University of Kentucky
UKnowledge

Mathematics Faculty Publications | Mathematics

5-22-2017

Simulating Within-Vector Generation of the Malaria Parasite Diversity

Lauren M. Childs
Virginia Tech

Olivia F. Prosper
University of Kentucky, olivia.prosper@uky.edu

Follow this and additional works at: https://uknowledge.uky.edu/math_facpub

 Part of the Disease Modeling Commons, Mathematics Commons, and the Parasitic Diseases Commons

Right click to open a feedback form in a new tab to let us know how this document benefits you.

Repository Citation

Childs, Lauren M. and Prosper, Olivia F., "Simulating Within-Vector Generation of the Malaria Parasite Diversity" (2017). *Mathematics Faculty Publications*. 22.
https://uknowledge.uky.edu/math_facpub/22

This Article is brought to you for free and open access by the Mathematics at UKnowledge. It has been accepted for inclusion in Mathematics Faculty Publications by an authorized administrator of UKnowledge. For more information, please contact UKnowledge@lsv.uky.edu.

Simulating Within-Vector Generation of the Malaria Parasite Diversity

Digital Object Identifier (DOI)
https://doi.org/10.1371/journal.pone.0177941

Notes/Citation Information
Published in *PLOS ONE*, v. 12, 5, e0177941, p. 1-23.

© 2017 Childs, Prosper.

This is an open access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

This article is available at UKnowledge: https://uknowledge.uky.edu/math_facpub/22

PLOS ONE

RESEARCH ARTICLE

Simulating within-vector generation of the malaria parasite diversity

Lauren M. Childs[1☯]*, **Olivia F. Prosper**[2☯]

1 Department of Mathematics, Virginia Tech, Blacksburg, VA, United States of America, **2** Department of Mathematics, University of Kentucky, Lexington, KY, United States of America

☯ These authors contributed equally to this work.
* lchilds@vt.edu

Check for updates

OPEN ACCESS

Citation: Childs LM, Prosper OF (2017) Simulating within-vector generation of the malaria parasite diversity. PLoS ONE 12(5): e0177941. https://doi.org/10.1371/journal.pone.0177941

Editor: Gordon Langsley, Institut national de la santé et de la recherche médicale - Institut Cochin, FRANCE

Received: December 7, 2016

Accepted: May 5, 2017

Published: May 22, 2017

Copyright: © 2017 Childs, Prosper. This is an open access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Data Availability Statement: All relevant data are within the paper and its Supporting Information files.

Funding: L.M.C. was partially supported by award number U54GM088558 from the National Institute of General Medical Sciences. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institute of General Medical Sciences or the National Institutes of Health. The authors also acknowledge the support of Virginia Tech's Open Access Subvention Fun. The funders had no role in

Abstract

Plasmodium falciparum, the most virulent human malaria parasite, undergoes asexual reproduction within the human host, but reproduces sexually within its vector host, the *Anopheles* mosquito. Consequently, the mosquito stage of the parasite life cycle provides an opportunity to create genetically novel parasites in multiply-infected mosquitoes, potentially increasing parasite population diversity. Despite the important implications for disease transmission and malaria control, a quantitative mapping of how parasite diversity entering a mosquito relates to diversity of the parasite exiting, has not been undertaken. To examine the role that vector biology plays in modulating parasite diversity, we develop a two-part model framework that estimates the diversity as a consequence of different bottlenecks and expansion events occurring during the vector-stage of the parasite life cycle. For the underlying framework, we develop the first stochastic model of within-vector *P. falciparum* parasite dynamics and go on to simulate the dynamics of two parasite subpopulations, emulating multiply infected mosquitoes. We show that incorporating stochasticity is essential to capture the extensive variation in parasite dynamics, particularly in the presence of multiple parasites. In particular, unlike deterministic models, which always predict the most fit parasites to produce the most sporozoites, we find that occasionally only parasites with lower fitness survive to the sporozoite stage. This has important implications for onward transmission. The second part of our framework includes a model of sequence diversity generation resulting from recombination and reassortment between parasites within a mosquito. Our two-part model framework shows that bottlenecks entering the oocyst stage decrease parasite diversity from what is present in the initial gametocyte population in a mosquito's blood meal. However, diversity increases with the possibility for recombination and proliferation in the formation of sporozoites. Furthermore, when we begin with two parasite subpopulations in the initial gametocyte population, the probability of transmitting more than two unique parasites from mosquito to human is over 50% for a wide range of initial gametocyte densities.

study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing interests: The authors have declared that no competing interests exist.

Introduction

Each year nearly 200 million people are infected with malaria causing parasites, resulting in over half a million deaths, mostly children under the age of five in sub-Saharan Africa [1]. The most virulent species, *Plasmodium falciparum*, exhibits significant diversity in the circulating parasite population [2] even in low endemicity settings [3]. The high level of diversity exhibited across the *P. falciparum* parasite population is unexpected due to the severe bottlenecks during transition between the vertebrate host and the mosquito vector. Furthermore, *P. falciparum* only has capacity to increase diversity via mutation throughout most of its life cycle due to its haploid state, particularly during the longest stage of parasite development, i.e. replication in human red blood cells. Only for a brief period in the mosquito, upon mating, do the parasites form a diploid state [4]. During this period, if multiple parasite populations are present, novel parasites can arise through the process of meiosis.

Mathematical modeling has been a critical component of malaria research. The majority of models have focused on transmission of parasites (e.g. [5–8]) or dynamics within the human host, specifically in the human red blood cells (e.g. [9–14]). The only works focused on parasite dynamics in the mosquito vector consider a genetically identical malaria parasite population and used deterministic transitions between stages in the life cycle [15, 16]. The progression of the life cycle within the mosquito, however, is much more varied with a bottleneck at the oocyst stage [17]. Parasites enter the mosquito through an infectious blood meal, consisting of a few μL of blood [18] which typically contains less than 100 gametocytes (the sexual stage of the parasite able to survive within the mosquito), although up to several thousand have been observed [19–22]. Once gametocytes enter the midgut of the mosquito and experience the change in environment such as a drop in temperature and rise in pH, the parasites begin gametogenesis, the transformation into gametes. In a matter of minutes, the male gametes divide three times to form up to eight motile microgametes while the female gametes mature into single stationary macrogametes. Microgametes search for macrogametes, and upon finding one, fuse to form a diploid zygote which eventually becomes a tetraploid ookinete [23]. The ookinetes migrate through the wall of the midgut and establish themselves as oocysts. Prior to settling as oocysts, they undergo meiosis [24, 25]. As oocysts mature, the parasites undergo 10—11 rounds of endomitosis [26] after which they burst releasing sporozoites into the surrounding extracellular space. The sporozoites must migrate to the salivary glands before they can be transferred to new hosts during an infectious bite. When an infectious mosquito bites a vertebrate host, typically 10-100 sporozoites are transferred into the skin of the new host [27].

A key complication in the understanding of population-level parasite diversity is multiply infected individuals. In fact, many malaria infections of humans, across various transmission settings, appear to contain multiple parasites [28–31]. Quantifying the diversity of parasites within individual infections is challenging as there is a lack of sensitive genetic markers to easily distinguish genetically similar parasites [32, 33]. Multiple parasites can be co-transmitted to a mosquito in a single infectious bite allowing for crossing of parasites in the mosquito midgut [33, 34]. Mixed infections, those harboring multiple distinct parasites, have been suggested to be more infectious [35, 36], despite observations to the contrary [37]. Although it has been generally accepted that recombination within the mosquito increases parasite diversity, there has been little analysis of how diversity generation proceeds mechanistically and its implications for genetic diversity of the population as a whole [38, 39].

In this paper, we address this knowledge gap concerning the mechanistic generation of diversity in the mosquito. To do so, we develop a modeling framework tracking malaria parasite diversity through the various parasite life cycle stages in the mosquito host. As part of this

framework, we introduce the first stochastic model of the development of the malaria parasite within the vector. We show that incorporating variation in the dynamics of life stages of the malaria parasite, *Plasmodium falciparum*, within the mosquito leads to a range of sporozoite levels within the salivary glands, with important implications for the capacity for onward transmission. To complete our framework, we couple our stochastic model of parasite number in the mosquito to a stochastic model of sequence diversity generation including recombination between parasites within the mosquito. From our simulations, we find that sequence diversity decreases from entrance of gametocytes in a blood meal to formation of oocysts but the number of unique sequences present in the sporozoite population found in the salivary glands is greatly increased. Those novel sequences are predicted to be frequently passed onto new infections.

Models

We develop a two-part modeling framework tracking malaria parasite diversity in the mosquito. The first part of this framework concerns the parasite life cycle dynamics in the mosquito, which we implement in two different ways (deterministic and stochastic). In the second part of our modeling framework, we build a stochastic model of sequence diversity generation. This model includes reassortment and recombination of the parasite genome during meiosis. When multiple genetically distinct (defined by at least one difference in a sequence position) parasite populations, are in the same mosquito, new genetically novel parasites can be generated. This model tracks the diversity of parasites within a mosquito through the various life stages in the mosquito, requiring output from the first part of the modeling framework.

Underlying life cycle model

Deterministic implementation. Teboh-Ewungkem and co-authors developed the first within-vector models of *P. falciparum* parasite dynamics in two similar formulations published in 2010 [15, 16]. Here, we utilize a hybrid of the two model formulations, which will serve as the deterministic comparison for our stochastic model developed in the next section. We briefly review the modified version of the Teboh-Ewungkem model below.

The parasite dynamics begin with an initial number of male and female gametes, determined by (1) the initial number G_0 of gametocytes in a bloodmeal, (2) the fraction m of gametocytes that are male, (3) the probabilities α and β that male and female gametocytes successfully undergo gametogenesis to produce gametes, and (4) the average numbers ρ and v of gametes resulting from one male or female gametocyte, respectively. Thus, the initial male, M_0, and female, F_0, gamete numbers, are given by:

$$\begin{aligned} M_0 &= m\rho\alpha G_0 \\ F_0 &= (1-m)v\beta G_0 \end{aligned}$$

The male and female gametes decay due to natural mortality at rates a and b, respectively. Otherwise, fusion of female F and male M gametes occurs at a fertilization rate r to produce new zygotes Z. These zygotes either die at a rate μ_z, or transform into ookinetes E at a rate σ_z. Subsequently, ookinetes will either die at a rate μ_e, or transform into oocysts O at a rate σ_e. Once established, oocysts undergo sporoblast formation, a critical stage in the parasite life-cycle, eventually releasing thousands of sporozoites when they rupture. Continuing with the notation of Teboh-Ewungkem *et al.*, we denote the oocyst mortality rate, the oocyst rupture rate, and the average number of sporozoites produced per oocyst by μ_o, $k(t)$, and n_0 (denoted by n in [15, 16]), respectively. Because oocysts only rupture following sporoblast formation,

the resulting delay in sporozoite release is captured by a step function, as in [15, 16]:

$$k(t) = \begin{cases} 0, & \text{if } 0 < t < t_0 \\ d, & \text{if } t \geq t_0 \end{cases}$$

where t_0 is the time until mature oocysts have formed, and d is the rate at which mature oocysts release sporozoites. Finally, only a fraction p of sporozoites will successfully migrate to, and invade, the salivary glands producing S sporozoites found there. The full deterministic model is described by the following system of non-autonomous, ordinary differential equations [15, 16], and all parameters are found in Table 1.

$$\begin{aligned} M' &= -aM - rMF \\ F' &= -bF - rMF \\ Z' &= rMF - (\sigma_z + \mu_z)Z \\ E' &= \sigma_z Z - (\sigma_e + \mu_e)E \\ O' &= \sigma_e E - (k(t) + \mu_o)O \\ S' &= n_0 p k(t) O. \end{aligned}$$

Stochastic implementation. To better evaluate within-vector parasite variation, we developed a continuous time, discrete state Markov Chain (CTMC) extension of the Teboh-Ewungkem models (see S1 Fig).

