Competition and immune selection of multi-strain Plasmodium falciparum malaria

Preliminary manuscript: 2/3/2019

David Gurarie (dxg5@case.edu)
Department of Mathematics, Applied Mathematics and Statistics
Center for Global Health and Diseases
Case Western Reserve University, Cleveland, OH 44106

Abstract

Malaria Plasmodium falciparum (Pf) species contains multiple parasite strains with different immunogenic profiles and expressed phenotypes. These strains interact (compete, cooperate) within host and in transmission (mosquito) environment, so host immunity and environment play important part in evolution and selection of parasite quasi-species. Conventional population-based models (e.g. Ross-MacDonald) have limited capacity to accommodate parasite diversity. An alternative Individual-based approach (IBM) offers greater flexibility.

Here we develop an IBM for multi-strain Pf malaria with genetically structured parasite and immune regulation. Parasite diversity is not constrained (could increase through recombination), yet the model allows efficient simulations of individual histories, as well as large communities and host ensembles over extended time period with multiple transmission cycles.

We employ our model to study competition and selection of mixed malaria strains in individual hosts and in host populations. The model offers different ways of quantifying *parasite fitness*, and selecting *most viable* types, under selective pressures. We explore the relationship between selected clones, in terms of their genetic makeup and associated fitness traits.

The model analysis revealed peculiar features of multi-strain malaria: winners of mixed competition in single-infection histories (individual hosts), may not succeed in host communities over multiple transmission cycles. So short-term *competitive fitness* does not guarantee strain survival. The latter in fact, was strongly associated with *cooperative behavior*, i.e. strain potential to co-exist and persist in different mixed combinations within-host. We examined network structure of such surviving *cooperative cliques*. These networks are found to have a few persistent *core nodes*, each one carrying a subordinate (transient) cluster. Our results shed new light on relative importance of *competitive* vs. *cooperative* behavior. Other potential applications of our model (future development) include spatially distributed host-mosquito systems, and putative vaccine that target dominant var-genes and persistent clusters.

Introduction

Multiple malaria strains circulate in host populations. Their makeup and population structure have important implications for parasite persistence, spread and control outcomes.

A successful parasite must succeed on two levels: (i) within-host, subject to immune pressure; (ii) in host communities and transmission environment (vector). In case of malaria, mosquito vector also serves to diversify parasite pool via mating and recombination. Strain diversity can contribute to parasite survival and adaption, particularly under stressful conditions.

There are different ways to define and quantify *strain diversity* (see e.g. [1]). Typically, it refers to internal stratification of a species into distinct subgroups (types), based on their *phenotypic expression* and/or underlying *genotype*.

Modeling approaches to host-parasite evolution (e.g. [2-5], [6], [7-9], [10]) often employ hypothetical 'fitness' traits, like virulence, transmissibility, persistence, and classify strains accordingly, e.g. 'virulent' vs. 'benign', 'drug- sensitive' vs. 'tolerant', etc. The underlying genetic makeup doesn't enter such models explicitly. However, the link between *genetic makeup* and *phenotypic expression*, is not straightforward, and often confounded by multiple factors, including host immune status, interacting parasite mixtures, timing of inoculae et al. In our approach we focus of asexual (haploid) parasite stages within-host, and define Pf *strains* (or clones) by their genetic makeup.

Host immunity is recognized as leading selective force, but combining it with population-level transmission can be challenging task, due to vastly different space /time scales, and limited capacity to accommodate parasite diversity and multiple pathways. Commonly used population models (e.g. [11],[12] [2-5], [6], [7-9], [10]) grossly oversimplify parasite makeup and immune control, which limits their scope for highly dynamic infection systems, like malaria.

On the opposite end are individual-based models with detailed resolution of in-host biology (see e.g. [13-17],[10, 18, 19]). But they are easily scalable to population level transmission.

