>	Pre-processing of reads in format required by GATK	CleanSam -VALIDATION_STRINGENCY =LENIENT SortSam -SO=COORDINATE -VALIDATION_STRINGENCY =LENIENT BuildBamIndex -VALIDATION_STRINGENCY =LENIENT CreateSequenceDictionary	(Broad Institute 2017)
<i>GenomeAnalysis ToolKit (GATK)</i>	<i>3.3.0</i>	Variant calling	-T UnifiedGenotyper --downsampling type NONE --downsample_to_coverage 1000 --genotype_likelihoods_model BOTH --computeSLOD -rf BadCigar --fix_misencoded_quality_scores	(Van der Auwera et al. 2013)
<i>GATK</i>	<i>3.3.0</i>	Quality filtering of variant calls; selects SNPs passing filters, which are biallelic only	-T VariantFiltration --filterExpression "QD<5.0  FS>60.0  MQ<40.0  HaplotypeScore>13.0  MappingQualityRankSum<8.0" --missingValuesInExpressionsShouldEvaluateAsFailing -T SelectVariants --selectexpressions "vc.isNotFiltered() &&	(Van der Auwera et al. 2013)

TABLE B.1 (CONTINUED)

<i>Software</i>	<i>Version</i>	<i>Function</i>	<i>Parameters</i>	<i>Reference</i>
			vc.isSNP()” --selectTypeToExclude INDEL --restrictAllelesTo BIALLELIC	
<i>VCFtools</i>	<i>0.1.15</i>	Dataset pruning	--mac 7 (at least 4 diploid individuals called at locus) --minGQ 30.0 (Genotype (GT) Quality of at least 30) --minDP 30.0 (at least 30 reads support GT) --singletons --max-missing 0.2 (exclude SNPs where missing data across populations > 80%)	(Danecek et al. 2011)
<i>PLINK</i>	<i>1.9</i>	Format data for PCA and global ancestry analysis	--allow-extra-chr --pca --recode 12	(Chang et al. 2015)
<i>R</i>	<i>3.2.1</i>	Apply plot functions to visualize data		(R Core Team 2014)
<i>ADMIXTURE</i>	<i>1.2.3</i>	Global ancestry analysis		(Alexander, Novembre, and Lange 2009)

APPENDIX C:

DETERMINATION OF THE 2LA INVERSION KARYOTYPE IN NATURAL  
POPULATIONS OF *ANOPHELES GAMBIAE* MOSQUITOES SAMPLED FROM THE  
NORTHWESTERN LAKE VICTORIA BASIN

TABLE C.1

2LA INVERSION KARYOTYPING RESULTS FOR *ANOPHELES GAMBIAE* MOSQUITOES  
 SAMPLED FROM SIX NATURAL POPULATIONS IN THE NORTHWESTERN LAKE  
 VICTORIA BASIN

<i>PCA group</i> <sup>*</sup>	<i>Sample ID</i> <sup>‡</sup>	<i>Allele length (bp)</i> <sup>§</sup>	<i>PCR result</i> <sup>a</sup>	<i>Sample karyotype</i> <sup>b</sup>	<i>Sample phenotype</i> <sup>c</sup>
1	BK.IR.003	492	Correct-call	2La/a	Inverted
1	BL.IR.003	492	Correct-call	2La/a	Inverted
1	BL.IR.009	492	Correct-call	2La/a	Inverted
1	EB.IR.001	492	Correct-call	2La/a	Inverted
1	EB.IR.003	492	Correct-call	2La/a	Inverted
1	EB.IR.004	492	Correct-call	2La/a	Inverted
1	EB.IR.007	492	Correct-call	2La/a	Inverted
1	EB.IR.008	492	Correct-call	2La/a	Inverted
1	EB.IR.009	492	Correct-call	2La/a	Inverted
1	EB.LC.006	492	Correct-call	2La/a	Inverted
1	EB.LC.008	492	Correct-call	2La/a	Inverted
1	NZ.IR.002	492	Correct-call	2La/a	Inverted
1	NZ.IR.004	492	Correct-call	2La/a	Inverted
1	NZ.IR.007	492	Correct-call	2La/a	Inverted
1	NZ.IR.008	492	Correct-call	2La/a	Inverted
1	NZ.IR.011	492	Correct-call	2La/a	Inverted
1	NZ.LC.M+1.007	492	Correct-call	2La/a	Inverted
1	NZ.LC.M+1.009	492	Correct-call	2La/a	Inverted
1	NZ.LC.M+1.011	492	Correct-call	2La/a	Inverted
1	SY.IR.2.003	492	Correct-call	2La/a	Inverted
1	SY.IR.2.004	492	Correct-call	2La/a	Inverted
1	SY.LC.1.001	492	Correct-call	2La/a	Inverted
1	SY.LC.1.003	492	Correct-call	2La/a	Inverted
1	SY.LC.1.015	492	Correct-call	2La/a	Inverted
1	SY.LC.2.001	492	Correct-call	2La/a	Inverted
1	SY.LC.2.002	492	Correct-call	2La/a	Inverted
1	SY.LC.2.003	492	Correct-call	2La/a	Inverted
1	SY.LC.2.004	492	Correct-call	2La/a	Inverted
1	SY.LC.2.005	492	Correct-call	2La/a	Inverted
1	SY.LC.2.006	492	Correct-call	2La/a	Inverted

TABLE C.1 (CONTINUED)

<i>PCA group</i> <sup>*</sup>	<i>Sample ID</i> <sup>‡</sup>	<i>Allele length (bp)</i> <sup>§</sup>	<i>PCR result</i> <sup>a</sup>	<i>Sample karyotype</i> <sup>b</sup>	<i>Sample phenotype</i> <sup>c</sup>
1	SY.LC.2.007	492	Correct-call	2La/a	Inverted
1	WL.IR.002	492	Correct-call	2La/a	Inverted
1	WL.IR.005	492	Correct-call	2La/a	Inverted
1	WL.IR.015	492	Correct-call	2La/a	Inverted
1	WL.IR.017	492	Correct-call	2La/a	Inverted
2	BK.IR.001	207/492	Correct-call	2La <sup>+</sup> /a	Hetero-
2	BK.IR.007	207/492	Correct-call	2La <sup>+</sup> /a	Hetero-
2	BK.IR.010	207/492	Correct-call	2La <sup>+</sup> /a	Hetero-
2	BK.IR.012	207/492	Correct-call	2La <sup>+</sup> /a	Hetero-
2	BK.IR.015	207/492	Correct-call	2La <sup>+</sup> /a	Hetero-
2	BL.IR.001	207/492	Correct-call	2La <sup>+</sup> /a	Hetero-
2	BL.IR.003	207/492	Correct-call	2La <sup>+</sup> /a	Hetero-
2	BL.IR.012	492/650 <sup>d</sup>	False-call	2La/a	Inverted
2	BL.IR.014	207/492	Correct-call	2La <sup>+</sup> /a	Hetero-
2	BL.IR.015	207/492	Correct-call	2La <sup>+</sup> /a	Hetero-
2	BL.IR.018	207/492	Correct-call	2La <sup>+</sup> /a	Hetero-
2	BL.IR.023	207/492	Correct-call	2La <sup>+</sup> /a	Hetero-
2	BL.IR.030	207/492	Correct-call	2La <sup>+</sup> /a	Hetero-
2	BL.IR.032	207/492	Correct-call	2La <sup>+</sup> /a	Hetero-
2	BL.IR.033	207/492	Correct-call	2La <sup>+</sup> /a	Hetero-
2	EB.IR.002	207/492	Correct-call	2La <sup>+</sup> /a	Hetero-
2	EB.IR.005	207/492	Correct-call	2La <sup>+</sup> /a	Hetero-
2	EB.IR.006	207/492	Correct-call	2La <sup>+</sup> /a	Hetero-
2	EB.LC.001	207/492	Correct-call	2La <sup>+</sup> /a	Hetero-
2	EB.LC.002	492 <sup>d</sup>	False-call	2La/a	Inverted
2	EB.LC.003	207/492	Correct-call	2La <sup>+</sup> /a	Hetero-
2	EB.LC.010	207/492	Correct-call	2La <sup>+</sup> /a	Hetero-
2	NZ.IR.005	207/492	Correct-call	2La <sup>+</sup> /a	Hetero-
2	NZ.IR.006	492 <sup>d</sup>	False-call	2La/a	Inverted
2	NZ.LC.001	207/492	Correct-call	2La <sup>+</sup> /a	Hetero-
2	NZ.LC.002	207/492	Correct-call	2La <sup>+</sup> /a	Hetero-
2	NZ.LC.006	207/492	Correct-call	2La <sup>+</sup> /a	Hetero-
2	NZ.LC.M+1.010	492 <sup>d</sup>	False-call	2La/a	Inverted
2	SY.IR.2.001	492 <sup>d</sup>	False-call	2La/a	Inverted
2	SY.IR.2.002	492 <sup>d</sup>	False-call	2La/a	Inverted
2	SY.IR.2.005	207/492	Correct-call	2La <sup>+</sup> /a	Hetero-
2	SY.LC.1.011	492 <sup>d</sup>	False-call	2La/a	Inverted
2	WL.IR.001	492/650 <sup>d</sup>	False-call	2La/a	Inverted
2	WL.IR.003	207/492	Correct-call	2La <sup>+</sup> /a	Hetero-
2	WL.IR.011	207/492	Correct-call	2La <sup>+</sup> /a	Hetero-

TABLE C.1 (CONTINUED)

<i>PCA group</i> <sup>*</sup>	<i>Sample ID</i> <sup>‡</sup>	<i>Allele length (bp)</i> <sup>§</sup>	<i>PCR result</i> <sup>a</sup>	<i>Sample karyotype</i> <sup>b</sup>	<i>Sample phenotype</i> <sup>c</sup>
2	WL.IR.012	492/650 <sup>d</sup>	False-call	2La/a	Inverted
2	WL.IR.018	207/492	Correct-call	2La <sup>+</sup> /a	Hetero-
2	WL.IR.020	207/492	Correct-call	2La <sup>+</sup> /a	Hetero-
2	WL.IR.022	207 <sup>d</sup>	False-call	2La <sup>+</sup> /a <sup>+</sup>	Standard
2	WL.IR.024	207/492	Correct-call	2La <sup>+</sup> /a	Hetero-
2	WL.IR.025	207/492	Correct-call	2La <sup>+</sup> /a	Hetero-
3	BK.IR.002	207	Correct-call	2La <sup>+</sup> /a <sup>+</sup>	Standard
3	BK.IR.005	207	Correct-call	2La <sup>+</sup> /a <sup>+</sup>	Standard
3	BK.IR.006	207	Correct-call	2La <sup>+</sup> /a <sup>+</sup>	Standard
3	BK.IR.008	207	Correct-call	2La <sup>+</sup> /a <sup>+</sup>	Standard
3	BK.IR.013	207	Correct-call	2La <sup>+</sup> /a <sup>+</sup>	Standard
3	BK.IR.018	207	Correct-call	2La <sup>+</sup> /a <sup>+</sup>	Standard
3	BK.IR.020	207	Correct-call	2La <sup>+</sup> /a <sup>+</sup>	Standard
3	BL.IR.011	650 <sup>e</sup>	False-call	Unknown	Unknown
3	BL.IR.013	207	Correct-call	2La <sup>+</sup> /a <sup>+</sup>	Standard
3	BL.IR.016	207	Correct-call	2La <sup>+</sup> /a <sup>+</sup>	Standard
3	BL.IR.031	207	Correct-call	2La <sup>+</sup> /a <sup>+</sup>	Standard
3	BL.IR.034	207	Correct-call	2La <sup>+</sup> /a <sup>+</sup>	Standard
3	EB.LC.004	207	Correct-call	2La <sup>+</sup> /a <sup>+</sup>	Standard
3	NZ.LC.005	207	Correct-call	2La <sup>+</sup> /a <sup>+</sup>	Standard
3	WL.IR.014	207	Correct-call	2La <sup>+</sup> /a <sup>+</sup>	Standard
3	WL.IR.016	207	Correct-call	2La <sup>+</sup> /a <sup>+</sup>	Standard
3	WL.IR.021	207	Correct-call	2La <sup>+</sup> /a <sup>+</sup>	Standard
2	BK.IR.009 <sup>f</sup>	-	-	-	-
3	BK.IR.014 <sup>f</sup>	-	-	-	-
3	BK.IR.019 <sup>f</sup>	-	-	-	-
2	NZ.IR.009 <sup>f</sup>	-	-	-	-