In this stochastic model, we assumed there exists a time interval Δt small enough that at most one event in the parasite dynamics can occur within this window. These events include deaths from each stage of the within-vector parasite life-cycle and progressions from one stage

Table 1. Parameter values for the underlying within-vector model. All values, apart from G_0 and d, were obtained from Teboh-Ewungkem *et al.* [15]. G_0 was varied. d was obtained from Teboh-Ewungkem *et al.* [16].

Symbol	Description	Default value
G_0	number of gametocytes in blood meal	{150, 200, . . ., 450}
m	percentage of gametocytes that are male	0.25
v	number of female gametes per female gametocyte	1
ρ	number of male gametes per male gametocytes	8
α	fraction of male gametes that are viable	0.39
β	fraction of female gametes that are viable	1
a	failure rate of male gametocytes (per day)	24/(20/60)
b	failure rate of female gametocytes (per day)	24/(25/60)
r	fertilization of male and female gametes (per parasite per day)	0.08
μ_z	death rate of zygotes (per day)	1
μ_e	death rate of ookinetes (per day)	1.4
μ_o	death rate of oocysts (per day)	0
σ_z	transformation rate of zygotes (per day)	24/19
σ_e	transformation rate of ookinetes (per day)	0.6
n_0	number of sporozoites per oocyst	3000
p	proportion of sporozoites that make it to salivary gland	0.2
t_0	time to mature oocyst formation (days)	10
d	rate mature oocysts release sporozoites (per day)	$\frac{1}{7}$

https://doi.org/10.1371/journal.pone.0177941.t001

Table 2. CTMC transition probabilities for the stochastic within-vector model. *Note that probabilities are denoted in the standard form used by [40]. For calculation, the term as written is divided by the sum of all transitions. See SI File for details.

Event	Probabilities*	Transitions
Death of Male gamete	$aM\Delta t + o(\Delta t)$	(−1, 0, 0, 0, 0, 0)
Mating	$rMF\Delta t + o(\Delta t)$	(−1, −1, 1, 0, 0, 0)
Death of Female gamete	$bF\Delta t + o(\Delta t)$	(0, −1, 0, 0, 0, 0)
Death of Zygote	$\mu_z Z\Delta t + o(\Delta t)$	(0, 0, −1, 0, 0, 0)
Zygote to Ookinete progression	$\sigma_z Z\Delta t + o(\Delta t)$	(0, 0, −1, 1, 0, 0)
Death of Ookinete	$\mu_e E\Delta t + o(\Delta t)$	(0, 0, 0, −1, 0, 0)
Ookinete to Oocyst progression	$\sigma_e E\Delta t + o(\Delta t)$	(0, 0, 0, −1, 1, 0)
Death of Oocyst	$\mu_o O\Delta t + o(\Delta t)$	(0, 0, 0, 0, −1, 0)
Bursting of Oocyst (sporozoite production)	$k(t)O\Delta t + o(\Delta t)$	(0, 0, 0, 0, −1, n)

https://doi.org/10.1371/journal.pone.0177941.t002

to the next. The simulations were carried out using a small modification to the traditional Gillespie Algorithm, for which the inter-event times are exponentially distributed, with probabilities and transitions described in Table 2 and the same initial conditions as in the deterministic model (see S1 Appendix for an outline of the algorithm). The step-function formulation for the oocyst bursting rate $k(t)$, introduced by Teboh-Ewungkem *et al.* and stated above in the deterministic formulation of the model, assumes all oocysts reach maturity at the same time. To be more biologically accurate, we relax this assumption and introduce a continuous time function with a similar profile: $\beta(t) = (1 + \exp(t_0 - t))^{-1}$, which is small for small values of t, 0.5 when $t = t_0$, and approaches one as t approaches infinity. Recall that d is the rupture rate of mature oocysts. Thus, the bursting rate function in our stochastic model is $k(t) = d\beta(t)$.

To introduce variability in the number of sporozoites n released from one bursting oocyst, we first drew the number of sporozoites per oocyst from a Poisson distribution with mean $n = n_0$. Because only a fraction of sporozoites successfully migrate to the vector salivary glands, we then, independently, selected a portion of these sporozoites from a binomial distribution with probability of success p.

Multiple genetically distinct parasite populations. Humans are often infected with multiple genetically distinct *Plasmodium* parasite subpopulations [28–30], and consequently, a mosquito blood meal can be composed of multiple genetically distinct parasites. A fitness advantage of different parasites within the same mosquito host has been exemplified experimentally [29]. To this end, we extended the stochastic life cycle model to incorporate N genetically distinct parasites with different survival probabilities. We assume that parasites with greater survival probability have a higher fitness. For example, in the case where $N = 2$ (which we consider in depth in this paper), that is, where we modeled two genetically distinct parasite subpopulations within the mosquito, we assumed that parasite 1 is more fit than parasite 2, with the mortality rate in each stage of the life cycle (beginning with the gamete stage) enhanced by 10% (or 50%) for parasite 2, referred to as 10% (or 50%) bias throughout. Four types of zygotes may result from the two genetically distinct parasites: $Z_{1,1}$, $Z_{1,2}$, $Z_{2,1}$, $Z_{2,2}$, where $Z_{i,j}$ denotes a zygote forming from the fertilization of a type i female gamete by a type j male gamete. As we have no knowledge of how the fitness of the progeny of two parasites will be effected, we assume the simplest function, i.e. that the fitness of each progeny is the average of their parents; for example, the mortality rate of $Z_{1,2}$ is the average of the mortality rates for $Z_{1,1}$ and $Z_{2,2}$.

The model keeps track of the number of zygotes, ookinetes, oocysts, and sporozoites resulting from the four possible sets of "parent parasites". The transition matrix is defined in a

similar fashion for the two parasite case as for the single genetically distinct parasite model, only now there are two types of 'Death of a male gamete' events: the death of a parasite 1 gamete, and the death of a parasite 2 gamete. For two parasites, there are four possible types of 'fertilization' events, resulting in the four types of zygotes discussed earlier. If the event in a particular time-step is the fertilization of a parasite 1 female gamete by a parasite 2 male gamete, then the transition matrix is defined such that the number of parasite 1 female gametes, and the number of parasite 2 male gametes both *decrease* by 1, and the number of parasite 1,2 zygotes ($Z_{1,2}$) *increases* by one. We present the complete transition matrix for the two parasite case in S1 Table. All code for the life-cycle model can be found in S1 File.

Model of diversity generation

We modeled parasite diversity as it changes during the parasite life cycle in the mosquito (see S2 Fig). We represent each parasite by a sequence of length L with each position taken from an alphabet of length A. We calculated the diversity of the simulated parasite population as entering gametocytes, ruptured oocysts, sporozoites in the salivary glands, and exiting sporozoites during a single mosquito infection. All code for the diversity model can be found in S2 File.

Definition of parasite barcode sequence. The genetic basis of this study is supported by the segregation of 24 positions in the genome ($L = 24$) that will be considered as biallelic ($A = 2$), motivated by the 24-position SNP (Single Nucleotide Polymorphism) barcode developed by Daniels *et al.* (2008) to uniquely identify parasite isolates [41]. In other words, we only simulated the barcode SNPs, which in the model appear as an ordered list of 24 positions each with just two possible alleles (0 or 1). For simplicity, the model remains bi-allelic despite evidence that one of the SNPs, SNP #15 (*Pf*_07_001415182), is tri-allelic [42]. These 24 positions are located throughout the 14 *P. falciparum* chromosomes at known points (see S1 Appendix for the exact locations). As the distance between SNPs will effect the likelihood of recombination, we considered the precise position of the barcode SNPs when we formalized recombination (see S1 Appendix for recombination algorithm). It is important to note that diversity, as we calculate it, can only be non-zero when multiple genetically distinct parasite populations are present. Furthermore, for ease of notation, throughout the manuscript we will refer to the simulated sequence of 24 positions representing each unique parasite as its *barcode sequence*.

Genetically distinct parasite populations. The starting pool of gametocytes, emulating an infectious blood meal from a human host, consisted of a single barcode sequence for each parasite subpopulation present. In the case of infections consisting of two parasites, the founding parasite subpopulation barcode sequences differed by $0 \leq p_B \leq 24$ positions. For example, when $N = 2$ there were two genetically distinct parasite populations whose numbers came from from the underlying life cycle model (in either its deterministic or stochastic implementation). Following fertilization between these two distinct parasites, the resulting barcode sequences were either of type $Z_{1,2}$ or $Z_{2,1}$ (described above in the life cycle model). The order of the subscript (female first, male second) denoted from which parasite population each parent came. The number of parasites in each stage up to the formation of oocysts occurred with loss, the magnitude of which was associated with the fitness bias of the starting parasite.

Simulating recombination and reassortment. While the barcode sequences were paired in the oocyst stage, we simulated recombination and reassortment. To simplify the genetic mechanism of this study, during recombination, crossovers were simulated on each chromosome according to a Poisson distribution [43, 44] with mean one [44]. When a crossover occurred, the location of the crossover was chosen uniformly across the length of the chromosome. Multiple crossover events occurred iteratively. Following each crossover event, the resulting individual haploid strands were chosen uniformly for the next crossover event and

then the location of the crossover was chosen uniformly across the length of the chromosome. As only the allele at the chromosomal position of the barcode SNP were recorded, not all crossovers changed the barcode sequence.

Following recombination, we simulated reassortment by selecting a single version of each chromosome to be packaged together. When only a single barcode SNP position appeared on a chromosome, whether the recorded allele changed depended solely on reassortment. When multiple barcode SNP positions appeared on the same chromosome, then recombination could separate the recorded alleles, with the probability dependent upon the distance between the SNP positions on the chromosome. In this way, whether or not exchanges of genetic material occurred between the two mating parasites was determined for each of the 14 chromosomes, producing up to four unique barcode sequences per mating event.

Measuring diversity. We measured the diversity of the within-vector parasite population in two ways: nucleotide diversity and number of unique barcode sequences. We used a standard measure of nucleotide diversity, π, to measure the genetic variation in the within-mosquito population, $\pi = \sum_{ij} \frac{n_i n_j H_{ij}}{L}$, where n_i and n_j are the frequency of genetically distinct parasite subpopulations i and j, H_{ij} is the Hamming distance between the two parasite subpopulations, and L is the total number of positions in each sequence. The number of unique barcode sequences that occurred in each mosquito represents a population richness type metric, commonly used in ecology. We chose these metrics to account for similarity to the entering parasite populations (identity by decent) due to recombination (see S1 Appendix for a further discussion of diversity metrics).

Comparisons of the diversity were made at each of the following stages: (1) gametocytes upon entering the mosquito, (2) burst oocysts in the mosquito, (3) sporozoites in the mosquito salivary glands, and (4) sporozoites found in an infectious bite. We assumed an infectious bite contained ten sporozoites randomly chosen from the total pool of salivary gland sporozoites [45].