Here we develop an agent-based model (IBM) for multi-strain Pf with sufficient biological details, yet computationally efficient, to allow large host ensembles and long history simulations. Our setup highlights genetically structured parasites, innate and specific immunity, transmissible blood stages (merozoite, gametocytes), mosquito uptake and recombination.

Specifically, a Pf clone is viewed as a collections of var-genes drawn from hypothetical 'natural pool'. Each var-gene in turn, has its specific growth/proliferation and immunogenic profile (stimulated antibodies). During asexual replication cycle parasites undergo antigenic variation by expressing different var-genes on the surface of infected RBC. Mixed strains interact within-host through shared resource (RBC), and via non-specific and specific (cross-reactive) immunity. Transmission by random mosquito biting allows parasite recombination and release of new clones.

A conceptual view on evolution in host-parasite systems, is often formulated via *fitness traits*, like virulence-transmissibility-persistence (V-T-P), whose trade-offs are believed to drive evolution (see e.g. [20-23]). Examples include 'fitness weights' (combination of reproductive / survival success) in the Fisher-Haldane-Wright selection theory (e.g. [24]), or 'fitness landscapes' [25].

In our setup 'fitness traits' can be assigned to parasite clones based on simulated infection histories. Among them are strain-specific *virulence* (RBC depletion), *infection duration* of detectable parasitemia, and most important for parasite survival, strain-specific *transmissibility* (infectivity) to mosquito. We employ the latter in our analysis of parasite selection, but we also show limitations of the conventional methodology. Indeed, 'fitness traits' identified in single-clone infections, may 'blend' or disappear in mixed infections with strong interaction. A case in point is 'virulence' measure by RBC depletion; here 'mixed infection' outcomes have no bearing on 'virulent profiles' of individual clones.

Multiple parasite strains 'compete' or 'cooperate' within-host and in host populations. We are interested in (short / long-term) outcomes of their interaction, its effect on parasite population structure and distribution in host communities, the makeup of 'best-fit' (most likely) selected clones, their immunogenic profiles and relationship.

Some recent papers studied makeup, diversity and organization of Pf clones from different perspectives. Paper [26] examined so-called *serum dominance* networks obtained from clinical isolates. They showed peculiar patterns and links among 'putative strains', but they didn't identify them explicitly. On the other side are detailed molecular analysis of var-gene samples collected from endemic areas (see e.g. [27-29]). They reveal high diversity of Pf clones, but also an organized structure. In all these works, and a subsequent model simulations [30], one observes highly non-random 'clustered' network structure of multiple clones, which bears qualitative resemblance to theoretical predictions of so-called 'strain theory' ([7-9, 31, 32]).

Our paper contributes to this line of work, and we add new perspective on immune-dominated in-host dynamics vs. 'mosquito transmission'; the role of 'competition' vs. 'cooperation'; var-gene selection, fitness traits.

There are several limitations of the current setup in terms of host population structure and parasite transmission. But we expect the essential qualitative conclusions should extend beyond our current setup. The future work will extend the scope of modeling and analysis to more realistic parasite-host systems and environments.

Materials and methods

Background

The key features of Pf asexual (merozoite) stage in human host are its multi-gene makeup, and immune evasion strategy via *antigenic variation* (AV) (see e.g. [32], [5, 18, 31, 33]). Each Pf parasite carries a family of var-antigens, expressed on the surface of infected red blood cells (iRBC), that play double role. On the one hand they expose parasite to host immunity by stimulating specific antibodies (Abs). On the other hand, they help it avoid immune clearing via cytoadherence in capillary tissues (a primary cause of Pf pathogenesis). Each Pf parasite carries 50-60 var genes, but only one of them is expressed at any given iRBC. During its 48-hour replication cycle a new generation of merozoites released from iRBC can switch their var-gene expression, and thus avoid immune clearing by accumulated Abs. The net outcome of AV are multiple waves of parasitemia, whereby a Pf-strain can extend its infective duration and increase transmission potential.