## NOTES

<sup>\*</sup> *PCA group* refers to the linear arrangements (numbered 1-3) as observed from top to bottom in a PCA plot of the 1,078 SNP loci on chromosome 2L in the six *An. gambiae* populations prior to dataset pruning (N=96)

<sup>‡</sup> Each mosquito is individually labeled (*sample ID*) according to entomological sampling site; collection method; other information i.e. village number (1/2) or date (M+1, which refers to a second collection one month after the first), and sample number in the collection. Entomological sampling sites are as follows: Bukasa (BK), Bugala (BL), Entebbe (EB), Nsadzi (NZ), Sserinya (SY), and Lake Wamala (WL). Collection methods refer to indoor resting (IR) or larval collection (LC)

<sup>§</sup> *Allele length* identifies 2La karyotypes by molecular weight of the DNA fragment in base pairs (bp) (B. White et al. 2007) through visualization of electrophoretic band sizes using the following diagnostic criteria: 207 bp (a<sup>+</sup>), 492 bp (a), and 207/492 bp (a<sup>+</sup>/a)

<sup>a</sup> *PCR results* were categorized as: no-calls (no band produced), false-calls (bands produced but inconsistent with karyotype), and correct-calls (bands produced consistent with karyotype)

<sup>b</sup> *Sample karyotype*: 2L refers to the left arm of the second chromosome; *a* is the four subdivisions (23-26 inclusive) length of the inversion (George, Sharakhova, and Sharakhov 2010)

## TABLE C.1 (CONTINUED)

<sup>c</sup> *Sample phenotype*: 2La phenotypes are categorized according to karyotype arrangements as described in allele length above. The homokaryotypic *standard* arrangement produces a single electrophoretic band at 207 bp (2La<sup>+</sup>/a<sup>+</sup>); *heterokaryotypes* produce two bands at 207, and 492 bp (2La<sup>+</sup>/a), while homokaryotypic *inverted* arrangements produce a single band at 492 bp (2La/a)

<sup>d</sup> Samples generating false-call PCR results amplified identical results upon repeat assay. The 650 bp product was classified as a PCR artifact

<sup>e</sup> Despite repeated attempts an unidentified 650 bp product was produced for sample BL.IR.011, which was inconsistent with any karyotype, and classified as a PCR artifact

<sup>f</sup> These samples were not physically processed due to human error



## APPENDIX D:

BIOINFORMATICS PIPELINE DETAILING SOFTWARE PROGRAMS AND PARAMETERS  
IMPLEMENTED IN THE ANALYSIS OF THE *ANOPHELES FARAUTI* WHOLE GENOME  
SEQUENCED DATASET

TABLE D.1

SOFTWARE AND PARAMETERS APPLIED IN THE BIOINFORMATICS ANALYSIS OF 16 WILD-TYPE *ANOPHELES FARAUTI*  
INDIVIDUALS SEQUENCED BY A WHOLE GENOME SHOTGUN APPROACH

<i>Software</i>	<i>Version</i>	<i>Function</i>	<i>Parameters</i>	<i>Reference</i>
<i>FastQC</i>	<i>0.10.1</i>	Sequencing quality check		(Babraham Institute 2011)
<i>Trimmomatic</i>	<i>0.30</i>	Remove Illumina sequencing adapters	ILLUMINACLIP:TruSeq3-PE.fa:2:30:10 LEADING:5 TRAILING:5 SLIDINGWINDOW:4:15 MINLEN:50	(Bolger, Lohse, and Usadel 2014)
<i>Burrows-Wheeler Alignment (BWA)</i>	<i>0.7.5</i>	Alignment of reads to <i>AfarF1</i> reference genome	-t 12 -q 5 -l 32 -k 3 -n 9 -o 1	(Li and Durbin 2009)
<i>SAMtools</i>	<i>0.1.18</i>	Pre-processing of reads		(Li et al. 2009)
<i>Picard Tools</i>	<i>1.92</i>	Pre-processing of reads	CleanSam -VALIDATION_STRINGENCY=LENIENT SortSam -SO=coordinate -VALIDATION_STRINGENCY=LENIENT	(Broad Institute 2013)

TABLE D.1 (CONTINUED)

<i>Software</i>	<i>Version</i>	<i>Function</i>	<i>Parameters</i>	<i>Reference</i>
			BuildBamIndex -VALIDATION_STRINGENCY=LENIENT CreateSequenceDictionary	
<i>GenomeAnalysis ToolKit (GATK)</i>	2.5.2	Variant calling	-T HaplotypeCaller --downsampling_type BY_SAMPLE --downsample_to_coverage 250 --genotyping_mode DISCOVERY --min_mapping_quality_score 20	(Van der Auwera et al. 2013)
<i>GATK</i>	2.5.2	Quality filtering of variant calls; selects SNPs passing filters, which are biallelic only	-T VariantFiltration --filterExpression "QD<2.0  MQ<40.0  FS>60.0  MQRankSum<12.5  ReadPosRankSum<-8.0" "QD<2.0  FS>200.0  ReadPosRankSum<20.0" -T SelectVariants --selectType SNP --excludeFiltered --restrictAllelesTo BIALLELIC	(Van der Auwera et al. 2013)
<i>VCftools</i>	0.1.15	Dataset pruning	--minGQ 30.0 ( <u>G</u> enotype <u>Q</u> uality of at least 30) --maf 0.05 ( <u>M</u> inor <u>A</u> llele <u>F</u> requency of 5%)	(Danecek et al. 2011)
<i>sharedVariants AllSamples.py</i>	<i>Custom</i>	Selects SNPs present in all samples		(Wiltshire 2019)
<i>PLINK</i>	1.9	Formats data for PCA and global	--allow-extra-chr --pca	(Chang et al. 2015)

TABLE D.1 (CONTINUED)

<i>Software</i>	<i>Version</i>	<i>Function</i>	<i>Parameters</i>	<i>Reference</i>
		ancestry analysis	--recode 12	
<i>R</i>	<i>3.2.1</i>	Apply plot functions to visualize data		(R Core Team 2014)
<i>ADMIXTURE</i>	<i>1.2.3</i>	Global ancestry analysis		(Alexander, Novembre, and Lange 2009)

## BIBLIOGRAPHY

- Aldrick, John. 1993. "The Susceptibility of Lands to Deterioration in the Solomon Islands." Working Paper 12. Honiara: Australian International Development Assistance Bureau, Ministry of Natural Resources.
- Alexander, David, John Novembre, and Kenneth Lange. 2009. "Fast Model-Based Estimation of Ancestry in Unrelated Individuals." *Genome Research* 19 (9): 1655–64. <https://doi.org/10.1101/gr.094052.109>.
- Alonso, Pedro, Graham Brown, Myriam Arevalo-Herrera, Fred Binka, Chetan Chitnis, Frank Collins, Ogobara Doumbo, et al. 2011. "A Research Agenda to Underpin Malaria Eradication." *PLoS Medicine* 8 (1): e1000406. <https://doi.org/10.1371/journal.pmed.1000406>.
- Ambrose, Luke, Robert Cooper, Tanya Russell, Thomas Burkot, Neil Lobo, Frank Collins, Jeffrey Hii, and Nigel Beebe. 2014. "Microsatellite and Mitochondrial Markers Reveal Strong Gene Flow Barriers for *Anopheles farauti* in the Solomon Archipelago: Implications for Malaria Vector Control." *International Journal for Parasitology* 44 (March): 225–33. <https://doi.org/10.1016/j.ijpara.2013.12.001>.
- Ambrose, Luke, C. Riginos, Robert Cooper, K. S. Leow, W. Ong, and Nigel Beebe. 2012. "Population Structure, Mitochondrial Polyphyly and the Repeated Loss of Human Biting Ability in Anopheline Mosquitoes from the Southwest Pacific." *Molecular Ecology* 21 (17): 4327–43. <https://doi.org/10.1111/j.1365-294X.2012.05690.x>.
- Amenya, D, R Naguran, T-C Lo, Hilary Ranson, B Spellings, O Wood, Basil Brooke, Maureen Coetzee, and Lizette Koekemoer. 2008. "Over Expression of a Cytochrome P450 (CYP6P9) in a Major African Malaria Vector, *Anopheles funestus*, Resistant to Pyrethroids." *Insect Molecular Biology* 17 (1): 19–25. <https://doi.org/10.1111/j.1365-2583.2008.00776.x>.
- Anderson, Atholl, and Sue O'Connor. 2008. "Indo-Pacific Migration and Colonization—Introduction." *Asian Perspectives* 47 (1): 2–11.
- Arnold, B., R. B. Corbett-Detig, Daniel Hartl, and K. Bomblies. 2013. "RADseq Underestimates Diversity and Introduces Genealogical Biases Due to

- Nonrandom Haplotype Sampling." *Molecular Ecology* 22 (11): 3179–90.  
<https://doi.org/10.1111/mec.12276>.
- Athrey, Giridhar, Theresa Hodges, Michael Reddy, Hans Overgaard, Abrahan Matias, Frances Ridl, Immo Kleinschmidt, Adalgisa Caccone, and Michel Slotman. 2012. "The Effective Population Size of Malaria Mosquitoes: Large Impact of Vector Control." *PLoS Genetics* 8 (12): e1003097.  
<https://doi.org/10.1371/journal.pgen.1003097>.
- Australian Government Bureau of Meteorology. 2018. "Pacific Climate Change Data Portal." 2018. <http://www.bom.gov.au/climate/pccsp/>.
- Avery, James. 1973. "A Review of the Malaria Eradication Programme in the British Solomon Islands 1970-1972." *Papua New Guinea Medical Journal* 17 (1): 50–60.
- Aweeka, Francesca, and Polina German. 2008. "Clinical Pharmacology of Artemisinin-Based Combination Therapies." *Clinical Pharmacokinetics* 476 (2): 91–102.
- Ayad, Hafez, and George Georghiou. 1979. "Resistance Pattern of *Anopheles albimanus* Wied. Following Selection by Parathion." *Mosquito News* 39 (March): 121–25.
- Babraham Institute. 2011. *FastQC: A Quality Control Tool for High-Throughput Sequence Data*. English.  
<http://www.bioinformatics.babraham.ac.uk/projects/fastqc/>.
- Baird, Nathan, Paul Etter, Tressa Atwood, Mark Currey, Anthony Shiver, Zachary Lewis, Eric Selker, William Cresko, and Eric Johnson. 2008. "Rapid SNP Discovery and Genetic Mapping Using Sequenced RAD Markers." *PLoS ONE* 3 (10): e3376. <https://doi.org/10.1371/journal.pone.0003376>.
- Balabanidou, Vasileia, Anastasia Kampouraki, Marina MacLean, Gary Blomquist, Claus Tittiger, M. Patricia Juárez, Sergio Mijailovsky, et al. 2016. "Cytochrome P450 Associated with Insecticide Resistance Catalyzes Cuticular Hydrocarbon Production in *Anopheles gambiae*." *Proceedings of the National Academy of Sciences* 113 (33): 9268–73.  
<https://doi.org/10.1073/pnas.1608295113>.
- Baruch, Dror, Britten Pasloske, Hardeep Singh, Xiahui Bi, Xin Ma, Michael Feldman, Theodore Taraschi, and Russell Howard. 1995. "Cloning the *P. falciparum* Gene Encoding PfEMP1, a Malarial Variant Antigen and Adherence Receptor on the Surface of Parasitized Human Erythrocytes." *Cell* 82 (1): 77–87.  
[https://doi.org/10.1016/0092-8674\(95\)90054-3](https://doi.org/10.1016/0092-8674(95)90054-3).