Results and discussion

Comparing the CTMC model to the deterministic model

Model of a single genetically distinct parasite population. The CTMC model of the parasite life cycle captures the average dynamics generated by the deterministic implementation of the model (Fig 1). However, the CTMC model is also able to capture variation in the life cycle. In the initial parasite stages, namely the gamete and zygote stages, the variation in the dynamics of the stochastic model is relatively small, and therefore the average dynamics captured by the deterministic model provides an accurate approximation to the population dynamics exhibited by its stochastic counterpart. However, as the parasites progress through successive stages, the variation in the stochastic output increases, and the mean captured by the deterministic model becomes less and less representative of the parasite population in any given mosquito. For example, using the baseline parameter values, the deterministic model results in a positive number of parasites at each stage of the within-vector life-cycle. However, 2.58% of the 10,000 stochastic simulations resulted in no sporozoites (Fig 1). The inability of the deterministic model to measure variation in parasite numbers could lead to erroneous estimates of parasite diversity at the population level. Furthermore, the fact that the deterministic model is continuous in state poses a challenge when linked to the sequence diversity model, since fractional oocysts still lead to the production of sporozoites.

We track the parasite dynamics within each mosquito for a time span of 21 days in all model simulations. As expected, sporozoite prevalence in the model, defined as the proportion out of 10,000 simulations resulting in at least one sporozoite in the salivary glands by day 21,

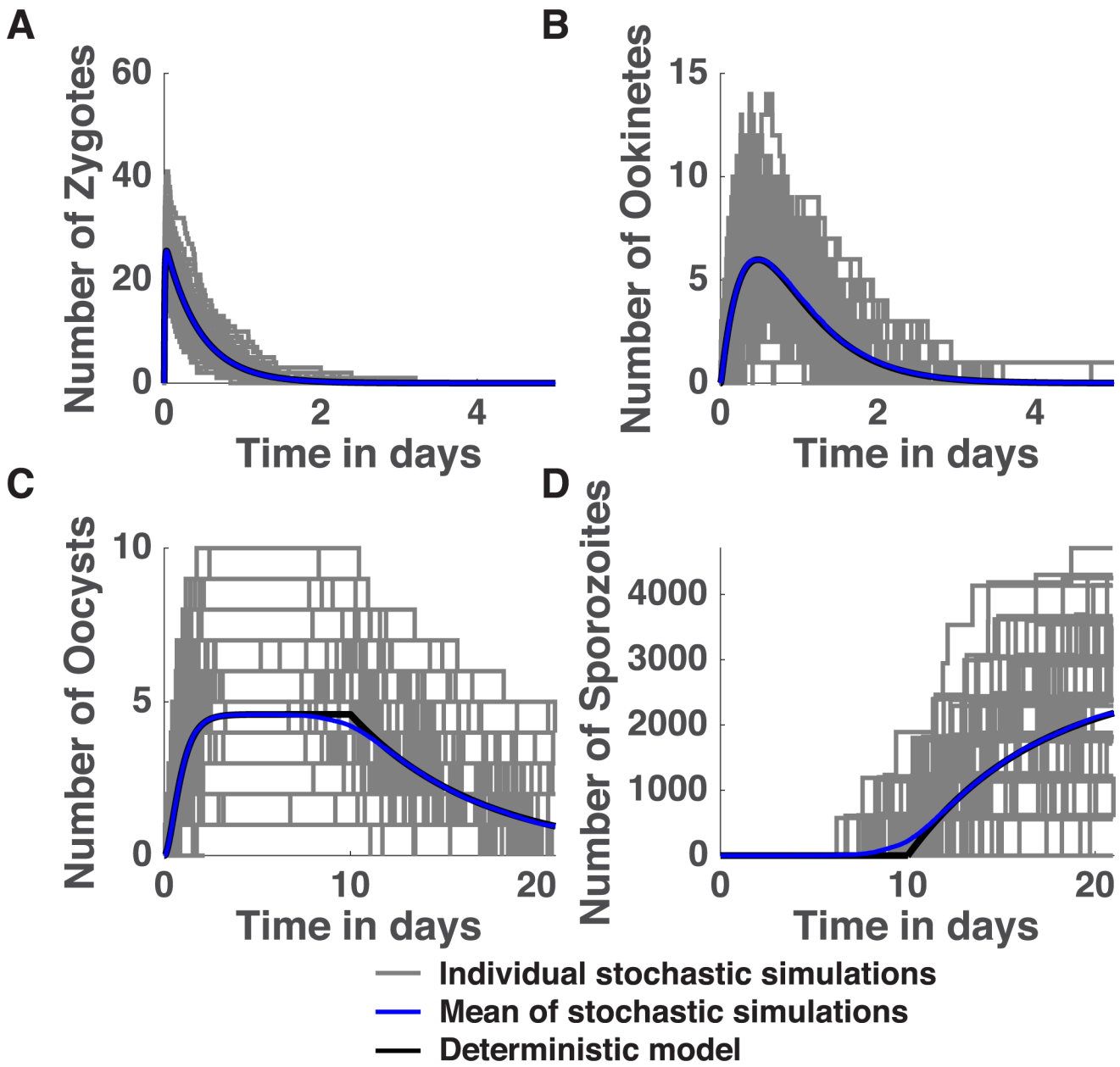


Fig 1. Dynamics of the within-vector model. The temporal dynamics of the deterministic model (black line) and the average dynamics across 1000 stochastic simulations (blue curve) overlays 100 simulations of the single-parasite population CTMC model (gray lines) with initial gametocyte number $G_0 = 300$ for zygotes (A), ookinetes (B), oocysts (C) and sporozoites (D). Observe that not all simulations produce a positive number of sporozoites. In fact, out of 10,000 simulations, 2.58% produce no sporozoites.

https://doi.org/10.1371/journal.pone.0177941.g001

increases with increasing initial gametocyte number G_0. In fact, for $G_0 = 150, 200, \ldots, 450$, the sporozoite prevalence is 62.90%, 82.98%, 92.95%, 97.42%, 99.47%, 99.84%, 99.95%, respectively. In S3 Fig, we show the dependence of oocyst and sporozoite prevalence as a function of G_0 (with G_0 ranging from 10 to 900), on days 7, 14, and 21. On day 14, all simulations with an initial gametocyte number of $G_0 = 500$ or greater harbor sporozoites in the salivary glands. However, on day 7 of a mosquito infection, most simulated mosquitoes will not be infectious to humans, even with an initial number of gametocytes equal to $G_0 = 900$.

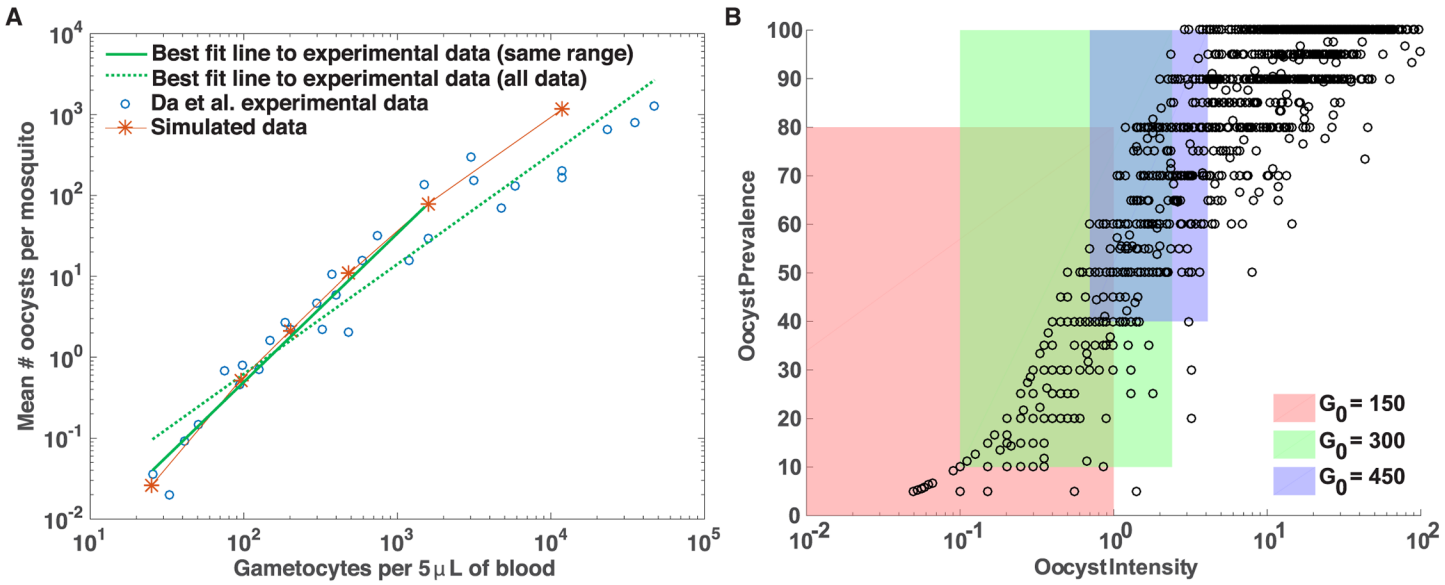


Fig 2. Comparison of life cycle model with published data. (A) Experimental data (blue circles) extracted (using Plot Digitizer) from Figure 1A in [46] along with the mean number of oocysts per mosquito calculated from our simulation data. The experimental data is scaled to be in units of gametocytes per five μL rather than gametocytes per one μL. A regression line to the log-log experimental data (dashed green) and to a subset of the experimental data over the range of initial gametocyte densities used to parameterize the model (solid green) is shown along with our model simulations (red stars). (B) Experimental data (black circles) from Figure 1 in [47] of oocyst intensity and prevalence (reproduced with permission) compared to results of the simulation starting from 150 (red shading), 300 (green shading) and 450 (blue shading) gametocytes. The colored areas represent the range of oocyst intensity and prevalence observed from simulations.

https://doi.org/10.1371/journal.pone.0177941.g002

Over the range of initial gametocyte densities that we consider in our model, oocyst densities are similar to reported by Da *et al* [46] suggesting that our model construction and parameterization produces realistic oocyst densities, at least in the initial gametocyte range $G_0 \in [0, 1000]$ (see Fig 2A). Da *et al* [46] dissected 1,636 female *Anopheles coluzzi* mosquitoes seven days after feeding them with blood containing gametocyte densities ranging between 80 and 9,520 gametocytes per microliter. For each initial gametocyte density, the mean number of oocysts per mosquito was reported. In our model, we assume a bloodmeal is five μL; thus, $G_0 = 450$ indicates that there are 450 gametocytes per 5 μL. In Fig 2A, the experimental data is scaled to be in units of gametocytes per five μL. Despite differences in mosquito species (our model assumes parameter values for *Anopheles gambiae*, while the experimental data consists of measurements from *Anopheles coluzzi*), the mean number of oocysts per mosquito as a function of initial gametocyte density per five μL of blood is remarkably consistent between the experimental data in [46] and our simulation data, particularly for $G_0 \in [0, 1000]$. Thus, we have confidence the simulation results presented in S3 Fig are realistic. When the full range of gametocyte densities considered in [46] is included in the regression analysis, the slope of the regression line is smaller, suggesting that the pattern observed for $G_0 \in [0, 1000]$ is different for gametocyte densities beyond this range, namely the number of oocysts produced per gametocyte is diminished at very high initial gametocyte densities.