There are different ways to accommodate strain diversity and AV (see e.g. [5, 18, 31, 34], [13], [14], [18]). Our approach here follows recent papers [35], [30]), but many details of modeling setup and simulations differ.

In-host model

Multi-strain parasite. Our model highlights multi-variant parasite makeup. Each strain has several loci (1,2,…,m), filled with multiple var-gene alleles, drawn from large variant pool $N \gg m$ (see Figure 1). Each variant (var-gene) has associated growth/replication factor r (= mean number of viable merozoites released by a schizont expressing r), and immunogenic profile – specific antibody u_r. Growth factor (r) can serve as label of the var-gene (we follow the setup of [35]). So we define parasite strain $s = (r_1, r_2, ..., r_m)$, as combination of its var-gene (growth) labels. The order of varaints (r_i) in strain (s) is important, as our AV employs sequential switching pattern, starting with locus 1. All variants are assumed immunogenically distinct.

Thus each strain has its specific *immunogenic* and *growth* profiles: antibodies $\mathbf{u} = (u_1, ..., u_m)$, replication factors $\mathbf{r} = (r_1, ..., r_m)$. Different clones can cross-react via shared var-genes.

Agent (in-host) model consists of cell populations: (i) target RBC x, (ii) parasite load (iRBC) $y_s = \sum_r y_{s,r}$, strain *s* partitioned into fraction $y_{s,r}$ expressing variant *r*; (iii) gametocytemia G_s of

strain s, (iv) immune effector variables $\{u_r\}$, for variants r. In numeric simulations (below) we rescale cell populations (x, y, G), relative to the normal RBC count = $5 \cdot 10^6 / \mu L$.

The basic dynamic processes include (i) invasion and depletion of RBC *x* by released merozoite (ii) parasite replication/growth via RBC invasion; (iii) immune stimulation and clearing of parasites by innate and specific immune effectors; (iv) antigenic variation (switching rate, pattern).

The model is run in discrete time steps, given by replication cycle (48 hrs for Pf). On each time step a new generation of merozoites $\{r\, y_{s,r}\}$ is released from the iRBC pool $\{y_{s,r}\}$. They invade the available RBC *x* in proportion to their densities. Some of newly released merozoites can switch their var-gene expression. The AV switching process has two inputs, rate $\varepsilon = 10^{-4} - 10^{-3}$, and AV-pattern. Different switching patterns we proposed. We shall consider 3 of them based on loci ordering, (i) cyclic sequential (unidirectional) $1 \to 2 \to ... \to m \to 1$; (ii) symmetric sequential (bidirectional) $1 \leftrightarrow 2 \leftrightarrow ... \leftrightarrow m \leftrightarrow 1$; (iii) one-to-many (see [18]). In most simulations below we use sequential patterns (i) and (ii).

Immune stimulation and clearing. All iRBC expressing a particular variant *r* can stimulate its specific antibodies u_r, depending on the combined 'variant density' $Y_r = \sum_s y_{s,r}$. At the same time all of them are partially cleared by immune effector u_r, depending on its level and specificity (affinity). The surviving iRBC fractions proceed to the next cycle, by releasing new generation of merozoites, and switching their var-genes (fraction ε). The details of immune stimulation – clearing (functions and parameters) are explained in the appendix. Here we mention that immunity has finite memory, so IE variables $\{u_r\}$ decay gradually in time (c.f. [35], [36], [30]). This has implication in the long run, e.g. life-long infection histories.

Further details on model variables, parameters, inputs and outputs are outlined in the Appendix.

Computational setup. The model was run for relatively small var-gene makeup ($m = 5 - 10$), as opposed to real count, 50-60. Our primarily goal was qualitative analysis, so realistic m-values may not be essential.

The *effective growth rates* of all variants were taken in the range $4 \le r \le 8$. Once again it appears lower than the known Pf proliferation (mean merozoite release/schizont) numbers, 10-20. But our r_i have somewhat different meaning: they combine proliferation rate with probability of RBC invasion. So the

disparity of variant-specific growth factors can be attributed to different invasion potential, rather than mean fecundity.