- Bass, Chris, and Linda Field. 2011. "Gene Amplification and Insecticide Resistance." *Pest Management Science* 67 (8): 886–90. <https://doi.org/10.1002/ps.2189>.
- Beebe, Nigel, and Robert Cooper. 2002. "Distribution and Evolution of the *Anopheles punctulatus* Group (Diptera: Culicidae) in Australia and Papua New Guinea." *International Journal for Parasitology* 32 (5): 563–74.
- Beebe, Nigel, Robert Cooper, Desmond Foley, and John Ellis. 2000. "Populations of the South-West Pacific Malaria Vector *Anopheles farauti* s.s. Revealed by Ribosomal DNA Transcribed Spacer Polymorphisms." *Heredity* 84 (February): 244–53.
- Beebe, Nigel, Tanya Russell, Thomas Burkot, Neil Lobo, and Robert Cooper. 2013. "Chapter 12. The Systematics and Bionomics of Malaria Vectors in the Southwest Pacific." In *Anopheles Mosquitoes - New Insights into Malaria Vectors*, edited by Sylvie Manguin, 357–94. London: InTechOpen. <https://www.intechopen.com/books/anopheles-mosquitoes-new-insights-into-malaria-vectors/the-systematics-and-bionomics-of-malaria-vectors-in-the-southwest-pacific>.
- Beebe, Nigel, and Allan Saul. 1995. "Discrimination of All Members of the *Anopheles punctulatus* Complex by Polymerase Chain Reaction-Restriction Fragment Length Polymorphism Analysis." *American Journal of Tropical Medicine and Hygiene* 53 (5): 478–81. <https://doi.org/10.4269/ajtmh.1995.53.478>.
- Belkin, John. 1962a. *Mosquitoes of the South Pacific I*. Berkeley: University of California Press.
- . 1962b. *Mosquitoes of the South Pacific II*. Berkeley: University of California Press.
- Bennett, G., P. Cohen, A. M. Schwarz, J. Albert, S. Lawless, C. Paul, and Z. Hilly. 2014. "Solomon Islands: Western Province Situation Analysis. Project Report: AAS-2014-15." Penang, Malaysia: CGIAR Research Program on Aquatic Agricultural Systems. [http://pubs.iclarm.net/resource\\_centre/AAS-2014-15.pdf](http://pubs.iclarm.net/resource_centre/AAS-2014-15.pdf).
- Bergey, Christina, Martin Lukindu, Rachel Wiltshire, Michael Fontaine, Jonathan Kayondo, and Nora Besansky. 2019. "Assessing Connectivity despite High Diversity in Island Populations of the Malaria Mosquito *Anopheles gambiae*." *BioRxiv*, February. <https://doi.org/10.1101/430702>.
- Besansky, Nora, Tovi Lehmann, G. Thomas Fahey, Didier Fontenille, Lawrence Braack, William Hawley, and Frank Collins. 1997. "Patterns of Mitochondrial Variation Within and Between African Malaria Vectors, *Anopheles gambiae* and *An. arabiensis*, Suggest Extensive Gene Flow." *Genetics* 147 (4): 1817–28.

- Bhatt, Samir, D. Weiss, E. Cameron, D. Bisanzio, B. Mappin, U. Dalrymple, Katherine Battle, et al. 2015. "The Effect of Malaria Control on *Plasmodium falciparum* in Africa between 2000 and 2015." *Nature* 526 (7572): 207–11. <https://doi.org/10.1038/nature15535>.
- Bian, Guowu, Deepak Joshi, Yuemei Dong, Peng Lu, Guoli Zhou, Xiaoling Pan, Yao Xu, George Dimopoulos, and Zhiyong Xi. 2013. "*Wolbachia* Invades *Anopheles stephensi* Populations and Induces Refractoriness to *Plasmodium* Infection." *Science* 340 (6133): 748–51. <https://doi.org/10.1126/science.1236192>.
- Bolger, Anthony, Marc Lohse, and Bjoern Usadel. 2014. "Trimmomatic: A Flexible Trimmer for Illumina Sequence Data." *Bioinformatics* 30 (15): 2114–20. <https://doi.org/10.1093/bioinformatics/btu170>.
- Bower, James, Mark Dowton, Robert Cooper, and Nigel Beebe. 2008. "Intraspecific Concerted Evolution of the rDNA ITS1 in *Anopheles farauti sensu stricto* (Diptera: Culicidae) Reveals Recent Patterns of Population Structure." *Journal of Molecular Evolution* 67 (4): 397–411. <https://doi.org/10.1007/s00239-008-9161-x>.
- Broad Institute. 2012a. *GAEMR* (version 1.0.1). Python. Cambridge, MA. <http://software.broadinstitute.org/software/gaemr/>.
- . 2012b. *Pilon*. Cambridge, MA. <https://github.com/broadinstitute/pilon/wiki>.
- . 2013. *Picard* (version 1.92). Java. Cambridge, MA. <https://github.com/broadinstitute/picard>.
- . 2017. *Picard*. Java. Cambridge, MA. <http://broadinstitute.github.io/picard/>.
- Bruce-Chwatt, Leonard. 1965. "Paleogenesis and Paleo-Epidemiology of Primate Malaria." *Bulletin of the World Health Organization* 32 (3): 363–87.
- Bryan, Joan, and Mario Coluzzi. 1971. "Cytogenetic Observations on *Anopheles farauti* Laveran." *Bulletin of the World Health Organization* 45 (2): 266–67.
- Burkot, Camilla, and Katherine Gilbert. 2017. "Reducing Malaria in Solomon Islands: Lessons for Effective Aid." Development Policy Centre Discussion Paper #64. Canberra: Crawford School of Public Policy, The Australian National University. <https://poseidon01.ssrn.com/delivery.php?ID=7470950950080000040810740100880150960540070080810880200251220070230091120950701240060380630220140490400050040030070130650960290230150690180030260930030000020891200950190930770071171250900070240020790>



30074097069007097008109100124073004005104115117065031&EXT=pdf.

- Burkot, Thomas, Hugo Bugoro, Allan Apairamo, Robert Cooper, Diego Echeverry, Danyal Odabasi, Nigel Beebe, et al. 2018. "Spatial-Temporal Heterogeneity in Malaria Receptivity Is Best Estimated by Vector Biting Rates in Areas Nearing Elimination." *Parasites & Vectors* 11 (November): 606.  
<https://doi.org/10.1186/s13071-018-3201-1>.
- Burkot, Thomas, Tanya Russell, Lisa Reimer, Hugo Bugoro, Nigel Beebe, Robert Cooper, Supraman Sukawati, Frank Collins, and Neil Lobo. 2013. "Barrier Screens: A Method to Sample Blood-Fed and Host-Seeking Exophilic Mosquitoes." *Malaria Journal* 12 (February): 49.  
<https://doi.org/10.1186/1475-2875-12-49>.
- Burt, Austin. 2014. "Heritable Strategies for Controlling Insect Vectors of Disease." *Philosophical Transactions of the Royal Society B* 369 (1645): 20130432.  
<https://doi.org/10.1098/rstb.2013.0432>.
- Busvine, James. 1978. "Current Problems in the Control of Mosquitoes." *Nature* 273 (June): 604–7. <https://doi.org/10.1038/273604a0>.
- Busvine, James, and Raman Pal. 1969. "The Impact on Insecticide Resistance on Control of Vectors and Vector-Borne Diseases." *Bulletin of the World Health Organization* 40 (5): 731–44.
- Capoluongo, Ettore. 2018. "Global Variome Shared LOVD G6PD." Database. Leiden Open Global Variation Database v.3.0 Build 21. 2018.  
<https://databases.lovd.nl/shared/genes/G6PD>.
- Carmody, Pádraig, and David Taylor. 2016. "Globalization, Land Grabbing, and the Present-Day Colonial State in Uganda: Ecolonization and Its Impacts." *Journal of Environment & Development* 25 (1): 100–126.  
<https://doi.org/10.1177/1070496515622017>.
- Carrasco, David, Thierry Lefèvre, Nicolas Moiroux, Cédric Penner, Fabrice Chandre, and Anna Cohuet. 2019. "Behavioural Adaptations of Mosquito Vectors to Insecticide Control." *Current Opinion in Insect Science* 34 (March): In press. <https://doi.org/10.1016/j.cois.2019.03.005>.
- Carter, Richard, and Kamini Mendis. 2002. "Evolutionary and Historical Aspects of the Burden of Malaria." *Clinical Microbiology Reviews* 15 (4): 564–94.  
<https://doi.org/10.1128/CMR.15.4.564-594.2002>.
- Chang, Christopher, Carson Chow, Laurent Tellier, Shashaank Vattikuti, Shaun Purcell, and James Lee. 2015. "Second-Generation PLINK: Rising to the

- Challenge of Larger and Richer Datasets." *GigaScience* 4 (1): 7.  
<https://doi.org/10.1186/s13742-015-0047-8>.
- Chen, Hong, Noboru Minakawa, John Beier, and Guiyun Yan. 2004. "Population Genetic Structure of *Anopheles gambiae* Mosquitoes on Lake Victoria Islands, West Kenya." *Malaria Journal* 3 (December): 48.  
<https://doi.org/10.1186/1475-2875-3-48>.
- Cheng, Changde, Bradley White, Colince Kamdem, Keithanne Mockaitis, Carlo Costantini, Matthew Hahn, and Nora Besansky. 2011. "Ecological Genomics of *Anopheles gambiae* Along a Latitudinal Cline: A Population-Resequencing Approach." *Genetics* 190 (4): 1417–32.  
<https://doi.org/10.1534/genetics.111.137794>.
- Christophides, George. 2005. "Transgenic Mosquitoes and Malaria Transmission." *Cellular Microbiology* 7 (3): 325–33. <https://doi.org/10.1111/j.1462-5822.2005.00495.x>.
- Clark-Carter, David. 2005. "Harmonic Mean." In *Encyclopedia of Statistics in Behavioral Science*, edited by Brian Everitt and David Howell, 2:783–84. Chichester: John Wiley & Sons, Ltd.  
<https://onlinelibrary.wiley.com/doi/pdf/10.1002/0470013192.bsa377>.
- Collins, Frank, and Susan Paskewitz. 1996. "A Review of the Use of Ribosomal DNA (rDNA) to Differentiate among Cryptic *Anopheles* Species." *Insect Molecular Biology* 5 (1): 1–9. <https://doi.org/10.1111/j.1365-2583.1996.tb00034.x>.
- Coluzzi, Mario, A. Sabatini, Vincenzo Petrarca, and Maria di Deco. 1979. "Chromosomal Differentiation and Adaptation to Human Environments in the *Anopheles gambiae* Complex." *Transactions of the Royal Society of Tropical Medicine and Hygiene* 73 (5): 483–97. [https://doi.org/10.1016/0035-9203\(79\)90036-1](https://doi.org/10.1016/0035-9203(79)90036-1).
- Conteh, Lesong, Elisa Sicuri, Fatuma Manzi, Guy Hutton, Benson Obonyo, Fabrizio Tediosi, Prosper Biao, et al. 2010. "The Cost-Effectiveness of Intermittent Preventive Treatment for Malaria in Infants in Sub-Saharan Africa." *PLoS ONE* 5 (6): e10313. <https://doi.org/10.1371/journal.pone.0010313>.
- Curtis, Chris. 1968. "Possible Use of Translocations to Fix Desirable Genes in Insect Pest Populations." *Nature* 218 (April): 368–69.  
<https://doi.org/10.1038/218368a0>.
- Danecek, Petr, Adam Auton, Goncalo Abecasis, Cornelis Albers, Eric Banks, Mark DePristo, Bob Handsaker, et al. 2011. "The Variant Call Format and VCFtools." *Bioinformatics* 27 (15): 2156–58.  
<https://doi.org/10.1093/bioinformatics/btr330>.