Model of two genetically distinct parasite populations. A range of dynamics are observed for the two parasite population CTMC model (Fig 3A): the order of which subpopulations are most numerous at each change can change while the deterministic model always gives identical ordering (Fig 3B). For example, while the deterministic model always produces a positive number of oocysts and sporozoites, some of the stochastic simulations produce no

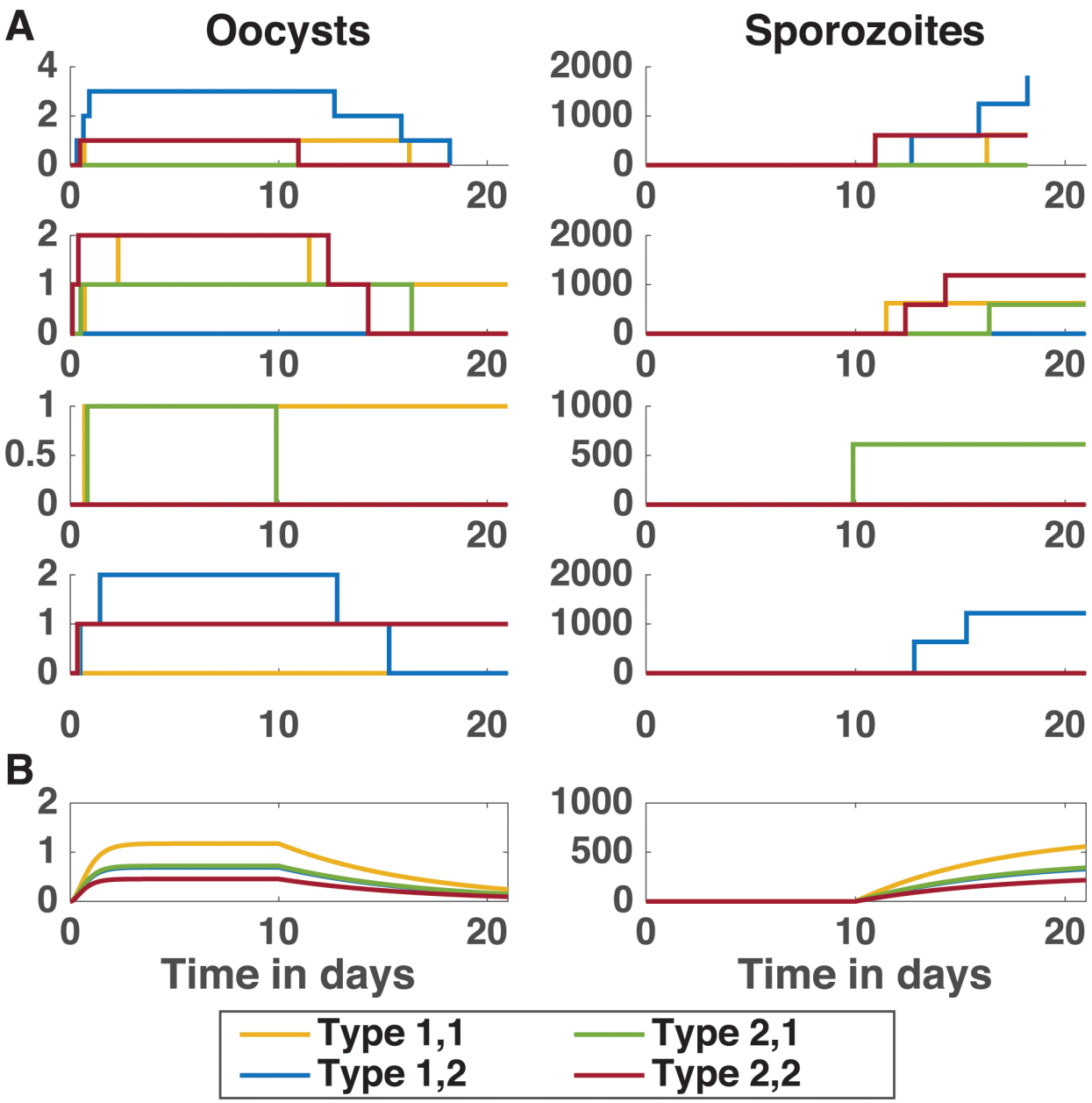


Fig 3. Comparison of stochastic and deterministic dynamics of oocysts and sporozoites. (A) Each of the first four rows illustrate the dynamics of the oocyst (left column) and sporozoite (right column) populations from individual simulations of the CTMC model using 50% fitness bias and $G_0 = 300$. (B) The dynamics of the deterministic model for oocysts (left) and sporozoites (right). Observe the variation in dynamics possible with the CTMC model.

https://doi.org/10.1371/journal.pone.0177941.g003

oocysts, and some of the oocysts that do arise never burst within the 21-day timespan. Furthermore, the deterministic model always results in type 1,1 outperforming type 2,1, followed closely by type 1,2, and finally type 2,2 parasites at each parasite stage. The stochastic model yields the same results on average (taken across all 10,000 simulations), but in contrast to the deterministic model, individual simulations can exhibit a wide range of behavior, as suggested by Fig 3A, including simulations where the least fit parasite type 2,2 is the only one surviving to the sporozoite stage.

While Fig 3 shows some example simulations, when all 10,000 simulations of the CTMC model are combined, the CTMC remains a good approximation of the deterministic model (Fig 4). The average cumulative number of oocysts formed by day 21 (Fig 4A) and the cumulative number of oocysts that rupture by day 21 (Fig 4B) are nearly identical for the stochastic and deterministic models, illustrated by colored circles and black crosses, respectively.

Recall that the deterministic model employs a discontinuous rupture function, which we approximate in the stochastic model by a continuous rupture function to allow some oocysts to rupture before day ten, with low probability. To ensure that differences were not the result of our choice of rupture function, we compared simulations of the stochastic model using the identical discontinuous rupture function. Our results revealed that assuming a continuous rupture function has little impact on cumulative oocyst and rupture numbers on day 21 (see S4 Fig) or the distribution of cumulative oocysts ruptured (see S5 Fig). Despite these similarities, the two types of rupture functions lead to different distributions of rupture timing, with most oocysts rupturing at intermediate values of time under the continuous rupture function while time to oocyst rupture decays with time under the step rupture function (see S6 Fig). This notable difference in the distribution of rupture times, and thus difference in the appearance of sporozoites in the salivary glands, has potentially important implications for onward transmission to a human host, given that female mosquitoes take multiple bloodmeals at different times throughout their life cycle.

We compared our simulations to experimental work conducted by Stone *et al* [47] in which prevalence and intensity of oocysts were measured. *Oocyst prevalence* is defined as the proportion of mosquitoes harboring oocysts (i.e. simulations producing at least one oocyst). *Oocyst intensity* is defined as the number of oocysts per mosquito (i.e. per simulation). Stone *et al.* calculated these quantities experimentally by computing the oocyst prevalence and mean oocyst intensity between days six and nine, post infection, for groups of *A. stephensi* and *A. gambiae* mosquitoes with a sample size greater than ten [47]. They included a total of 21,240 mosquitoes in these calculations. To emulate these experiments, we randomly drew ten mosquitoes (that is, ten stochastic realizations) from our 10,000 simulations of the stochastic model and computed the oocyst prevalence and intensity for these ten mosquitoes on day seven (to be consistent with experimental comparisons [47]). We then repeated this calculation 1000 times to produce the scatter plots and distributions in Fig 5. Consequently, each point in the scatter plot represents an average taken over a group of ten mosquitoes (simulations), and the resulting plots are analogous to Fig 1 in Stone *et al.* [47] (Note: the x-axes of the scatter plots are restricted to the interval [0, 4] for illustrative purposes; the range exceeds [0, 4] in some simulation experiments). Under a 0% bias, approximately 63.0% and 99.9% of simulations resulted in ruptured oocysts, when $G_0 = 150$ and $G_0 = 450$, respectively. Under a 10% bias, the prevalence of ruptured oocysts is 60.3% and 99.9%, respectively. Under a 50% bias, the prevalence of ruptured oocysts is 48.6% and 99.5%, respectively.

To facilitate a comparison between our simulation results presented in Fig 5 and the experimental results in Stone et al. [47], we have recreated Figure 1 in [47], with summary results of our simulations overlaying the scatter plot of the experimental data (see Fig 2B). Although Figs 5 and 2B differ quantitatively, the figures are qualitatively similar. Furthermore, our summary statistics and those of the Stone *et al* experiments are surprisingly comparable, particularly when taking into consideration the timespan of the Stone *et al* experiments. The rectangles depicted in Fig 2B illustrate the range of oocyst intensity and oocyst prevalence values obtained from each group of 10 simulated mosquitoes in Fig 5, for initial gametocyte densities $G_0 = 150$, 300, and 450 and a 10% fitness bias. For example, for $G_0 = 300$ and a 10% fitness bias, the range in mean oocyst intensity is 0.1 to 2.4, the range in mean oocyst prevalence is 0.1 to 1, and the median number of oocysts across all 10,000 simulated mosquitoes is 1. The range in

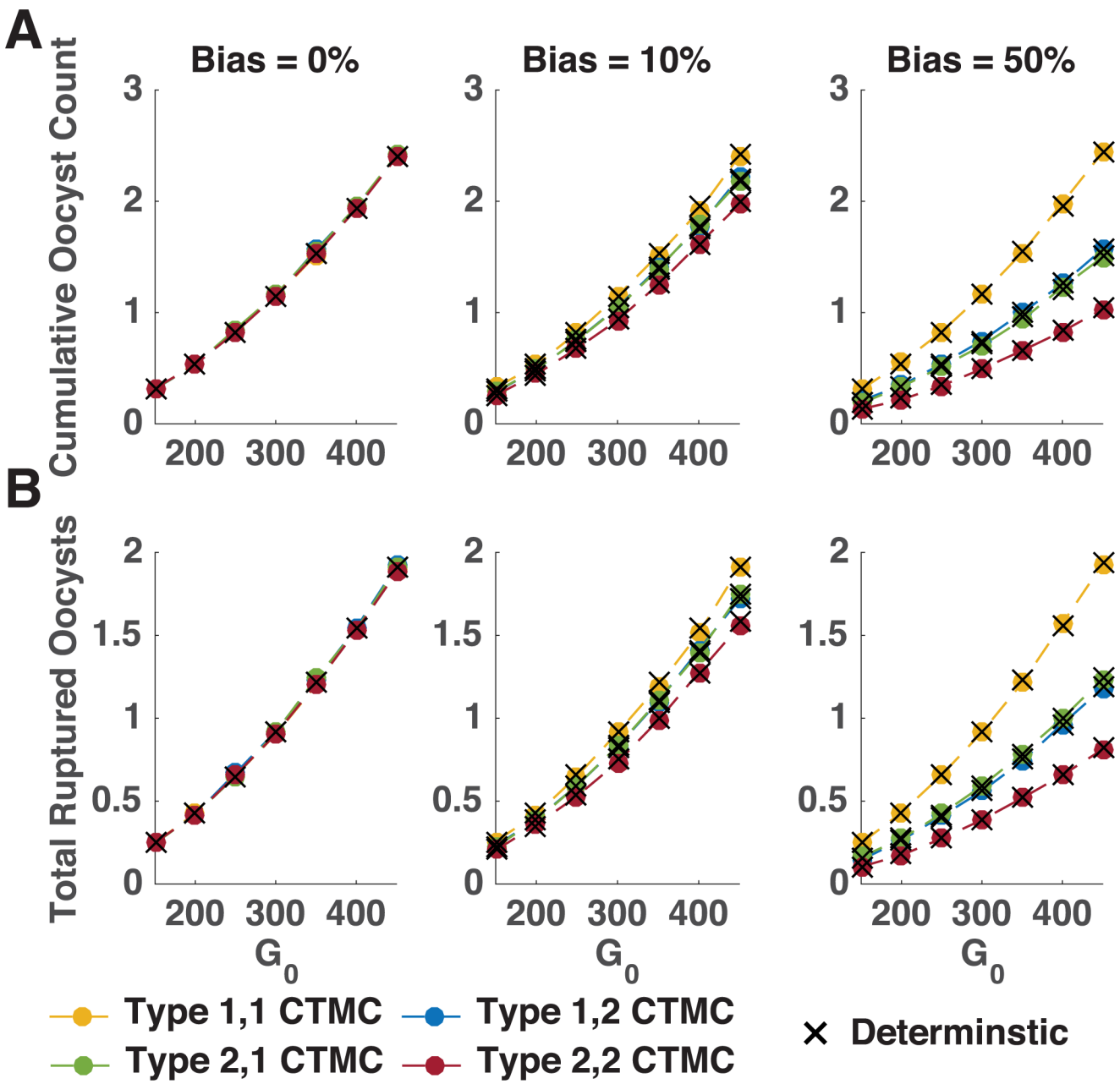


Fig 4. Cumulative and ruptured oocysts. The mean cumulative number (on day 21) of oocysts formed (A) and the total number of ruptured oocysts (B) resulting from the 10,000 simulations of the two-parasite population CTMC model (dots) and deterministic model (crosses) is illustrated for initial gametocyte number G_0. Here, the CTMC model utilizes a more biologically realistic continuous rupture function for oocyst bursting, compared to the step function used in the deterministic model. (See S4 Fig for the comparison of models both using a step bursting function.) Left to right the columns illustrate the results for 0%, 10% and 50% fitness biases, respectively.