Immune parameters (stimulation, clearing, waning rates) in our model were chosen so that parasitemia levels, and RBC depletion in simulated infection histories, are compatible with the know malaria therapy (MT) data (e.g. [37], [38]), though we made no attempt to calibrate our model with MT- data.

All dynamic simulations, data processing and analysis, were implemented and run on Wolfram Mathematica platform (version 11).

Results

Single and mixed infection histories in naïve host

A typical naive infection history of a single 5-variant clone is shown in Figure 2. The panels include (a) RBC depletion reaching its peak (virulence = 20%) by day 37; (b) iRBC (parasitemia waves) with 5 expressed variants (thin curves) taking over in sequential pattern; (c) stimulated specific immune effectors; (d) gametocytemia; (e) probability of mosquito infection P(t).

The latter, P(t) is computed from gametocytemia curve G(t) panel (d), via a threshold-type sigmoid function $P = \phi(G)$, that predicts probability of mosquito infection in term circulating gametocyte level (see e.g. [39-41] and Appendix).

Function P(t) gives rise to an important strain 'fitness output' called *transmission potential* (TP). The latter is defined by integrating P(t) over a prescribed time window (e.g. infection duration [0, tf]), so TP = AUC (area under curve) of function P(t). Panel (e) shows TP is approximately equal to 'host infectivity' times 'infective duration'.

Assuming uniform mosquito biting over time-range [0, tf], higher TP-value would imply higher likelihood of mosquito uptake and transmission, hence provide competitive advantage on population level.

We proceed to analyze naïve host histories, using different combinations of single and mixed-strain inoculae. Multiple factors can affect infection outcomes and estimated 'phenotype values': virulence (V), duration (D), TP. They include strain 'immune size' *d* = number of distinct var-genes ($d \le m$ - number of var loci), AV switching pattern, initial inoculum, and host immune status and competence.

For single-strain infections fitness traits (V, D, TP) are well defined, estimated from simulated histories in identical hosts. All 3 are correlated, higher V implies longer duration and higher TP. Initial dose has marginal effect, just shifting history by a few cycles. The AV-switching pattern had more pronounced effect (see Figure 2): symmetric sequential (ii) and one-to-many (iii) (not shown) give shorter D and lower TP, compared to direct sequential (i). Immune size *d* also matters, in general m-strain (number of var loci) of reduced diversity ($d < m$), behaves like d-strain.

The above 'phenotypic' strain classification breaks down for mixed infections, which exhibit more complicated dynamic pattern and multiple outcomes. Several factors can affect mixed infection histories, strains' makeup, immune cross-reactivity (var-gene overlap), sequential order of shared var-genes, time-lag between inoculae, et al.

We demonstrate it with a few examples of mixed infections and their TP-outcomes. Typically, mixed-strain competition would reduce individual TP (up to 50% or more), compared to uncontested (singe-strain) run (Figure 3, Table 1). But in some cases, we observe 'cooperative behavior', with some strains gaining TP, while others loosing (Appendix).

Immune-modulated parasite selection

Our main goal is evolution and selection of genetically structured malaria parasite in host populations. A realistic large scale simulation would require structured host community (demographic, geographic distribution) coupled to mosquito environment. A community of human agents is coupled to 'mosquito agents' by randomly biting and transmitting circulating clones from 'donor' to 'recipient' host. Transmitted clones would undergo immune-regulated within-host process, as described above. In addition, they could recombine to generate new clones. Two processes, immune mediated 'competition' in humans, and 'recombinant transmission' in mosquito, have opposite effect on parasite diversity. The former tends to reduce it via immune clearing, the latter can potentially increase it. The two driving forces would shape parasite population structure and select 'most viable' type, or viable 'cooperating cliques' as we shall demonstrate below. To implement such fully coupled human-mosquito evolutionary system is challenging task, that we shall address in the future work.