- Davidson, G. 1962. "*Anopheles gambiae* Complex." *Nature* 196 (4857): 907.  
<https://doi.org/10.1038/196907a0>.
- . 1964. "*Anopheles gambiae*, a Complex of Species." *Bulletin of the World Health Organization* 31 (5): 625–34.
- Dekker, Teun, Willem Takken, and Ring Cardé. 2001. "Structure of Host-Odour Plumes Influences Catch of *Anopheles gambiae* s. s. and *Aedes aegypti* in a Dual-Choice Olfactometer." *Physiological Entomology* 26 (2): 124–34.  
<https://doi.org/10.1046/j.1365-3032.2001.00225.x>.
- Dethier, V. G., L. B. Browne, and C. N. Smith. 1960. "The Designation of Chemicals in Terms of the Responses They Elicit from Insects." *Journal of Economic Entomology* 53 (1): 134–36. <https://doi.org/10.1093/jee/53.1.134>.
- Ding, Yunchuan, Nicola Hawkes, Janet Meredith, Paul Eggleston, Janet Hemingway, and Hilary Ranson. 2005. "Characterization of the Promoters of Epsilon Glutathione Transferases in the Mosquito *Anopheles gambiae* and Their Response to Oxidative Stress." *Biochemical Journal* 387 (3): 879–88.  
<https://doi.org/10.1042/BJ20041850>.
- Ding, Yunchuan, Federica Ortelli, Louise Rossiter, Janet Hemingway, and Hilary Ranson. 2003. "The *Anopheles gambiae* Glutathione Transferase Supergene Family: Annotation, Phylogeny and Expression Profiles." *BMC Genomics* 4 (August): 35. <https://doi.org/10.1186/1471-2164-4-35>.
- Do, Chi, Robin Waples, D. Peel, G. M. Macbeth, B. J. Tillett, and J. R. Ovenden. 2014. "NeEstimator v2: Re-Implementation of Software for the Estimation of Contemporary Effective Populations Size (Ne) from Genetic Data." *Molecular Ecology Resources* 14 (1): 209–14. <https://doi.org/10.1111/1755-0998.12157>.
- Dönitz, Wilhelm. 1901. "Nachrichten aus dem Berliner Entomologischen Verein." *Insektenbörse* 18: 36–38.
- Du, W, Taiwo Awolola, Paul Howell, Lizette Koekemoer, Basil Brooke, Mark Benedict, Maureen Coetzee, and Liangbiao Zheng. 2005. "Independent Mutations in the Rdl Locus Confer Dieldrin Resistance to *Anopheles gambiae* and *An. arabiensis*." *Insect Molecular Biology* 14 (2): 179–83.  
<https://doi.org/10.1111/j.1365-2583.2005.00544.x>.
- Edi, Constance, Luc Djogbénou, Adam Jenkins, Kimberly Regna, Marc Muskavitch, Rodolphe Poupardin, Christopher Jones, et al. 2014. "CYP6 P450 Enzymes and ACE-1 Duplication Produce Extreme and Multiple Insecticide Resistance in the Malaria Mosquito *Anopheles gambiae*." *PLoS Genetics* 10 (3): e1004236.  
<https://doi.org/10.1371/journal.pgen.1004236>.

- Essandoh, John, Alexander Yawson, and David Weetman. 2013. "Acetylcholinesterase (Ace-1) Target Site Mutation 119S Is Strongly Diagnostic of Carbamate and Organophosphate Resistance in *Anopheles gambiae* s.s. and *Anopheles coluzzii* across Southern Ghana." *Malaria Journal* 12 (November): 404. <https://doi.org/10.1186/1475-2875-12-404>.
- Failloux, Anna-Bella, Michel Raymond, Andre Ung, Christine Chevillon, and Nicole Pasteur. 1997. "Genetic Differentiation Associated with Commercial Traffic in the Polynesian Mosquito, *Aedes polynesiensis* Marks 1951." *Biological Journal of the Linnean Society* 60 (1): 107–18. <https://doi.org/10.1111/j.1095-8312.1997.tb01486.x>.
- Fillinger, Ulrike, Bryson Ndenga, Andrew Githeko, and Steven Lindsay. 2009. "Integrated Malaria Vector Control with Microbial Larvicides and Insecticide-Treated Nets in Western Kenya: A Controlled Trial." *Bulletin of the World Health Organization* 87: 655–65. <https://doi.org/10.2471/BLT.08.055632>.
- Flannery, Tim. 1995. *Mammals of the South-West Pacific & Moluccan Islands*. New York: Cornell University Press.
- Foley, Desmond, Robert Cooper, and Joan Bryan. 1995. "A New Species within the *Anopheles punctulatus* Complex in Western Province, Papua New Guinea." *Journal of the American Mosquito Control Association* 11 (1): 122–27.
- Foley, Desmond, S. R. Meek, and Joan Bryan. 1994. "The *Anopheles punctulatus* Group of Mosquitoes in the Solomon Islands and Vanuatu Surveyed by Allozyme Electrophoresis." *Medical and Veterinary Entomology* 8 (4): 340–50. <https://doi.org/10.1111/j.1365-2915.1994.tb00098.x>.
- Foley, Desmond, R. Paru, H. Dagoro, and Joan Bryan. 1993. "Allozyme Analysis Reveals Six Species within the *Anopheles punctulatus* Complex of Mosquitoes in Papua New Guinea." *Medical and Veterinary Entomology* 7 (1): 37–48. <https://doi.org/10.1111/j.1365-2915.1993.tb00649.x>.
- Fontaine, Michael, James Pease, Aaron Steele, Robert Waterhouse, Daniel Neafsey, Igor Sharakov, Xiaofang Jiang, et al. 2015. "Extensive Introgression in a Malaria Vector Species Complex Revealed by Phylogenomics." *Science* 347 (6217): 1258524. <https://doi.org/10.1126/science.1258524>.
- Fonzi, Eugenio, Yukiko Higa, Arlene Bertuso, Kyoko Futami, and Noboru Minakawa. 2015. "Human-Mediated Marine Dispersal Influences the Population Structure of *Aedes aegypti* in the Philippine Archipelago." *PLoS Neglected Tropical Diseases* 9 (6): e0003829. <https://doi.org/10.1371/journal.pntd.0003829>.

- Gantz, Valentino, Nijole Jasinskiene, Olga Tatarenkova, Aniko Fazekas, Vanessa Macias, Ethan Bier, and Anthony James. 2015. "Highly Efficient Cas9-Mediated Gene Drive for Population Modification of the Malaria Vector Mosquito *Anopheles stephensi*." *Proceedings of the National Academy of Sciences* 112 (49): E6736–43. <https://doi.org/10.1073/pnas.1521077112>.
- Garrett-Jones, C. 1964. "The Prognosis for Interruption of Malaria Transmission through Assessment of the Mosquito's Vectorial Capacity." *Nature* 204 (4964): 1173–75.
- Gatton, Michelle, Nakul Chitnis, Thomas Churcher, Martin Donnelly, Azra Ghani, H. Charles Godfray, Fred Gould, et al. 2013. "The Importance of Mosquito Behavioural Adaptations to Malaria Control in Africa." *Evolution* 67 (4): 1218–30. <https://doi.org/10.1111/evo.12063>.
- Geier, Martin, Oliver Bosch, and Jürgen Boeckh. 1999. "Influence of Odour Plume Structure on Upwind Flight of Mosquitoes Towards Hosts." *Journal of Experimental Biology* 202: 1639–48.
- George, P., Maria Sharakhova, and Igor Sharakhov. 2010. "High-Resolution Cytogenetic Map for the African Malaria Vector *Anopheles gambiae*." *Insect Molecular Biology* 19 (5): 675–82. <https://doi.org/10.1111/j.1365-2583.2010.01025.x>.
- Gilbert, Kimberley, and Michael Whitlock. 2015. "Evaluating Methods for Estimating Local Effective Population Size with and without Migration." *Evolution* 69 (8): 2154–66. <https://doi.org/10.1111/evo.12713>.
- Gillies, Michael. 1961. "Studies on the Dispersion and Survival of *Anopheles gambiae* Giles in East Africa, by Means of Marking and Release Experiments." *Bulletin of Entomological Research* 52 (1): 99–127. <https://doi.org/10.1017/S0007485300055309>.
- . 1980. "The Role of Carbon Dioxide in Host-Finding by Mosquitoes (Diptera: Culicidae): A Review." *Bulletin of Entomological Research* 70: 525–32.
- Gillies, Michael, and Botha de Meillon. 1968. *The Anophelinae of Africa South of the Sahara (Publication No. 54)*. Johannesburg: South African Institute for Medical Research.
- Gnerre, Sante, Iain MacCallum, Dariusz Przybylski, Filipe Ribeiro, Joshua Burton, Bruce Walker, Ted Sharpe, et al. 2011. "High-Quality Draft Assemblies of Mammalian Genomes from Massively Parallel Sequence Data." *Proceedings of the National Academy of Sciences* 108 (4): 1513–18. <https://doi.org/10.1073/pnas.1017351108>.