https://doi.org/10.1371/journal.pone.0177941.g004

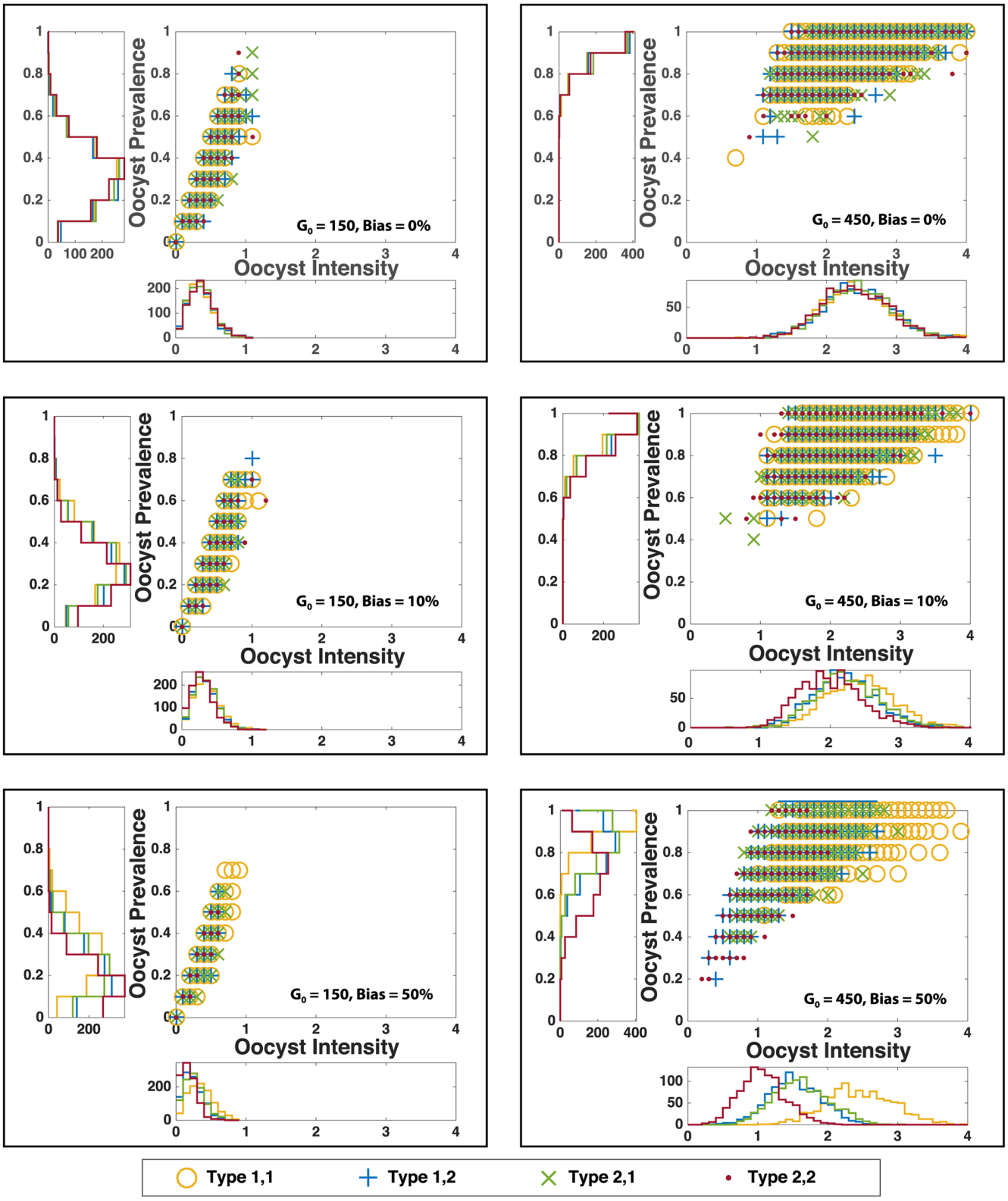


Fig 5. Oocyst intensity and prevalence. The scatter plots show mean oocyst intensity (x-axis) versus mean oocyst prevalence (y-axis), calculated 1000 times. *Oocyst prevalence* is defined as the proportion of mosquitoes harboring oocysts (i.e. simulations producing at least one oocyst). *Oocyst intensity* is defined as the number of oocysts per mosquito (i.e. per simulation). To produce each point, we randomly drew ten mosquitoes (that is, ten

stochastic realizations) from our 10,000 simulations of the stochastic model and computed the oocyst prevalence and intensity for these ten mosquitoes on day seven. The histograms show the density of points in the scatter plot for different values of oocyst intensity (horizontal graph) and oocyst prevalence (vertical graph). In the left and right columns, the initial gametocyte number is $G_0 = 150$ and $G_0 = 450$, respectively. Top to bottom the rows illustrate the results for 0%, 10% and 50% fitness biases, respectively. Observe that as G_0 increases, the differences in distribution of the genetically distinct parasite populations becomes more pronounced.

https://doi.org/10.1371/journal.pone.0177941.g005

oocyst numbers across all 10,000 simulated mosquitoes is 1 to 7 for these starting conditions. These summary statistics for each initial gametocyte number and for each fitness bias are presented in S2 Table. For comparison, the Stone *et al.* experiments yielded an infection prevalence, measured at seven days post infection, ranging from 33% to 86.5%, and mean oocyst intensity ranging between 0.57 and 4.7. Furthermore, our simulation results yielded approximately 56%, 41%, and 21%, on average, of oocysts remained intact on days 7, 14, and 21, respectively, whereas the experiments in [47] showed that 25% of oocysts remained intact 21 days post infection, with few oocysts rupturing between day 14 and day 21.

Diversity results

During the life cycle of the parasite in the mosquito host, there is a severe bottleneck at the oocyst stage with only a handful of parasites able to successfully mate and produce sporozoites. Due to this bottleneck, even with the presence of reassortment and recombination among chromosomes, the sporozoite population would not be expected to have an increased nucleotide diversity compared to the entering gametocyte population, in the absence of new mutations. However, the variation exhibited in the oocyst and sporozoite populations may include newly generated combinations of existing genetic pieces, creating novel variation. Naturally, this variation cannot occur when only a single genetically distinct parasite population is present because any recombination or reassortment would be with the same original barcode sequence.

Modeling two genetically distinct parasite populations. Fig 6 shows an increase in the number of new genetically distinct parasite populations as quantified by the number of unique parasite barcode sequences present at the oocyst stage. New barcode sequences arise through the reassortment and recombination of existing alleles. As the two initial parasite populations decrease in similarity, i.e. the number of differences in barcode sequences (p_B) increases, the number of unique barcode sequences rapidly increases before leveling off. The sharp increase is observed until approximately 10 differences exist between the parasite populations. As the initial barcode sequences become more disparate, the possible ways to create new parasite barcode sequences is so great that almost every new barcode sequence generated is unique and the number of unique barcode sequences is limited by the total number of barcode sequences created during oocyst formation. Fig 6B shows that at every initial gametocyte number there is a substantial peak of simulations that harbor only a single genetically distinct parasite population, the result of parasites only mating with their identical siblings. As the initial gametocyte number increases, more of the simulations result in a higher number of unique barcode sequences, i.e. the density of peaks shifts to the right. Because recombination between two parasites generally results in one or four unique barcode sequences (if they are identical or different, respectively) certain numbers of unique barcode sequences are not possible in a single mosquito. For example, it is not possible to obtain three unique sequences in a single mosquito when starting with two parasite subpopulations. Our results are consistent with experimental

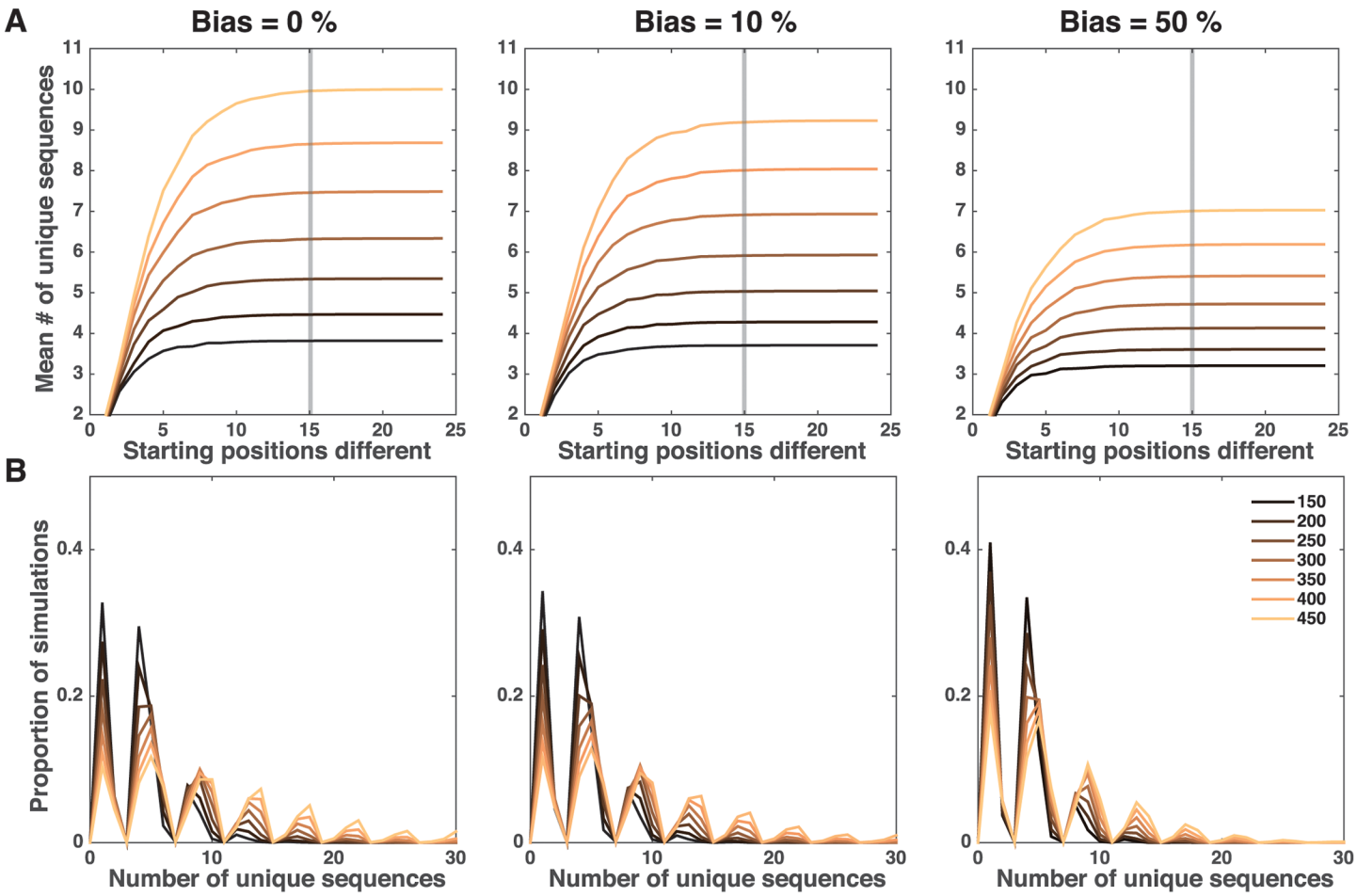


Fig 6. Number of unique barcode sequences. (A) The number of unique barcode sequences from two genetically distinct parasite populations as a function of the number of barcode positions that initially differ. (B) Distribution of the unique barcode sequences when looking at 15 differences between the two initial parasite populations. Colors indicate the number of starting gametocytes with darker colors representing lower initial gametocyte numbers. Left to right the columns illustrate the results for 0%, 10% and 50% fitness biases, respectively. Gray bars in (A) show the similarity in parasite populations for the distributions in (B).