Our current analysis is done under simplifying assumptions: (i) human population made of naïve hosts, (ii) no parasite recombination in mosquito. So mosquito part is limited to direct transmission of 'infective' parasite clones from donor to recipient. Both assumptions could be approximately valid under special conditions. For instance, naïve hosts can dominate transmission cycle in populations with high

turnover rate relative to infection duration (so most of them experience a single bout of infection). Another example could be populations with the bulk of transmission carried by young children. Our reason for imposing the above assumptions is mostly technical (computational), rather than biological. The future work will extend the scope of modeling to more realistic host communities and environments.

The analysis of evolution and selection of multi-strain malaria will run in 2 steps: (i) *primary* selection in naïve host ensembles; (ii) *serial passage* (SP) in host lines and communities, over multiple transmission cycles.

Fitness selection in naïve host ensembles

We want to identify most-fit parasite clones that could dominate in different patterns of mixed-infection. The model employs a hypothetical ('natural') var-gene pool, made of N=21 growth-types uniformly distributed over the range ($4 \le r \le 8$). We use it to generate a random repertoire of 200 clones $\{s = (r_1, r_2, ..., r_5)\}$, each one carrying 5 variants (cf. [35]). The 200 selected clones will comprise the entire parasite repertoire, as no recombination are allowed in the current analysis. All var-genes $\{r_i : i = 0,1,...,20\}$ are assumed immunogenically disjunct (no cross-reaction), but clones sharing identical variants could interact via immune regulation.

To select 'most-viable' clones from the repertoire, we run multiple simulations of single and mixed-infections with different strain combinations. The key *selectable trait* will be transmission potential (TP), introduced in the previous section.

Each clone from the repertoire is subjected to a range of single/mixed contests (double, triple etc.), and we estimate the resulting TP scores. The number of all possible contests (doublets, triplets et al) of 200 grows rapidly with the contest size (= 2, 3, …). So we won't run all doublets or triplets, but limit our trial to representative 'fair contests', whereby each clone would compete against an equal number of randomly drawn contestants for each pool size (see Table 2).

TP-fitness is a well-defined trait for a single contest. But as we mentioned earlier, mixed-infections can give a range of TP-values, depending on the contest-pool makeup, host immune status, relative timing of inoculate et al (see Appendix). In such cases, we propose to define strain- TP by its mean value <TP > over all contests of a given type. Specifically, for each clone *s* we estimate its TP1-fitness ('single contest'), TP2 –mean TP in double contests, TP3 – mean TP in triple contests (see Table 2).

Such mean <TP> gives relative success rate of strain *s* against multiple competitors drawn from large hypothetical test pool, where different mixtures are present in roughly equal proportions. Frequency of mosquito uptake of strain *s* in such system is proportional to its mean-TP, a combination of (TP1, TP2, TP3, …) weighted by fraction in the contest pool. So statistical selection of mixed-contest outcomes can identify clones most-likely transmitted from *primary* (initial) host pool to the next generation.

For our statistical analysis we simulated 3 mixed contests (single, double, triple), to get for each clone its mean TP-values: TP1 (singlet), TP2 (8 doublets), TP3 (45 triplets). We comparted distributions of mean TP-values. The overall pattern was drop of <TP> with contest-size (single vs. double vs. triple). Figure 4 (a) shows distribution of TP_n for (n=1,2,3). A relatively narrow TP-range of singlets (natural TP-fitness), moves to lower and more broad range for doublets (about 15% drop), and further down for triples (20%). Panel (b) illustrates singlet-to-doublet loss by their scatter plot.