- Grant, AJ, BE Wigton, JG Aghajanian, and RJ O'Connell. 1995. "Electrophysiological Responses of Receptor Neurons in Mosquito Maxillary Palp Sensilla to Carbon Dioxide." *Journal of Comparative Physiology A: Neuroethology, Sensory, Neural, and Behavioral Physiology* 177: 389–96.
- Hammond, Andrew, Roberto Galizi, Kyros Kyrou, Alekos Simoni, Carla Siniscalchi, Dimitris Katsanos, Matthew Gribble, et al. 2016. "A CRISPR-Cas9 Gene Drive System Targeting Female Reproduction in the Malaria Mosquito Vector *Anopheles gambiae*." *Nature Biotechnology* 34 (1): 78–83. <https://doi.org/10.1038/nbt.3439>.
- Hay, Simon, David Smith, and Robert Snow. 2008. "Measuring Malaria Endemicity from Intense to Interrupted Transmission." *The Lancet Infectious Diseases* 8 (6): 369–78. [https://doi.org/10.1016/S1473-3099\(08\)70069-0](https://doi.org/10.1016/S1473-3099(08)70069-0).
- Hemingway, Janet. 1982. "The Biochemical Nature of Malathion Resistance in *Anopheles stephensi* from Pakistan." *Pesticide Biochemistry and Physiology* 17 (2): 149–55. [https://doi.org/10.1016/0048-3575\(82\)90019-0](https://doi.org/10.1016/0048-3575(82)90019-0).
- . 1983. "Biochemical Studies on Malathion Resistance in *Anopheles arabiensis* from Sudan." *Transactions of the Royal Society of Tropical Medicine and Hygiene* 77 (4): 477–80. [https://doi.org/10.1016/0035-9203\(83\)90118-9](https://doi.org/10.1016/0035-9203(83)90118-9).
- Hoeppli, R. 1959. *Parasites and Parasitic Infections in Early Medicine and Science*. Singapore: University of Malaya Press.
- Holt, Robert, G. Mani Subramanian, Aaron Halpern, Granger Sutton, Rosane Charlab, Deborah Nusskern, Patrick Wincker, et al. 2002. "The Genome Sequence of the Malaria Mosquito *Anopheles gambiae*." *Science* 298 (5591): 129–49. <https://doi.org/10.1126/science.1076181>.
- Howes, Rosalind, Frédéric Piel, Anand Patil, Oscar Nyangiri, Peter Gething, Mewahyu Dewi, Mariana Hogg, et al. 2012. "G6PD Deficiency Prevalence and Estimates of Affected Populations in Malaria Endemic Countries: A Geostatistical Model-Based Map." *PLoS Medicine* 9 (11): e1001339. <https://doi.org/10.1371/journal.pmed.1001339>.
- Ingham, Victoria, Patricia Pignatelli, Jonathan Moore, Simon Wagstaff, and Hilary Ranson. 2017. "The Transcription Factor Maf-S Regulates Metabolic Resistance to Insecticides in the Malaria Vector *Anopheles gambiae*." *BMC Genomics* 18 (August): 669. <https://doi.org/10.1186/s12864-017-4086-7>.
- Jaffrés, Jasmine, Chris Cuff, Cecily Rasmussen, and Aimee Hesson. 2018. "Teleconnection of Atmospheric and Oceanic Climate Anomalies with Australian Weather Patterns: A Review of Data Availability." *Earth-Science*

Reviews 176 (January): 117–46.  
<https://doi.org/10.1016/j.earscirev.2017.08.010>.

- James, Stefanie, Frank Collins, Philip Welkhoff, Claudia Emerson, Charles Godfray, Michael Gottlieb, Brian Greenwood, et al. 2018. "Pathway to Deployment of Gene Drive Mosquitoes as a Potential Biocontrol Tool for Elimination of Malaria in Sub-Saharan Africa: Recommendations of a Scientific Working Group." *American Journal of Tropical Medicine and Hygiene* 98 ((Suppl 6)): 1–49. <https://doi.org/10.4269/ajtmh.18-0083>.
- Kabbale, Fredrick, Anne Akol, John Kaddu, and Ambrose Onapa. 2013. "Biting Patterns and Seasonality of *Anopheles gambiae sensu lato* and *Anopheles funestus* Mosquitoes in Kamuli District, Uganda." *Parasites & Vectors* 6 (1): 340. <https://doi.org/10.1186/1756-3305-6-340>.
- Katovai, Eric. 2016. "Regeneration and Recovery Dynamics of Logged Forests in the Solomon Islands." PhD, Australia: James Cook University.  
<https://researchonline.jcu.edu.au/49985/1/49985-katovai-2016-thesis.pdf>.
- Katovai, Eric, Alana Burley, and Margaret Mayfield. 2012. "Understory Plant Species and Functional Diversity in the Degraded Wet Tropical Forests of Kolombangara Island, Solomon Islands." *Biological Conservation* 145 (1): 214–24. <https://doi.org/10.1016/j.biocon.2011.11.008>.
- Kawada, Hitoshi, Kazunori Ohashi, Gabriel Dida, George Sonye, Sammy Njenga, Charles Mwandawiro, and Noboru Minakawa. 2014. "Insecticidal and Repellent Activities of Pyrethroids to the Three Major Pyrethroid-Resistant Malaria Vectors in Western Kenya." *Parasites & Vectors* 7 (May): 208. <https://doi.org/10.1186/1756-3305-7-208>.
- Kayentao, Kassoum, Paul Garner, Anne Maria van Eijk, Inbarani Naidoo, Cally Roper, Abdunoor Mulokozi, John MacArthur, et al. 2013. "Intermittent Preventive Therapy for Malaria during Pregnancy Using 2 vs 3 or More Doses of Sulfadoxine-Pyrimethamine and Risk of Low Birth Weight in Africa." *Journal of the American Medical Association* 309 (6): 594–604. <https://doi.org/10.1001/jama.2012.216231>.
- Kayondo, Jonathan, Louis Mukwaya, Aram Stump, Andrew Michel, Mamadou Coulibaly, Nora Besansky, and Frank Collins. 2005. "Genetic Structure of *Anopheles gambiae* Populations on Islands in Northwestern Lake Victoria, Uganda." *Malaria Journal* 4 (December): 59. <https://doi.org/10.1186/1475-2875-4-59>.
- Kayser, Manfred. 2010. "The Human Genetic History of Oceania: Near and Remote Views of Dispersal." *Current Biology* 20 (4): R194–201. <https://doi.org/10.1016/j.cub.2009.12.004>.

- Kellogg, FE. 1970. "Water Vapour and Carbon Dioxide Receptors in *Aedes aegypti*." *Journal of Insect Physiology* 16: 99–108.
- Koekemoer, Lizette, Luna Kamau, Richard Hunt, and Maureen Coetzee. 2002. "A Cocktail Polymerase Chain Reaction Assay to Identify Members of the *Anopheles funestus* (Diptera: Culicidae) Group." *American Journal of Tropical Medicine and Hygiene* 66 (6): 804–11.  
<https://doi.org/10.4269/ajtmh.2002.66.804>.
- Koella, Jacob, Flemming Sørensen, and RA Anderson. 1998. "The Malaria Parasite, *Plasmodium falciparum*, Increases the Frequency of Multiple Feeding of Its Mosquito Vector, *Anopheles gambiae*." *Proceedings of the Royal Society London B* 265 (1398): 763–68. <https://doi.org/10.1098/rspb.1998.0358>.
- Kyrou, Kyros, Andrew Hammond, Roberto Galizi, Nace Kranjc, Austin Burt, Andrea Beaghton, Tony Nolan, and Andrea Crisanti. 2018. "A CRISPR–Cas9 Gene Drive Targeting Doublesex Causes Complete Population Suppression in Caged *Anopheles gambiae* Mosquitoes." *Nature Biotechnology* 36 (September): 1062–66. <https://doi.org/10.1038/nbt.4245>.
- Laveran, NA. 1902. "Sur des culicides de Nouvelles-Hébrides." *Comptes Rendus des Séances et Mémoires de la Société de Biologie* 54: 908–10.
- Lawniczak, Mara, Scott Emrich, A. K. Holloway, A. P. Regier, M. Olson, B. White, Seth Redmond, et al. 2010. "Widespread Divergence between Incipient *Anopheles gambiae* Species Revealed by Whole Genome Sequences." *Science* 330 (6003): 512–14. <https://doi.org/10.1126/science.1195755>.
- Lee, David. 1987. *The Culicidae of the Australasian Region. Volume 5*. Edited by Margaret Debenham. Canberra: Australian Government Publishing Service.
- Lee, Kim-Sung, Paul Divis, Siti Khatijah Zakaria, Asmad Matusop, Roynston Julin, David Conway, Janet Cox-Singh, and Balbir Singh. 2011. "*Plasmodium knowlesi*: Reservoir Hosts and Tracking the Emergence in Humans and Macaques." *PLoS Pathogens* 7 (4): e1002015.  
<https://doi.org/10.1371/journal.ppat.1002015>.
- Lehmann, Tovi, C. R. Blackston, Nora Besansky, A. A. Escalante, Frank Collins, and William Hawley. 2000. "The Rift Valley Complex as a Barrier to Gene Flow for *Anopheles gambiae* in Kenya: The MtDNA Perspective." *Journal of Heredity* 91 (2): 165–68. <https://doi.org/10.1093/jhered/91.2.165>.
- Lehmann, Tovi, William Hawley, Heather Grebert, and Frank Collins. 1998. "The Effective Population Size of *Anopheles gambiae* in Kenya: Implications for Population Structure." *Molecular Biology and Evolution* 15 (3): 264–76.  
<https://doi.org/10.1093/oxfordjournals.molbev.a025923>.



- Lehmann, Tovi, William Hawley, Heather Grebert, M. Danga, F. Atieli, and Frank Collins. 1999. "The Rift Valley Complex as a Barrier to Gene Flow for *Anopheles gambiae* in Kenya." *Journal of Heredity* 90 (6): 613–21. <https://doi.org/10.1093/jhered/90.6.613>.
- Lehmann, Tovi, William Hawley, Luna Kamau, Didier Fontenille, Frédéric Simard, and Frank Collins. 1996. "Genetic Differentiation of *Anopheles gambiae* Populations from East and West Africa: Comparison of Microsatellite and Allozyme Loci." *Heredity* 77: 192–208.
- Li, Heng, and Richard Durbin. 2009. "Fast and Accurate Short Read Alignment with Burrows-Wheeler Transform." *Bioinformatics* 25 (4): 1754–60. <https://doi.org/10.1093/bioinformatics/btp324>.
- Li, Heng, Bob Handsaker, Alec Wysoker, Tim Fennell, Jue Ruan, Nils Homer, Gabor Marth, Goncalo Abecasis, Richard Durbin, and 1000 Genome Project Data Processing Subgroup. 2009. "The Sequence Alignment/Map Format and SAMtools." *Bioinformatics* 25 (16): 2078–79. <https://doi.org/10.1093/bioinformatics/btp352>.
- Lorrey, Andrew, Giovanni Dalu, James Renwick, Howard Diamond, and Marco Gaetani. 2012. "Reconstructing the South Pacific Convergence Zone Position during the Presatellite Era: A La Niña Case Study." *Monthly Weather Review* 140 (November): 3653–68. <https://doi.org/10.1175/MWR-D-11-00228.1>.
- Lukindu, Martin, Christina Bergey, Rachel Wiltshire, Scott Small, Brian Bourke, Jonathan Kayondo, and Nora Besansky. 2018. "Spatio-Temporal Genetic Structure of *Anopheles gambiae* in the Northwestern Lake Victoria Basin, Uganda: Implications for Genetic Control Trials in Malaria Endemic Regions." *Parasites & Vectors* 11 (April): 246. <https://doi.org/10.1186/s13071-018-2826-4>.
- Mahon, Rodney. 1983. "Identification of the Three Sibling Species of *Anopheles farauti* Laveran by the Banding Pattern of Their Polytene Chromosomes." *Journal of the Australian Entomological Society* 22 (1): 31–34. <https://doi.org/10.1111/j.1440-6055.1983.tb01835.x>.
- Mahon, Rodney, and P Miethke. 1982. "Anopheles Farauti No. 3, a Hitherto Unrecognized Biological Species of Mosquito within the Taxon *A. farauti* Laveran (Diptera: Culicidae)." *Transactions of the Royal Society of Tropical Medicine and Hygiene* 76 (1): 8–12.
- Manichaikul, Ani, Josyf Mychaleckyj, Stephen Rich, Kathy Daly, Michèle Sale, and Wei-Min Chen. 2010. "Robust Relationship Inference in Genome-Wide Association Studies." *Bioinformatics* 26 (22): 2867–73. <https://doi.org/10.1093/bioinformatics/btq559>.