https://doi.org/10.1371/journal.pone.0177941.g006

results, which used microsatellite markers to quantify the parasite diversity within a mosquito. They found fluctuations of genotypes, specific parasite sequences, throughout development and the presence of genotypes in the sporozoite population that were not detected in the gametocyte population [29].

In Fig 7, the nucleotide diversity increases linearly as the two genetically distinct parasite populations decrease in similarity, as measured by the number of simulated barcode positions which differ. The rate of linear increase, however, depends on the size of the initial gametocyte population (line colors) as well as the fitness bias of one parasite population (0% on the left to 50% on the right). The increase with higher initial gametocyte number is further enhanced as more initial positions differ between the two parasite populations. Comparatively, the strength of the fitness bias only slightly impacts the increase in nucleotide diversity, with greater differences observed when the initial parasite populations are most disparate (Fig 7, right panel).

This decrease in diversity observed in populations where one parasite population has a high fitness bias is due to the decreased survival of each of the life stages for the low fitness

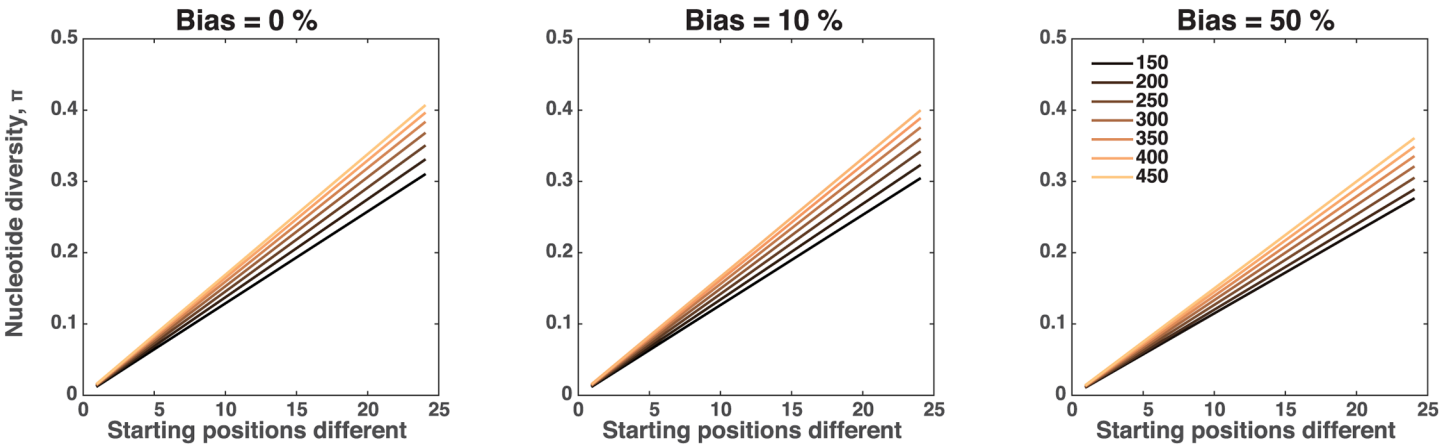


Fig 7. Nucleotide diversity of ruptured oocysts. Nucleotide diversity, π, by differing number of starting positions between genetically distinct parasite populations. Left to right the columns illustrate the results for 0%, 10% and 50% fitness biases, respectively. Colors refer to the initial gametocyte number with darker colors representing lower gametocyte numbers.

https://doi.org/10.1371/journal.pone.0177941.g007

population. When the success of one parasite population is severely limited, the potential for diversity in the full population is also limited as no new diversity can be created, either by reassortment or recombination, when a parasite mates with an identical sibling. The increase in nucleotide diversity observed as the number of differing positions increases is due to the reduction in the proportion of the population that exhibits no diversity, i.e. the proportion where only identical sequences are found (see S7 Fig). Regardless of the size of the initial population of gametocytes, the location of the dominant peak of nucleotide diversity does not shift, given a level of starting diversity between the two parasite populations. The parasite diversity within a mosquito does not significantly change between the oocyst stage and the sporozoites found in the salivary glands (see S8 Fig) despite the loss of 80% of sporozoites following oocyst bursting and travel of the sporozoites to the salivary glands. In fact, we observe no loss in the number of genetically distinct parasites from burst oocysts to sporozoites in the salivary glands, so the number of unique sequences does not change (not pictured). However, the nucleotide diversity shifts slightly depending on if the survival of parasites increases or decreases the evenness of populations.

Fig 8 illustrates that even though only ten sporozoites compose an infectious bite (mean from [48]) the complexity of infection (COI) frequently remains the same or even increases with multiple parasites transmitted simultaneously. With just 150 initial gametocytes from an infection with COI two, nearly 66% of infectious bites contain at least two genetically distinct parasites for onward transmission and over 50% of which increase the COI. As the initial gametocyte number grows nearly all infectious bites contain multiple genetically distinct parasites, maintaining or frequently increasing the COI that is transmitted. Despite observing high levels of COI, several studies have noted that the observed heterozygosity falls below the expected heterozygosity, indicating that novel genotypes may not be generated and transmitted as frequently as predicted [49, 50].

Diversity generation using the underlying deterministic model. As the deterministic model, by definition, predicts the same number of gametes, ookinetes, oocysts and sporozoites for a given starting gametocyte population, the resulting diversity observed is severely constrained. When the initial gametocyte population is small ($G_0 < 350$), the deterministic model predicts fractions of oocysts. In order to compare to results based on life cycle numbers from

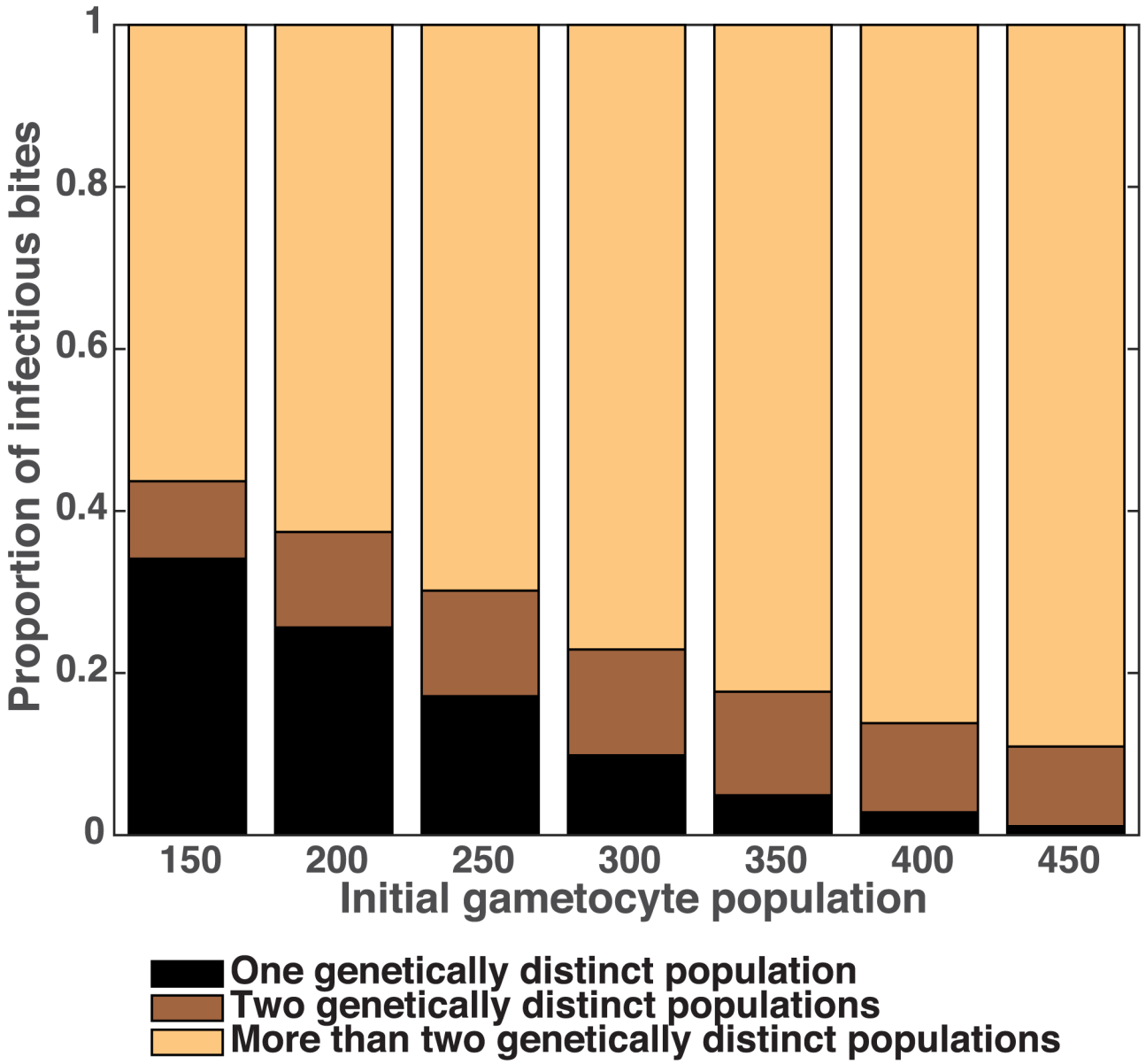


Fig 8. Frequency that multiple genetically distinct parasites are passed in an infectious bite. The fraction of infectious bites that harbor one (black), two (brown) or multiple (beige) parasites with distinct barcode sequences. When the mosquito is infected with a small number of gametocytes, 34% of infectious bites are composed of a single sequence. At larger initial gametocyte numbers the proportion of infectious bites passing only a single genetically distinct parasite population falls to nearly 0. Ten sporozoites were assumed to be present in a single infectious bite.

https://doi.org/10.1371/journal.pone.0177941.g008

the CTMC model, we round up to the next highest oocyst number. Thus, exactly four oocysts survive, (one of each type of cross between the two parasite populations), leading to an identical number of unique barcode sequences when $G_0 < 350$ (see S9 Fig). For larger initial gametocyte numbers, the rounded deterministic model predicts eight oocysts—two of each parasite pairing. This also results in an identical number, albeit higher that what was found with $G_0 < 350$, of unique barcode sequences—so the nucleotide diversity is identical, regardless of the starting gametocyte number (see S10 Fig). The clear differences between the predicted

diversity when using the life cycle numbers from implementation with the underlying stochastic and deterministic models underscores the necessity of considering variation in the parasite development within the mosquito.