Each contest (TP1- TP2- TP3) would select its own best-fit quantile clones. Their analysis reveals highly non-random primary selection. Indeed, the overlap of the 80% best quantiles of (TP12-TP2-TP3) contained 13 common clones, well above the expected (random) value 1.6. Furthermore, the best 13 types have shown only minor drop of their TP-values between (TP1-TP2-TP3). It suggests 'highly transmissible' types can maintain their dominance and TP-value, in multi mixed-strain contests (Figure 4(a)). In contrast, the bottom 20% quantiles are nearly random (overlap =2), and exhibit marked drop of the TP, from natural 'singlet' values TP1 to 'doublet /triplet' (TP2-TP3) (Figure 4(a)).

The selected 80% -quantile overlap was also found to correlate with high *mean-growth* phenotype ($\langle r \rangle = \frac{1}{m}\sum_1^m r_i$), and high combined TP-score (TP1+ TP2+ TP3) (Appendix).

Besides clone selection, our mixed-contest runs have revealed highly non-random var-gene. To identify them, we assigned each variant *v* its frequency (multiplicity) in a given strain pool (# of clones containing *v)*. It differs from the usual allele frequency count based on host population fractions. The original random repertoire of 200 was unbiased, with approximately equi-distributed var-gene frequency (mean = 47.6). The primary selection however, produced strong bias towards high-growth variants (see Figure 6).

So mixed-contest immune competition produced highly non-random outcomes, both on the parasite level (selected clones), and in terms of their var-gene frequencies.

One way to think of our primary selection, is hypothetical naïve host population invaded by parasite infection, drawn large 'natural' repertoire, and organized into mixed-contest pools. Assuming uniform initial distribution of contest-pools, the output (transmissible clones) proved highly uneven, biased toward 'competitive types', which are more likely get transmitted to the next generation.

In the next section we turn to evolutionary dynamics of multi-strain malaria over multiple transmission cycles and generations. The *primary* selected *pool* will be used for initialization of evolutionary systems.

Serial-passage evolution in host lines and ensembles

A serial passage (SP) line follows a sequence of transmission cycles from donor to recipient. Each SP-agent runs its infection course from time 0 to d_P (passage day), then a collection of transmissible gametocytes (strain mixture) at d_P is transferred to the next in line. The passage day can be fixed, or vary randomly over a prescribed time window ($d_0 \le d_P \le d_2$). In our experiments passage day serves as proxy of transmission intensity (EIR), shorter d_P meaning higher EIR. In most experiments below we use random d_p .

Parasite mixtures undergo selective processes, as they follow through SP-lines and multiple transmission cycles, whereby successful types or groups take over. While individual SP-lines may differ, statically significant results arise on population level, i.e. ensembles of SP-lines run in parallel.

Evolutionary time scale used in our analysis, requires some clarification. Transmission cycles serve as discreet time units, as each line proceeds over generations (1st, 2nd …). Though different SP-lines of an ensemble may not be synchronous in real physical time (due to random d_P), a fixed generation of all SP-lines can be viewed as (approximate) 'instantaneous' snapshot of the entire community.

We are interested in statistical evolution of parasite population structure in such SP host ensembles. More specifically, we look at persistence and frequency distribution of single and mixed parasite clusters, the resulting relationship (network structure) among selected clones, their var- gene makeup and expressed phenotypes.

The analysis makes a few simplifying assumptions on transmission dynamics. We drop Intermediate parasite stages and processes in human and mosquito, so 'transmissible gametocyte pool' from donor turns into recipient 'merozoites/iRBC inoculum'. The net effect of omitted stages and processes are

time-lags in the onset of infection, which are negligible on evolutionary time scale (generations). As above, all hosts in SP-lines are assumed naïve.

We conducted two types of SP experiments using parasite repertoire of the previous section. In the first experiment, the 95% best quantile of primary selection, was used to initialize all SP- lines in the host ensemble. No recombination or external inputs were allowed, so initial mixture could only get depleted in such evolutionary process.

In the second experiment, we allowed an external input on each SP-step, drawn randomly from the complete parasite repertoire (an additional mosquito bite). This case allows parasite diversity to increase and sample more broadly the entire 200 clone repertoire.