- Mantel, Nathan. 1967. "The Detection of Disease Clustering and a Generalized Regression Approach." *Cancer Research* 27 (2): 209–20.
- Marsden, Clare, Anthony Cornel, Yoosook Lee, Michelle Sanford, Laura Norris, Parker Goodell, Catelyn Nieman, et al. 2013. "An Analysis of Two Island Groups as Potential Sites for Trials of Transgenic Mosquitoes for Malaria Control." *Evolutionary Applications* 6 (4): 706–20.  
<https://doi.org/10.1111/eva.12056>.
- Mathews, Christopher, Kensal van Holde, and Kevin Ahern. 2000. *Biochemistry*. 3rd ed. San Francisco: Benjamin Cummings.
- Max Planck Institute Bioinformatics Group. 2010. *Tools for Multiplex Sequencing on the Illumina Platform*. Python. <https://bioinf.eva.mpg.de/multiplex/>.
- Mboera, Leonard, and Willem Takken. 1997. "Carbon Dioxide Chemotropism in Mosquitoes (Diptera: Culicidae) and Its Potential in Vector Surveillance and Management Programmes." *Medical and Veterinary Entomology* 7: 355–68.
- Meyer Steiger, D. B., Scott Ritchie, and S. G. W. Laurance. 2014. "Overcoming the Challenges of Mosquito (Diptera: Culicidae) Sampling in Remote Localities: A Comparison of CO<sub>2</sub> Attractants on Mosquito Communities in Three Tropical Forest Habitats." *Journal of Medical Entomology* 51 (1): 39–45.
- Miller, Michael, Joseph Dunham, Angel Amores, William Cresko, and Eric Johnson. 2007. "Rapid and Cost-Effective Polymorphism Identification and Genotyping Using Restriction Site Associated DNA (RAD) Markers." *Genome Research* 17: 240–48. <https://doi.org/10.1101/gr.5681207>.
- Minucci, Angelo, Kamran Moradkhani, Ming Jing Hwang, Cecilia Zuppi, Bruno Giardina, and Ettore Capoluongo. 2012. "Glucose-6-Phosphate (G6PD\_ Mutations Database: Review of the 'Old' and Update of the New Mutations." *Blood Cells, Molecules, and Diseases* 48 (3): 154–65.  
<https://doi.org/10.1016/j.bcmd.2012.01.001>.
- Mukiama, T. K., and R. W. Mwangi. 1989. "Seasonal Population Changes and Malaria Transmission Potential of *Anopheles Pharoensis* and the Minor Anophelines in Mwea Irrigation Scheme, Kenya." *Acta Tropica* 46 (3): 181–89.  
[https://doi.org/10.1016/0001-706X\(89\)90035-1](https://doi.org/10.1016/0001-706X(89)90035-1).
- Müller, Pie, Emma Warr, Bradley Stevenson, Patricia Pignatelli, John Morgan, Andrew Steven, Alexander Yawson, et al. 2008. "Field-Caught Permethrin-Resistant *Anopheles gambiae* Overexpress *CYP6P3*, a P450 That Metabolises Pyrethroids." *PLoS Genetics* 4 (11): e1000286.  
<https://doi.org/10.1371/journal.pgen.1000286>.

- Mweresa, Collins, Philemon Omusula, Bruno Otieno, Joop van Loon, Willem Takken, and Wolfgang Mukabana. 2014. "Molasses as a Source of Carbon Dioxide for Attracting the Malaria Mosquitoes *Anopheles gambiae* and *Anopheles funestus*." *Malaria Journal* 13 (April): 160. <https://doi.org/10.1186/1475-2875-13-160>.
- Nabarro, David, and Elizabeth Tayler. 1998. "The 'Roll Back Malaria' Campaign." *Science* 280 (5372): 2067–68. <https://doi.org/10.1126/science.280.5372.2067>.
- National Center for Biotechnology Information. 2013. "Illumina Whole Genome Shotgun Sequencing of Genomic DNA Paired-End Library 'FARA1\_L' Containing Sample 'FARA1-L' from Individual 'MR4\_An. Farauti\_3' (SRR980932)." NCBI-Sequence Read Archive. 2013. [https://www.ncbi.nlm.nih.gov/sra/SRX349764\[accn\]](https://www.ncbi.nlm.nih.gov/sra/SRX349764[accn]).
- National Environment Management Authority (NEMA). 2005. "Kalangala District State of Environment Report, 2005." Kampala, Uganda: National Environment Management Authority (NEMA). [http://www.nemaug.org/district\\_reports/Kalangala\\_DSOER\\_2004.pdf](http://www.nemaug.org/district_reports/Kalangala_DSOER_2004.pdf).
- National Geodetic Survey. 2018. "NGS Coordinate Conversion and Transformation Tool (NCAT)." 2018. <https://www.ngs.noaa.gov/NCAT/>.
- National Malaria Control Program. 2018. "On the Road to a Malaria-Free Uganda - Second Universal Coverage Mosquito Net Distribution Campaign Offers Hope to Uganda." Newsletter. Uganda: Ministry of Health. <http://health.go.ug/sites/default/files/LLIN%20News%20letter%20final.pdf>.
- National Vector Borne Disease Control Programme. 2013. "Malaria Indicator Survey 2011." Honiara, Solomon Islands: Ministry of Health and Medical Services. [http://www.atoifiresearch.org.sb/sites/www.atoifiresearch.org.sb/files/uploaded/2011.4.SOL\\_1.MVP\\_FinalReport.1.Mar102014-8\\_32am-1.pdf](http://www.atoifiresearch.org.sb/sites/www.atoifiresearch.org.sb/files/uploaded/2011.4.SOL_1.MVP_FinalReport.1.Mar102014-8_32am-1.pdf).
- . 2016. "Annual Malaria Program Report (2016)." Honiara, Solomon Islands: Ministry of Health and Medical Services.
- Neafsey, Daniel, George Christophides, Frank Collins, Scott Emrich, Michael Fontaine, William Gelbart, Matthew Hahn, et al. 2013. "The Evolution of the *Anopheles* 16 Genomes Project." *G3: Genes, Genomes, Genetics* 3 (7): 1191–94. <https://doi.org/10.1534/g3.113.006247>.
- Neafsey, Daniel, Robert Waterhouse, Frank Collins, Scott Emrich, Michael Fontaine, William Gelbart, Matthew Hahn, Daniel Lawson, Igor Sharakov, and Nora Besansky. 2015. "Highly Evolvable Malaria Vectors: The Genomes of 16

- Anopheles* Mosquitoes." *Science* 347 (6217): 1258522.  
<https://doi.org/10.1126/science1258522>.
- Notre Dame Bioinformatics Lab. 2014. *Trimmer*. Python.  
<https://bitbucket.org/NDBL/hot-rad>.
- O'Loughlin, Samantha, Stephen Magesa, Charles Mbogo, Frank Mosha, Janet Midega, Susan Lomas, and Austin Burt. 2014. "Genomic Analyses of Three Malaria Vectors Reveals Extensive Shared Polymorphism but Contrasting Population Histories." *Molecular Biology and Evolution* 31 (4): 889–902.
- Oxborough, Richard. 2016. "Trends in US President's Malaria Initiative-Funded Indoor Residual Spray Coverage and Insecticide Choice in Sub-Saharan Africa (2008–2015): Urgent Need for Affordable, Long-Lasting Insecticides." *Malaria Journal* 15 (1): 146. <https://doi.org/10.1186/s12936-016-1201-1>.
- Paik, Yang, and James Avery. 1973. "Problem Areas in the Malaria Eradication Programme in the British Solomon Islands." *Papua New Guinea Medical Journal* 17 (1): 61–67.
- Parchman, Thomas, Zachariah Gompert, Joann Mudge, Faye Schilkey, Craig Benkman, and C. Alexander Buerkle. 2012. "Genome-Wide Association Genetics of an Adaptive Trait in Lodgepole Pine." *Molecular Ecology* 21 (12): 2991–3005. <https://doi.org/10.1111/j.1365-294X.2012.05513.x>.
- Paskewitz, Susan, Dawn Wesson, and Frank Collins. 1994. "The Internal Transcribed Spacers of Ribosomal DNA in Five Members of the *Anopheles gambiae* Species Complex." *Insect Molecular Biology* 2 (4): 247–57.  
<https://doi.org/10.1111/j.1365-2583.1994.tb00144.x>.
- Pritchard, Jonathan, Matthew Stephens, and Peter Donnelly. 2000. "Inference of Population Structure Using Multilocus Genotype Data." *Genetics* 155 (2): 945–59.
- R Core Team. 2014. *R: A Language and Environment for Statistical Computing* (version 3.2.1). R. Vienna, Austria: R Foundation for Statistical Computing.  
<http://www.R-project.org/>.
- Ragland, Gregory. 2017. *SampToSam.pl*. Perl.  
<https://github.com/gjragland/perlScripts/blob/master/sampToSam.pl>.
- Ranson, Hilary, B Jensen, John Vulule, X Wang, Janet Hemingway, and Frank Collins. 2000. "Identification of a Point Mutation in the Voltage-gated Sodium Channel Gene of Kenyan *Anopheles gambiae* Associated with Resistance to DDT and Pyrethroids." *Insect Molecular Biology* 9 (5): 491–97.  
<https://doi.org/10.1046/j.1365-2583.2000.00209.x>.

- Ranson, Hilary, and Natalie Lissenden. 2016. "Insecticide Resistance in African *Anopheles* Mosquitoes: A Worsening Situation That Needs Urgent Action to Maintain Malaria Control." *Trends in Parasitology* 32 (3): 187–96. <https://doi.org/10.1016/j.pt.2015.11.010>.
- Reeves, W. C. 1951. "Field Studies on Carbon Dioxide as a Possible Host Stimulant to Mosquitoes." *Experimental Biology and Medicine* 77 (1): 64–66.
- Reiff, Dana, Akira Kaneko, George Taleo, Motunrayo Amos, and J. Koji Lum. 2007. "Population Structure and Gene Flow of *Anopheles farauti* s.s. (Diptera: Culicidae) among Ten Sites on Five Islands of Vanuatu: Implications for Malaria Control." *Journal of Medical Entomology* 44 (4): 601–7. [https://doi.org/10.1603/0022-2585\(2007\)44\[601:PSAGFO\]2.0.CO;2](https://doi.org/10.1603/0022-2585(2007)44[601:PSAGFO]2.0.CO;2).
- Roll Back Malaria Partnership. 2008. "The Global Malaria Action Plan for a Malaria-Free World." Geneva, Switzerland: The Roll Back Malaria Partnership. <http://www.unhcr.org/4afac5629.pdf>.
- . 2010. "Malaria Funding and Resource Utilization: The First Decade of Roll Back Malaria." 1. Progress & Impact Series. Geneva, Switzerland: World Health Organization. [https://apps.who.int/iris/bitstream/handle/10665/44281/9789241599160\\_eng.pdf?sequence=1&isAllowed=y](https://apps.who.int/iris/bitstream/handle/10665/44281/9789241599160_eng.pdf?sequence=1&isAllowed=y).
- Rousset, François. 1997. "Genetic Differentiation and Estimation of Gene Flow from F Statistics Under Isolation by Distance." *Genetics* 145 (4): 1219–28.
- Rozeboom, Lloyd, and Kenneth Knight. 1946. "The Punctulatus Complex of *Anopheles* (Diptera: Culicidae)." *Journal of Parasitology* 32 (2): 95–131.
- RTS,S Clinical Trials Partnership. 2015. "Efficacy and Safety of RTS,S/AS01 Malaria Vaccine with or without a Booster Dose in Infants and Children in Africa: Final Results of a Phase 3, Individually Randomised, Controlled Trial." *The Lancet* 386 (9988): 31–45. [https://doi.org/10.1016/S0140-6736\(15\)60721-8](https://doi.org/10.1016/S0140-6736(15)60721-8).
- Rural Development Division. 2001. *Western Province Development Profile (August 2001)*. Honiara, Solomon Islands: Ministry of Provincial Government and Rural Development.
- Russell, Paul. 1955. *Man's Mastery of Malaria*. London: Oxford University Press.
- Russell, Tanya, Nigel Beebe, Hugo Bugoro, Allan Apairamo, Robert Cooper, Frank Collins, Neil Lobo, and Thomas Burkot. 2016. "Determinants of Host Feeding Success by *Anopheles farauti*." *Malaria Journal* 15 (March): 152. <https://doi.org/10.1186/s12936-016-1168-y>.