Conclusion

We have introduced a stochastic model of the development of the malaria parasite within the vector to understand the development of diversity in the malaria parasite population. Our results indicate that incorporating variation in the dynamics of life stages of *Plasmodium falciparum* within the mosquito vector can impact the level and diversity of sporozoites produced during the life span of the mosquito. The variation in sporozoite populations predicted, which differ from the implicit assumption of identical sporozoite numbers made by many transmission models [5–8], have potential consequences for the capacity for onward transmission. When modeling diversity generation, we find that the presence of multiple parasite populations within a single mosquito is essential for rapidly increasing diversity in a parasite population. Overall the severe bottlenecks in the life cycle of the parasite reduce population level diversity, but the opportunity for sexual recombination can reintroduce variation when multiple genetically distinct parasite populations are present simultaneously.

Here, we have only considered the dynamics within individual mosquitoes. The next step will be to couple the dynamics within individual mosquitoes with population-level dynamics of transmission, which will allow for inference on infection prevalence. This requires including the human host as well as the vector. Understanding of the potential for diversity generation in the mosquito vector is essential for monitoring and predicting the spread of genetic resistance of the parasite population.

Supporting information

S1 Fig. Diagram of the underlying within-vector parasite dynamics model. Male (M) gametes fertilize female (F) gametes, creating new zygotes (Z) at rate rMF. Z zygotes progress to the ookinete stage (E) at rate $\sigma_z Z$, and E ookinetes progress to the oocyst stage (O) at rate $\sigma_e E$. Finally, sporoblast formation occurs within the oocysts, leading to $k(t)O$ oocysts rupturing, releasing $n_0 k(t)O$ sporozoites, a fraction p of which successfully migrate to the mosquito's salivary glands (depicted in red). All rates are per unit time.
(EPS)

S2 Fig. Diagram of the within-vector parasite diversity generation model. Male (M) and female (F) gametes enter with their unique barcode sequence depicted by gray or white boxes. Each of the 24 positions of the barcode sequence are represented. Adjacent boxes on segments indicate they are found on the same chromosome while ones separated by a small space are on different chromosomes. Only 12 segments are shown as two chromosomes do not have barcode SNPs. The gametes form a new zygote (Z) bringing the genetic material of both parasites together. Within the oocyst stage (O) the parasites undergo meiosis, briefly forming tetraploid cells where reassortment and recombination of sequences is possible. A cross (x) denotes an example of a position of recombination; in this case all positions to the left of the cross on the same chromosome are swapped between the two barcode sequences. Dashed lines show reassortment of chromosomes, swapping all positions on the chromosome. The new barcode sequences then proliferate asexually within the oocyst, which ultimately rupture producing sporozoites (S), each containing one of up to four unique barcode sequences. Pictured here are male and female gametes with maximally different barcode sequences, i.e. all white or all gray boxes, respectively. For simplicity only a single recombination event and a several

reassortments are depicted. If the male and female gametes entering are siblings and thus genetically identical, no new genetic combinations are possible and all sporozoites will contain identical barcode sequences. We do not show stages where the parasite diversity is unchanged, such as the ookinete stage.
(EPS)

S3 Fig. Oocyst and sporozoite prevalence versus initial gametocyte number. The proportion of mosquitoes harboring oocysts (A) and sporozoites (B) on days 7, 14, and 21 are plotted as a function of the initial gametocyte number G_0 (from $G_0 = 10$ to 900).
(EPS)

S4 Fig. Cumulative formed and ruptured oocysts with step function. The mean cumulative number (on day 21) of oocysts formed (A) and the total number of ruptured oocysts (B) resulting from the 10,000 simulations of the two-parasite population CTMC model (dots) and deterministic model (crosses) is illustrated for initial gametocyte number G_0. Here, the CTMC model utilizes a step rupture function for oocyst bursting, analogous to what is used in the deterministic model. In comparison, Fig 3 uses a continuous time bursting function that is more biologically realistic. Left to right the columns illustrate the results for 0%, 10% and 50% fitness biases, respectively. The stochastic implementation of the life cycle model accurately replicates the mean of the number of formed and burst oocysts.
(EPS)

S5 Fig. Comparison of cumulative rupture counts for continuous and step rupture functions. A cumulative count of rupture times using the continuous rupture function (A) and the step rupture function (B) resulting from the 10,000 simulations of the two-parasite population CTMC model. Left to right the columns illustrate the results for 0%, 10% and 50% fitness biases, respectively. The initial gametocyte number is $G_0 = 300$ in all cases.
(EPS)

S6 Fig. Comparison of rupture time frequencies for continuous and step rupture functions. The frequencies of rupture time for the continuous rupture function (A) and the step rupture function (B) resulting from the 10,000 simulations of the two-parasite population CTMC model. Left to right the columns illustrate the results for 0%, 10% and 50% fitness biases, respectively. The initial gametocyte number is $G_0 = 300$ in all cases.
(EPS)

S7 Fig. Composition of nucleotide diversity. (A) Mean nucleotide diversity, π, as gametocyte number changes at varying number of starting positions between barcode sequences of the two parasite populations (left to right: 5, 10, 15, 20). When the nucleotide diversity is zero, it is excluded from the mean. 0% fitness bias (blue pluses), 10% fitness bias (red circles) and 50% fitness bias (black crosses) are considered. (B) Histogram of nucleotide diversity, π, in the population with 10% bias at varying number of differing of starting positions between barcode sequences of the two parasite populations (left to right: 5, 10, 15, 20). Observe that increasing the initial gametocyte number only affects the distribution of the population but not the diversity at the peak.
(EPS)

S8 Fig. Nucleotide diversity comparison between oocysts and sporozoites. Nucleotide diversity, π, of the burst oocyst population (x-axis) and sporozoite population in the salivary gland (y-axis). The nucleotide diversity is nearly identical between the two measures, with differences accounted for by changes in population size. Neither appearance nor disappearance of barcode sequences is observed. Points have been subsampled randomly from all initial gametocyte

number and biases. A guide ($x = y$) line is provided in red for reference.
(EPS)

S9 Fig. Number of unique sequences with the deterministic model. (A) The number of unique sequences produced as the number of starting positions differs between barcode sequences of the two parasite populations. Left to right the columns illustrate the results for 0%, 10% and 50% fitness biases, respectively. Observe that for initial gametocyte numbers $G_0 < 350$ the lines are overlapping. (B) Distribution of the unique sequences when looking at 15 differences in the barcode sequences of the two parasite populations. Left to right the columns illustrate the results for 0%, 10% and 50% fitness biases, respectively. Colors indicate the number of starting gametocytes. Gray bars in (A) show amount of similarity in the two parasite populations for the distributions in (B).
(EPS)

S10 Fig. Nucleotide diversity with the deterministic model. Nucleotide diversity, π, by differing number of starting positions between barcode sequences of the two parasite populations. Left to right the columns illustrate the results for 0%, 10% and 50% fitness biases, respectively. Colors indicate different initial gametocyte numbers but lines are overlapping.
(EPS)

S1 Appendix. Description of simulation method and diversity generation protocol. We present the details of the simulation method for the underlying life-cycle model and the protocol for the generation of parasite diversity. Furthermore, we describe the process of recombination of parasites within-mosquito, and discuss alternative diversity metrics.
(PDF)

S1 File. MATLAB code for life-cycle model. MATLAB code for the underlying stochastic life-cycle model.
(PDF)

S2 File. MATLAB code for diversity model. MATLAB code for the underlying stochastic diversity model.
(PDF)

S1 Table. Two-parasite population CTMC transition matrix.
(PDF)

S2 Table. Oocyst prevalence and intensity summary statistics. The proportion of oocysts intact on day 7 ranges from approximately 56 to 59%; on day 21 the range is approximately 20-21%.
(PDF)

S3 Table. Position of barcode SNPs. The location of the barcode SNPs by chromosome and within each chromosome [41].
(PDF)

Acknowledgments

The authors thank T. Bousema and W. Stone for sharing the data for replication of Figure 1 in [47]. L.M.C. was supported by award number U54GM088558 from the National Institute of General Medical Sciences. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institute of General Medical Sciences or the National Institutes of Health. The authors acknowledge the Odyssey cluster supported by

the FAS Division of Science, Research Computing Group at Harvard University and Advanced Research Computing at Virginia Tech for providing computational resources and technical support that have contributed to the results reported within this paper.

Author Contributions

Conceptualization: LMC OFP.

Data curation: LMC OFP.

Formal analysis: LMC OFP.

Funding acquisition: LMC OFP.

Investigation: LMC OFP.

Methodology: LMC OFP.

Project administration: LMC OFP.

Resources: LMC OFP.

Software: LMC OFP.

Supervision: LMC OFP.

Validation: LMC OFP.

Visualization: LMC OFP.

Writing – original draft: LMC OFP.

Writing – review & editing: LMC OFP.

References

1. WHO. World malaria report 2013. World Health Organization; 2014.
2. Manske M, Miotto O, Campino S, Auburn S, Almagro-Garcia J, Maslen G, et al. Analysis of *Plasmodium falciparum* diversity in natural infections by deep sequencing. Nature. 2012; 487(7407):375–379. https://doi.org/10.1038/nature11174 PMID: 22722859
3. Bei AK, Diouf A, Miura K, Larremore DB, Ribacke U, Tullo G, et al. Immune Characterization of *Plasmodium falciparum* Parasites with a Shared Genetic Signature in a Region of Decreasing Transmission. Infect Immun. 2015; 83(1):276–285. https://doi.org/10.1128/IAI.01979-14 PMID: 25368109
4. Sinden R. Malaria, sexual development and transmission: retrospect and prospect. Parasitology. 2009; 136(12):1427–1434. https://doi.org/10.1017/S0031182009990667 PMID: 19660156
5. Ross R. The prevention of malaria. London: Murray; 1911.
6. Macdonald G. The epidemiology and control of malaria. Oxford University Press; 1957.
7. Koella JC. On the use of mathematical models of malaria transmission. Acta Trop. 1991; 49(1):1–25. https://doi.org/10.1016/0001-706X(91)90026-G PMID: 1678572
8. Smith DL, Battle KE, Hay SI, Barker CM, Scott TW, McKenzie FE. Ross, Macdonald, and a theory for the dynamics and control of mosquito-transmitted pathogens. PLoS Pathog. 2012; 8(4):e1002588. https://doi.org/10.1371/journal.ppat.1002588 PMID: 22496640
9. McQueen PG, McKenzie FE. Age-structured red blood cell susceptibility and the dynamics of malaria infections. P Natl Acad Sci USA. 2004; 101(24):9161–9166. https://doi.org/10.1073/pnas.0308256101
10. Maire N, Smith T, Ross A, Owusu-Agyei S, Dietz K, Molineaux L. A model for natural immunity to asexual blood stages of *Plasmodium falciparum* malaria in endemic areas. Am J Trop Med Hyd. 2006; 75:19–31. https://doi.org/10.4269/ajtmh.2006.75.19
11. Buckee CO, Recker M, Watkins ER, Gupta S. Role of stochastic processes in maintaining discrete strain structure in antigenically diverse pathogen populations. P Natl Acad Sci USA. 2011; 108 (37):15504–15509. https://doi.org/10.1073/pnas.1102445108