In both cases we study long term evolution of parasite population structure, the resulting statistical and dynamic patterns, including stable dominant strains and mixtures, their genetic makeup, phenotypic profiles et al.

SP evolution for confined parasite pool.

The first set of SP-experiments employs 10 best-fit strains of the primary selection (Table 3) for initialization. Each SP-lines was run over multiple cycles, and we generated evolving host community of 200 SP-lines. To study the effect of transmission intensity we used 2 time windows for random passage day: (i) $d_P = 15-30$ (high intensity), (ii) $d_P = 25-40$ (low intensity).

A typical SP-line history with high intensity random d_P, is shown in Figure 7. On each cycle we measured strain PT (probability of mosquito transmission) at the passage day. It varied in the course of history, with most strains maintaining highest level 1, while other (#7) getting depleted after 30 cycle. Figure 8 shows evolution of two SP-communities with high and low transmission intensity. Each curve shows host population fraction, carrying a particular strain mixture. In both cases (i-ii) evolution drives such system to fixation. High intensity case (i) maintained elevated strain diversity (only one clone, #7 got depleted). Low intensity case (ii) splits into several small (4- and 5-strain) persisting clusters, with mean diversity dropping from 10 to 6.

Strain organization and connectivity. There different ways to describe strain patterns in terms of their structure and dynamics. A typical dynamic evolution produces long term persistent groups of 'cooperating' strains (see e.g. Figure 8). Structurally, they consist of overlapping clusters, each cluster in turn can be described by its immunogenic profile, represented by cross-reactivity (CR) matrix or graph.

We define CR-graph of strain complex $\{a,b,...\}$ (nodes), by linking clones that share common var-genes; each link $a \leftrightarrow b$ is weighted by the number (C_{ab}) of shared variants. Figure 9 illustrates the CR-network of 'best-fit' pool of Table 3. This network is tightly bound, with weights varying from 0 to 4 (e.g. $C_{1\leftrightarrow 9}=1$, $C_{2\leftrightarrow 10}=4$). V*ertex degree* C_a is weighted network is defined, as the sum of all link weights ($\sum_b C_{ab}$), connected to *a*. Mean cross-reactivity of cluster $\{a,b,...\}$ is defined as total CR ($\sum_a C_a$) divided by the cluster size. For the primary selection pool of Table 3, mean CR $=8$.

The loss of parasite diversity in SP-evolution is expected to lower its mean CR. Figure 8 (c) shows the CR-drop in both cases (i-ii). Such CR-loss is qualitatively consistent with theoretical predictions of strain theory, whereby immune selection should drive parasite distribution towards disjunct clusters (see e.g. ([7-9]).

More close analysis of the terminal evolution state (Figure 8(b)) shows two-level organization into persistent interconnected clusters. The three dominant clusters of Figure 8(b), CL1= (5, 6,8,10), CL2=(4,5,6,8,10), CL3=(6,8,9,10) share common core (6,8,10) (Figure 10) . Theirs CR-graphs (Figure 10 (b) show reduced connectivity, compared to strongly linked initial system (Figure 9).

Table 4 summarized mean CR-connectivity of core-strains (6,8,10) in different configurations: complete 6-node terminal network of Figure 10 (column "Total"), followed by 3 dominant clusters (CL1, CL2, CL3), and the overlap (core) triplet (6,8,10). We observe a consistent drop of mean CR. It suggests evolution of the initial 10-strain system towards lower CR-connectivity state, consistent with general principles of immune selection in mixed infections, and the strain theory (c.f. [18, 27, 30, 35]).

SP-evolution with external source.

Our second experiment involves the entire 200 repertoire of the previous section (primary selection). Here we use as an external source of inoculae, rather than initialization, which will play secondary role. So a serial passage will combine transmissible donor output, as before, supplemented with randomly drawn clone from the entire 200-repertoire. Such community can be viewed as subpopulation (e.g. children) submerged into larger transmission environment with multiple circulating strains.