- Russell, Tanya, Thomas Burkot, Hugo Bugoro, Allan Apairamo, Nigel Beebe, Weng Chow, Robert Cooper, Frank Collins, and Neil Lobo. 2016a. "Data Archive: Larval Habitats of the *Anopheles farauti* and *Anopheles lungae* Complexes in the Solomon Islands." Data files. James Cook University Tropical Data Hub. 2016. 10.4225/28/56C6511C9B57D.
- . 2016b. "Larval Habitats of the *Anopheles farauti* and *Anopheles lungae* Complexes in the Solomon Islands." *Malaria Journal* 15 (March): 164. <https://doi.org/10.1186/s12936-016-1196-7>.
- Saitoh, Yasuhide, Junko Hattori, Shiro Chinone, Naoko Nihel, Yoshio Tsuda, Hiromu Kurahashi, and Mutsuo Kobayashi. 2004. "Yeast-Generated CO<sub>2</sub> as a Convenient Source of Carbon Dioxide for Adult Mosquito Sampling." *Journal of the American Mosquito Control Association* 20 (3): 261–64.
- Scott, Julie, William Brogdon, and Frank Collins. 1993. "Identification of Single Specimens of the *Anopheles gambiae* Complex by the Polymerase Chain Reaction." *American Journal of Tropical Medicine and Hygiene* 49 (4): 520–29. <https://doi.org/10.4269/ajtmh.1993.49.520>.
- Sharakhova, Maria, Martin Hammond, Neil Lobo, Jaroslaw Krzywinski, Maria Unger, Maureen Hillenmeyer, Robert Bruggner, Ewan Birney, and Frank Collins. 2007. "Update of the *Anopheles gambiae* PEST Genome Assembly." *Genome Biology* 8 (January): R5. <https://doi.org/10.1186/gb-2007-8-1-r5>.
- Shaukat, A. M., J. G. Breman, and F. G. McKenzie. 2010. "Using the Entomological Inoculation Rate to Assess the Impact of Vector Control on Malaria Parasite Transmission and Elimination." *Malaria Journal* 9 (May): 122. <https://doi.org/10.1186/1475-2875-9-122>.
- Simard, Frédéric, Didier Fontenille, Tovi Lehmann, Romain Girod, Laurent Brutus, Raheem Gopaul, Christian Dournon, and Frank Collins. 1999. "High Amounts of Genetic Differentiation between Populations of the Malaria Vector *Anopheles arabiensis* from West Africa and Eastern Outer Islands." *American Journal of Tropical Medicine and Hygiene* 60 (6): 1000–1009. <https://doi.org/10.4269/ajtmh.1999.60.1000>.
- Sinka, Marianne, Michael Bangs, Sylvie Manguin, Theeraphap Chareonviriyaphap, Anand Patil, William Temperley, Peter Gething, et al. 2011. "The Dominant *Anopheles* Vectors of Human Malaria in the Asia-Pacific Region: Occurrence Data, Distribution Maps and Bionomic Précis." *Parasites & Vectors* 4 (May): 89. <https://doi.org/10.1186/1756-3305-4-89>.
- Sinka, Marianne, Michael Bangs, Sylvie Manguin, Yasmin Rubio-Palis, Theeraphap Chareonviriyaphap, Maureen Coetzee, Charles Mbogo, et al. 2012. "A Global

- Map of Dominant Malaria Vectors." *Parasites & Vectors* 5 (1): 69.  
<https://doi.org/10.1186/1756-3305-5-69>.
- Smallgange, Renate, Wolfgang Schmied, Karel van Roey, Niels Verhulst, Jeroen Spitzen, Wolfgang Mukabana, and Willem Takken. 2010. "Sugar-Fermenting Yeast as an Organic Source of Carbon Dioxide to Attract the Malaria Mosquito *Anopheles gambiae*." *Malaria Journal* 9 (October): 292.  
<https://doi.org/10.1186/1475-2875-9-292>.
- Smithsonian Institution. 2018. "Global Volcanism Program." 2018.  
<http://volcano.si.edu/volcano.cfm?vn=255805>.
- Solomon Islands National Malaria Control Program. 2007. "National Malaria Strategic Vision 2007-2016 Solomon Islands."  
<http://apmen.org/apmen/Country%20Partners/Solomon%20Islands/Solomon%20Islands%20-%20National%20Vision%20Statement.pdf>.
- Solomon Islands National Statistical Office. 2009. "2009 Population & Housing Census National Report." Honiara, Solomon Islands.  
<http://www.statistics.gov.sb/statistics/demographic-statistics/census>.
- Ssegawa, Paul, and David Nkuutu. 2006. "Diversity of Vascular Plants on Ssesse Islands in Lake Victoria, Central Uganda." *African Journal of Ecology* 44 (1): 22–29. <https://doi.org/10.1111/j.1365-2028.2006.00609.x>.
- Syed, Zainulabeuddin, and Walter Leal. 2007. "Maxillary Palps Are Broad Spectrum Odorant Detectors in *Culex quinquefasciatus*." *Chemical Senses* 32 (8): 727–38. <https://doi.org/10.1093/chemse/bjm040>.
- Taylor, Brian. 1975a. "Observations on Malaria Vectors of the *Anopheles Punctulatus* Complex in the British Solomon Islands Protectorate." *Journal of Medical Entomology* 11 (6): 677–87. <https://doi.org/10.1093/jmedent/11.6.677>.
- . 1975b. "Changes in the Feeding Behaviour of a Malaria Vector, *Anopheles Farauti* Lav., Following Use of DDT as a Residual Spray in Houses in the British Solomon Islands Protectorate." *Ecological Entomology* 127 (3): 277–92. <https://doi.org/10.1111/j.1365-2311.1975.tb00576.x>.
- The *Anopheles gambiae* 1000 Genomes Consortium. 2017. "Natural Diversity of the Malaria Vector *Anopheles gambiae*." *Nature* 552 (7683): 96–100.  
<https://doi.org/10.1038/nature24995>.
- The Global Fund. 2018. "List of Indoor Residual Sprays (IRS) That Meet Global Fund Quality Assurance Requirements for Use against Malaria Vectors."  
[https://www.theglobalfund.org/media/5857/psm\\_indoorresidualsprayirsgf\\_list\\_en.pdf?u=636651116840000000](https://www.theglobalfund.org/media/5857/psm_indoorresidualsprayirsgf_list_en.pdf?u=636651116840000000).

- The Malaria Atlas Project. 2019. *The Malaria Atlas Project*. English. University of Oxford. <https://map.ox.ac.uk>.
- The malERA Consultative Group on Drugs. 2011. "A Research Agenda for Malaria Eradication: Drugs." *PLoS Medicine* 8 (1): e1000402. <https://doi.org/10.1371/journal.pmed.1000402>.
- The malERA Consultative Group on Vaccines. 2011. "A Research Agenda for Malaria Eradication: Vaccines." *PLoS Medicine* 8 (1): e1000398. <https://doi.org/10.1371/journal.pmed.1000398>.
- The malERA Consultative Group on Vector Control. 2011. "A Research Agenda for Malaria Eradication: Vector Control." *PLoS Medicine* 8 (1): e1000401. <https://doi.org/10.1371/journal.pmed.1000401>.
- The PATH Malaria Vaccine Initiative. 2016. "Fact Sheet: RTS,S Malaria Vaccine Candidate (Mosquirix™)." [https://www.malariavaccine.org/files/content/page/files/RTSS%20vaccine%20candidate%20Factsheet\\_FINAL.pdf](https://www.malariavaccine.org/files/content/page/files/RTSS%20vaccine%20candidate%20Factsheet_FINAL.pdf).
- Thomas, A. S. 1941. "The Vegetation of the Sese Islands. Uganda: An Illustration of Edaphic Factors in Tropical Ecology." *Journal of Ecology* 29 (2): 330–53. <https://doi.org/10.2307/2256396>.
- Touré, Yeya, Guimogo Dolo, Vincenzo Petrarca, S. F. Traoré, M. Bouaré, Adama Dao, John Carnahan, and Charles Taylor. 1998. "Mark-Release-Recapture Experiments with *Anopheles gambiae* s. l. in Banambani Village, Mali, to Determine Population Size and Structure." *Medical and Veterinary Entomology* 12 (1): 74–83. <https://doi.org/10.1046/j.1365-2915.1998.00071.x>.
- Touré, Yeya, Ayoade Oduola, and Carlos Morel. 2004. "The *Anopheles gambiae* Genome: Next Steps for Malaria Vector Control." *Trends in Parasitology* 20 (3): 142–49. <https://doi.org/10.1016/j.pt.2004.01.008>.
- Trampuz, Andrej, Matjaz Jereb, Igor Muzlovic, and Rajesh Prabhu. 2003. "Clinical Review: Severe Malaria." *Critical Care* 7 (April): 315. <https://doi.org/10.1186/cc2183>.
- Uganda Bureau of Statistics. 2016. "The National Population and Housing Census 2014 - Main Report." Kampala, Uganda: Uganda Bureau of Statistics. <http://www.ubos.org/onlinefiles/uploads/ubos/NPHC/2014%20National%20Census%20Main%20Report.pdf>.