12. Eckhoff P. *P. falciparum* infection durations and infectiousness are shaped by antigenic variation and innate and adaptive host immunity in a mathematical model. PLoS One. 2012; 7(9):e44950. https://doi.org/10.1371/journal.pone.0044950 PMID: 23028698

13. Johnston GL, Smith DL, Fidock DA. Malaria's missing number: calculating the human component of R_0 by a within-host mechanistic model of *Plasmodium falciparum* infection and transmission. PLoS Comput Biol. 2013; 9(4):e1003025. https://doi.org/10.1371/journal.pcbi.1003025 PMID: 23637586

14. Childs LM, Buckee CO. Dissecting the determinants of malaria chronicity: why within-host models struggle to reproduce infection dynamics. J R Soc Interface. 2015; 12(104):20141379. https://doi.org/10.1098/rsif.2014.1379 PMID: 25673299

15. Teboh-Ewungkem MI, Yuster T. A within-vector mathematical model of *Plasmodium falciparum* and implications of incomplete fertilization on optimal gametocyte sex ratio. J Theor Biol. 2010; 264(2):273–286. https://doi.org/10.1016/j.jtbi.2009.12.017 PMID: 20122943

16. Teboh-Ewungkem M, Yuster T, Newman N. A mathematical model of the within-vector dynamics of the *Plasmodium falciparum* protozoan parasite. In: Tchuenche JM, Chiyaka C, editors. Infectious Disease Modelling Research Progress. New York: Nova Science Publishers; 2010; pp.177–199.

17. Vaughan JA, Noden BH, Beier JC. Population dynamics of *Plasmodium falciparum* sporogony in laboratory-infected *Anopheles gambiae*. J Parasitol.1992; 716–724. https://doi.org/10.2307/3283550 PMID: 1635032

18. Roitberg BD, Gordon I. Does the *Anopheles* blood meal-fecundity curve, curve? J Vector Ecol. 2005; 30 (1):83. PMID: 16007959

19. Eichner M, Diebner HH, Molineaux L, Collins WE, Jeffery GM, Dietz K. Genesis, sequestration and survival of *Plasmodium falciparum* gametocytes: parameter estimates from fitting a model to malariatherapy data. T Roy Soc Trop Med H. 2001; 95(5):497–501. https://doi.org/10.1016/S0035-9203(01)90016-1

20. Collins WE, Jeffery GM. A retrospective examination of mosquito infection on humans infected with *Plasmodium falciparum*. Am J Trop Med Hyg. 2003; 68(3):366–371. PMID: 12685646

21. Jeffery GM, Eyles DE. Infectivity to mosquitoes of *Plasmodium falciparum* as related to gametocyte density and duration of infection. Am J Trop Med Hyg. 1955; 4(5):781–789. https://doi.org/10.4269/ajtmh.1955.4.781 PMID: 13259002

22. Churcher TS, Bousema T, Walker M, Drakeley C, Schneider P, Ouédraogo AL, et al. Predicting mosquito infection from *Plasmodium falciparum* gametocyte density and estimating the reservoir of infection. Elife. 2013; 2:e00626. https://doi.org/10.7554/eLife.00626 PMID: 23705071

23. Reininger L, Billker O, Tewari R, Mukhopadhyay A, Fennell C, Dorin-Semblat D, et al. A NIMA-related protein kinase is essential for completion of the sexual cycle of malaria parasites. J Biol Chem. 2005; 280(36):31957–31964. https://doi.org/10.1074/jbc.M504523200 PMID: 15970588

24. Sinden R. Gametocytogenesis in *Plasmodium spp.*, and observations on the meiotic division. Ann Soc Belg Med Trop. 1985; 65(2):21–23. PMID: 4073978

25. Sinden R. The cell biology of malaria infection of mosquito: advances and opportunities. Cell Microbiol. 2015; 17(4):451–466. https://doi.org/10.1111/cmi.12413 PMID: 25557077

26. Gerald N, Mahajan B, Kumar S. Mitosis in the human malaria parasite *Plasmodium falciparum*. Eukaryot Cell. 2011; 10(4):474–482. https://doi.org/10.1128/EC.00314-10 PMID: 21317311

27. Rosenberg R, Wirtz RA, Schneider I, Burge R. An estimation of the number of malaria sporozoites ejected by a feeding mosquito. T Roy Soc Trop Med H. 1990; 84(2):209–212. https://doi.org/10.1016/0035-9203(90)90258-G

28. Branch OH, Sutton PL, Barnes C, Castro JC, Hussin J, Awadalla P, et al. *Plasmodium falciparum* genetic diversity maintained and amplified over 5 years of a low transmission endemic in the Peruvian Amazon. Mol Biol Evol. 2011; 28(7):1973–1986. https://doi.org/10.1093/molbev/msq311 PMID: 21109587

29. Morlais I, Nsango SE, Toussile W, Abate L, Annan Z, Tchioffo MT, et al. *Plasmodium falciparum* mating patterns and mosquito infectivity of natural isolates of gametocytes. PloS One. 2015; 10(4):e0123777. https://doi.org/10.1371/journal.pone.0123777 PMID: 25875840

30. Carter TE, Malloy H, Existe A, Memnon G, Victor YS, Okech BA, et al. Genetic diversity of *Plasmodium falciparum* in Haiti: Insights from microsatellite markers. PloS One. 2015; 10(10):e0140416. https://doi.org/10.1371/journal.pone.0140416 PMID: 26462203

31. Juliano JJ, Porter K, Mwapasa V, Sem R, Rogers WO, Ariey F, et al. Exposing malaria in-host diversity and estimating population diversity by capture-recapture using massively parallel pyrosequencing. P Natl Acad Sci USA. 2010; 107(46):20138–20143. https://doi.org/10.1073/pnas.1007068107

32. Ross A, Koepfli C, Li X, Schoepflin S, Siba P, Mueller I, et al. Estimating the numbers of malaria infections in blood samples using high-resolution genotyping data. PloS One. 2012; 7(8):e42496. https://doi.org/10.1371/journal.pone.0042496 PMID: 22952595

33. Nkhoma SC, Nair S, Cheeseman IH, Rohr-Allegrini C, Singlam S, Nosten F, et al. Close kinship within multiple-genotype malaria parasite infections. P Roy Soc B-Biol Sci. 2012; 279(1738):2589–2598. https://doi.org/10.1098/rspb.2012.0113

34. Sutton PL, Torres LP, Branch O. Sexual recombination is a signature of a persisting malaria epidemic in Peru. Malaria J. 2011; 10(1):329. https://doi.org/10.1186/1475-2875-10-329

35. Taylor L, Walliker D, Read A. Mixed-genotype infections of the rodent malaria *Plasmodium chabaudi* are more infectious to mosquitoes than single-genotype infections. Parasitology. 1997; 115(2):121–132. https://doi.org/10.1017/S0031182097001145 PMID: 10190168

36. Nassir E, Abdel-Muhsin AMA, Suliaman S, Kenyon F, Kheir A, Geha H, et al. Impact of genetic complexity on longevity and gametocytogenesis of *Plasmodium falciparum* during the dry and transmission-free season of eastern Sudan. Int J Parasitol. 2005; 35(1):49–55. https://doi.org/10.1016/j.ijpara.2004.10.014 PMID: 15619515

37. Nsango SE, Abate L, Thoma M, Pompon J, Fraiture M, Rademacher A, et al. Genetic clonality of *Plasmodium falciparum* affects the outcome of infection in *Anopheles gambiae*. Int J Parasitol. 2012; 42 (6):589–595. https://doi.org/10.1016/j.ijpara.2012.03.008 PMID: 22554991

38. Anderson TJC, Haubold B, Williams JT, Estrada-Franco JG, Richardson L, Mollinedo R, et al. Microsatellite markers reveal a spectrum of population structures in the malaria parasite *Plasmodium falciparum*. Mol Biol Evol. 2000; 17(10): 1467–1482. https://doi.org/10.1093/oxfordjournals.molbev.a026247 PMID: 11018154

39. Conway DJ, Roper C, Oduola AMJ, Arnot DE, Kremsner PG, Grobusch MP, et al. High recombination rate in natural populations of *Plasmodium falciparum*. P Natl Acad Sci USA. 1999: 96(8) 4506–4511. https://doi.org/10.1073/pnas.96.8.4506

40. Allen LJS. An introduction to stochastic processes with applications to biology. Boca Raton: CRC Press; 2010.

41. Daniels R, Volkman SK, Milner DA, Mahesh N, Neafsey DE, Park DJ, et al. A general SNP-based molecular barcode for *Plasmodium falciparum* identification and tracking. Malaria J. 2008; 7(1):223. https://doi.org/10.1186/1475-2875-7-223

42. Dwivedi A, Khim N, Reynes C, Ravel P, Ma L, Tichit M, et al. *Plasmodium falciparum* parasite population structure and gene flow associated to anti-malarial drugs resistance in Cambodia. Malaria J. 2016; 15(1):319. https://doi.org/10.1186/s12936-016-1370-y

43. Handel MA. Meiosis and gametogenesis. Elsevier; 1997.

44. Jiang H, Li N, Gopalan V, Zilversmit MM, Varma S, Nagarajan V, et al. High recombination rates and hotspots in a *Plasmodium falciparum* genetic cross. Genome Biol. 2011; 12(4):1. https://doi.org/10.1186/gb-2011-12-4-r33

45. Kappe SH, Vaughan AM, Boddey JA, Cowman AF. That was then but this is now: malaria research in the time of an eradication agenda. Science. 2010; 328(5980):862–866. https://doi.org/10.1126/science.1184785 PMID: 20466924

46. Da DF, Churcher TS, Yerbanga RS, Yaméogo B, Sangaré I, Ouedraogo JB, et al. Experimental study of the relationship between *Plasmodium* gametocyte density and infection success in mosquitoes; implications for the evaluation of malaria transmission-reducing interventions. Exp Parasitol. 2015; 149:74–83. https://doi.org/10.1016/j.exppara.2014.12.010 PMID: 25541384

47. Stone WJ, Eldering M, van Gemert GJ, Lanke KH, Grignard L, van de Vegte-Bolmer MG, et al. The relevance and applicability of oocyst prevalence as a read-out for mosquito feeding assays. Sci Rep-UK. 2013; 3.

48. Medica DL, Sinnis P. Quantitative dynamics of *Plasmodium yoelii* sporozoite transmission by infected anopheline mosquitoes. Infect Immun. 2005; 73(7):4363–4369. https://doi.org/10.1128/IAI.73.7.4363-4369.2005 PMID: 15972531

49. Annan Z, Durand P, Ayala FJ, Arnathau C, Awono-Ambene P, Simard F, et al. Population genetic structure of *Plasmodium falciparum* in the two main African vectors, *Anopheles gambiae* and *Anopheles funestus*. P Natl Acad Sci USA. 2007; 104(19):7987–7992 https://doi.org/10.1073/pnas.0702715104

50. Sandeu MM, Abate L, Tchioffo MT, Bayibéki AN, Awono-Ambéné PH, Nsango SE, et al. Impact of exposure to mosquito transmission-blocking antibodies on *Plasmodium falciparum* population genetic structure. Infect Genet Evol. 2016; 45: 138–144. https://doi.org/10.1016/j.meegid.2016.08.025 PMID: 27566334