- Uganda Ministry of Health. 2016. "Uganda Malaria Quarterly Bulletin Issue 16: October-December 2016." Quarterly Bulletin 16. Kampala, Uganda: Uganda Ministry of Health. <http://uphfp.org/?mdocs-file=2733>.
- Van der Auwera, Geraldine, Mauricio Carneiro, Christopher Hartl, Ryan Poplin, Guillermo del Angel, Ami Levy-Moonshine, Tadeusz Jordan, et al. 2013. "From FastQ Data to High-Confidence Variant Calls: The Genome Analysis Toolkit Best Practices Pipeline." In *Current Protocols in Bioinformatics*, 11.10.1-33. John Wiley & Sons, Inc. <http://onlinelibrary.wiley.com/doi/10.1002/0471250953.bi1110s43/full>.
- VectorBase. 2013. *Anopheles-Farauti-FAR1\_SCAFFOLDS\_AfarF1.fa.gz*. English. <https://www.vectorbase.org/downloadinfo/anopheles-farauti-far1scaffoldsafarf1fagz>.
- . 2014. *Anopheles-Gambiae-PEST\_AgamP4\_agp.gz*. English. <https://www.vectorbase.org/downloadinfo/anopheles-gambiae-pestchromosomesagamp4fagz>.
- . 2017. "VectorBase: Bioinformatics Resource for Invertebrate Vectors of Human Pathogens." 2017. <http://biomart.vectorbase.org/biomart/martview/39809899a4a2f01f48761e0ae1fce499>.
- Vincent, Dayton. 1994. "The South Pacific Convergence Zone (SPCZ): A Review." *Monthly Weather Review* 122 (September): 1949–70. [https://doi.org/10.1175/1520-0493\(1994\)122<1949:TSPCZA>2.0.CO;2](https://doi.org/10.1175/1520-0493(1994)122<1949:TSPCZA>2.0.CO;2).
- Visit Solomons. 2018. "How to Get Around." Visit Solomons. 2018. <http://www.visitsolomons.com.sb/plan/how-to-get-around>.
- Vythilingam, Indra, Yusuf NoorAzian, Tan Cheong Huat, Adela Ida Jiram, Yusof Yusri, Abdul Azahari, Ismail NorParina, Abdullah NoorRain, and Sulaiman LokmanHakim. 2008. "*Plasmodium knowlesi* in Humans, Macaques and Mosquitoes in Peninsular Malaysia." *Parasites & Vectors* 1 (1): 26. <https://doi.org/10.1186/1756-3305-1-26>.
- Vythilingam, Indra, C. H. Tan, M. Asmad, S. T. Chan, K. S. Lee, and Balbir Singh. 2006. "Natural Transmission of *Plasmodium knowlesi* to Humans by *Anopheles latens* in Sarawak, Malaysia." *Transactions of the Royal Society of Tropical Medicine and Hygiene* 100 (11): 1087–88. <https://doi.org/10.1016/j.trstmh.2006.02.006>.
- Wairiu, Morgan, and Rattan Lal. 2003. "Soil Organic Carbon in Relation to Cultivation and Topsoil Removal on Sloping Lands of Kolombangara, Solomon Islands."

- Soil & Tillage Research* 70 (1): 19–27. [https://doi.org/10.1016/S0167-1987\(02\)00116-2](https://doi.org/10.1016/S0167-1987(02)00116-2).
- Wallace, Alfred. 1863. “On the Physical Geography of the Malay Archipelago.” *Journal of the Royal Geographical Society of London* 33: 217–34. <https://doi.org/10.2307/1798448>.
- Wang, Jingliang. 2016. “A Comparison of Single-Sample Estimators of Effective Population Sizes from Genetic Marker Data.” *Molecular Ecology* 25 (19): 4692–4711. <https://doi.org/10.1111/mec.13725>.
- Wang, Sibao, and Marcelo Jacobs-Lorena. 2013. “Genetic Approaches to Interfere with Malaria Transmission by Vector Mosquitoes.” *Trends in Biotechnology* 31 (3): 185–93. <https://doi.org/10.1016/j.tibtech.2013.01.001>.
- Waples, Robin, and Chi Do. 2008. “LDNE: A Program for Estimating Effective Population Size from Data on Linkage Disequilibrium.” *Molecular Ecology* 8 (4): 753–56. <https://doi.org/10.1111/j.1755-0998.2007.02061.x>.
- Weber, Max. 1902. *Der indo-australische Archipel und die Geschichte seiner Tierwelt*. Jena: G. Fischer.
- Weir, Bruce, and C. Clark Cockerham. 1984. “F-Statistics for the Analysis of Population Structure.” *Evolution* 38 (6): 1358–70. <https://doi.org/10.2307/2408641>.
- White, Bradley, Frank Collins, and Nora Besansky. 2011. “Evolution of *Anopheles gambiae* in Relation to Humans and Malaria.” *Annual Review of Ecology, Evolution, and Systematics* 42 (December): 111–32. <https://doi.org/10.1146/annurev-ecolsys-102710-145028>.
- White, Bradley, Federica Santolamazza, Luna Kamau, Marco Pombi, Olga Grushko, Karine Mouline, Cecile Brengues, et al. 2007. “Molecular Karyotyping of the 2LA Inversion in *Anopheles gambiae*.” *American Journal of Tropical Medicine and Hygiene* 76 (2): 334–39. <https://doi.org/10.4269/ajtmh.2007.76.334>.
- White, Nicholas. 2009. “Chapter 73. Malaria.” In *Manson’s Tropical Diseases*, edited by Gordon Cook and Alimuddin Zumla, 22nd ed., 1201–1300. Edinburgh: Saunders.
- WHO Working Group. 1989. “Glucose-6-Phosphate Dehydrogenase Deficiency.” *Bulletin of the World Health Organization* 67 (6): 601–11.
- Wiltshire, Rachel. 2019. *SharedVariantsAllSamples.py*. Python. University of Notre Dame. <https://github.com/rachelwiltshire/Python-scripts/blob/master/sharedVariantsAllSamples.py>.

- Wiltshire, Rachel, Christina Bergey, Jonathan Kayondo, Josephine Birungi, Louis Mukwaya, Scott Emrich, Nora Besansky, and Frank Collins. 2018. "Reduced-Representation Sequencing Identifies Small Effective Population Sizes of *Anopheles gambiae* in the North-Western Lake Victoria Basin, Uganda." *Malaria Journal* 17 (August): 285. <https://doi.org/10.1186/s12936-018-2432-0>.
- World Health Organization. 1993. "Implementation of the Global Malaria Control Strategy." 839. WHO Technical Report Series. Geneva, Switzerland: World Health Organization. [https://apps.who.int/iris/bitstream/handle/10665/37106/WHO\\_TRS\\_839\\_eng.pdf](https://apps.who.int/iris/bitstream/handle/10665/37106/WHO_TRS_839_eng.pdf).
- . 2010. "World Malaria Report 2010." Geneva, Switzerland: World Health Organization. [http://www.who.int/malaria/world\\_malaria\\_report\\_2010/worldmalariareport2010.pdf?ua=1](http://www.who.int/malaria/world_malaria_report_2010/worldmalariareport2010.pdf?ua=1).
- . 2012. "Global Plan for Insecticide Resistance Management in Malaria Vectors." Geneva, Switzerland: World Health Organization. [http://apps.who.int/iris/bitstream/handle/10665/44846/9789241564472\\_eng.pdf?sequence=1](http://apps.who.int/iris/bitstream/handle/10665/44846/9789241564472_eng.pdf?sequence=1).
- . 2013. "Seasonal Malaria Chemoprevention with Sulfadoxine-Pyrimethamine plus Amodiaquine in Children: A Field Guide." Geneva, Switzerland: World Health Organization. [http://apps.who.int/iris/bitstream/handle/10665/85726/9789241504737\\_eng.pdf?sequence=1](http://apps.who.int/iris/bitstream/handle/10665/85726/9789241504737_eng.pdf?sequence=1).
- . 2014. "Guidance Framework for Testing of Genetically Modified Mosquitoes." Geneva, Switzerland: World Health Organization. [http://apps.who.int/iris/bitstream/handle/10665/127889/9789241507486\\_eng.pdf?sequence=1](http://apps.who.int/iris/bitstream/handle/10665/127889/9789241507486_eng.pdf?sequence=1).
- . 2015a. "Global Technical Strategy for Malaria 2016-2030." Geneva, Switzerland: World Health Organization. <https://www.who.int/malaria/publications/atoz/9789241564991/en/>.
- . 2015b. "Guidelines for the Treatment of Malaria." Geneva, Switzerland. [http://apps.who.int/iris/bitstream/handle/10665/162441/9789241549127\\_eng.pdf?sequence=1](http://apps.who.int/iris/bitstream/handle/10665/162441/9789241549127_eng.pdf?sequence=1).
- . 2015c. "World Malaria Report 2015." Geneva, Switzerland: World Health Organization. <http://www.who.int/malaria/publications/world-malaria-report-2015/report/en/>.

- . 2017a. "Data Requirements and Methods to Support the Evaluation of New Vector Control Products." Geneva, Switzerland: World Health Organization. <http://apps.who.int/iris/bitstream/handle/10665/259536/WHO-HTM-GMP-2017.14-eng.pdf?sequence=1>.
  - . 2017b. "Report of the Twentieth WHOPES Working Group Meeting, WHO/HQ, Geneva, 20-24 March 2017: Review of Interceptor G2LN, DawaPlus 3.0 LN, DawaPlus 4.0LN, SumiLarv 2 MR, Chlorfenapyr 240 SC." Geneva: World Health Organization. <http://apps.who.int/iris/bitstream/handle/10665/258921/WHO-HTM-NTD-WHOPES-2017.04-eng.pdf?sequence=1>.
  - . 2017c. "Malaria Prevention Works." Geneva, Switzerland: World Health Organization. <http://apps.who.int/iris/bitstream/handle/10665/254991/WHO-HTM-GMP-2017.6-eng.pdf;jsessionid=29F6B7F3A2DBD344998EF6DADA21109D?sequence=1>.
  - . 2017d. "Malaria: Intermittent Preventive Treatment in Infants (IPTi)." World Health Organization. May 1, 2017. [http://www.who.int/malaria/areas/preventive\\_therapies/infants/en/](http://www.who.int/malaria/areas/preventive_therapies/infants/en/).
  - . 2017e. "World Malaria Report 2017." Geneva, Switzerland: World Health Organization. <http://www.who.int/malaria/publications/world-malaria-report-2017/report/en/>.
  - . 2018a. "Prequalification Vector Control Products." World Health Organization. <http://www.who.int/pq-vector-control/prequalified-lists/Prequalifiedproducts8June2018.pdf?ua=1>.
  - . 2018b. "Malaria: Intermittent Preventive Treatment in Pregnancy (IPTp)." World Health Organization. June 21, 2018. [http://www.who.int/malaria/areas/preventive\\_therapies/pregnancy/en/](http://www.who.int/malaria/areas/preventive_therapies/pregnancy/en/).
  - . 2018c. "Q&A on the Malaria Vaccine Implementation Programme (MVIP)." World Health Organization. <http://www.who.int/malaria/media/malaria-vaccine-implementation-qa/en/>.
  - . 2018d. "World Malaria Report 2018." Geneva, Switzerland: World Health Organization. <https://www.who.int/malaria/publications/world-malaria-report-2018/en/>.
- Wright, Sewall. 1943. "Isolation By Distance." *Genetics* 28 (2): 114–38.

- . 1978. *Evolution and the Genetics of Populations: A Treatise in Four Volumes*. Vol. 4: Variability within and among natural populations. 4 vols. University of Chicago Press.
- Yang, Jian, Beben Benyamin, Brian McEvoy, Scott Gordon, Anjali Henders, Dale Nyholt, Pamela Madden, et al. 2010. "Common SNPs Explain a Large Proportion of the Heritability for Human Height." *Nature Genetics* 42 (7): 565–69. <https://doi.org/10.1038/ng.608>.
- Yaya, Sanni, Olalekan Uthman, Agbessi Amouzou, and Ghose Bishwajit. 2018. "Use of Intermittent Preventive Treatment among Pregnant Women in Sub-Saharan Africa: Evidence from Malaria Indicator Surveys." *Tropical Medicine and Infectious Disease* 3 (1): 18. <https://doi.org/10.3390/tropicalmed3010018>.