As above our goal is evolution of parasite population structure in multiple SP-lines and ensembles. The analysis of dynamic patterns becomes more involved now, compared to the previous 'limited-pool' case, as multiple 'cooperative clusters' can arise and persist here.

We start with strain-diversity analysis, then move to identifying persistent 'cooperative cliques', and their structure: strain grouping CR-connectivity, var-gene makeup.

As above, we run simulations of SP-lines, over long evolutionary path (100 to 200 generations), with high (15-20) and low (25-40) transmission intensity. Figure 11 (a-b) shows typical SP-histories for high/low intensity cases. Multiple clones from the 200-pool enter these histories, but they exhibit vastly different duration patterns, from few cycle to the entire period. Their temporal distribution is strain-specific and highly non-random. The core-nodes of the previous selection (##6-8-10) persist throughout the entire period (they correspond to 3 long dark lines at the bottom of (a)). But other clones enter SP-lines by random source, and some of them persist over multiple cycles.

We want to examine statistically stable patterns, i.e. long persisting strains and groups, that have statistical significance in the SP-ensemble. Specifically, we looked at ensemble frequency above threshold (> 25%), and called them *cooperating cliques*. In the previous case, core nodes ##6-8-10 played this role.

Two cases, 'limited-pool' and 'open-source', exhibit markedly different dynamic patterns. The former had all SP-lines go to fixation. The latter can produce stable patterns only in a statistical sense, in terms of ensemble frequencies.

Thus one could ask for statistical fixation (population frequency) of individual clones. Figure 12 shows evolution and selection of high-frequency clones, defined by host count > 50/200 (>25%), at the late stage of evolution (t > 150). Altogether we got 14 'high frequency' clones, but only three of them (##196-198-200) came from original best-fit pool, used for initialization of SP-lines. As expected they were core triplet (labelled ##6-8-10) in the previous 'limited pool' case. The core triplet 6-8-10 was joined by 11 new clones coming from the open source.

Individual-clone selection of Figure 12 however, gives only partial view of the SP-evolution. Among other factors contributing to strain persistence is 'cooperative behavior'. Here we define persistent *cooperative clique* (singlet, doublet, triplet, etc.) by taking sufficiently long time-window $[t_0, t_1]$ of evolutionary history, and selecting persisting clusters within each SP- line.

In our experiment, we took an ensemble of 200 SP-lines, run over 200 cycles, with random passage date, $15 \le d_P \le 30$. The time window chosen as [50-150] passages, gave 196 persistent SP-lines with the following distribution of their multiplicities

1	2	3	4	5	6	7
25	54	51	38	23	4	1

i.e. 25 persistent singlets, 54 –doublets, etc. The most crowded clique had 7 cooperating clones.

Furthermore, close examination of multi-clone persistent cliques has shown many of them sharing the same limited pool of *core nodes* (Table 5). It appears successful clones are able to collaborate with multiple partners in different mixed combinations.

Figure 13 (a) shows the resulting CR-network of the core nodes , and panel (b) their population frequencies in the SP-ensemble, over time-window [50-150]. Some of them, ##32-68-172, attain high frequency values: (60-70) /200 host lines. It means large population fractions of the SP - community (30-35%) will carry 3 dominant clones over long periods of evolutionary history.

Overall, 11 cooperative clones exhibit relatively high CR-connectivity (Figure 13 (a)), and some peculiar patterns of their genetic makeup (Table 5). Namely, locus 1 is occupied by the high-growth vars (##18-20), while the rest fall into middle-low growth range. Individually, such clones would produce 'aggressive' initial parasitemia wave, and relatively high virulence (RBC-depletion)

We also observe best 'cooperative clones' of the open-source SP (Table 5) differ from the previous (limited-pool) selection (Table 3).

Each core-node of Table 5 carries associated high multiplicity (population frequency) *cliques*, made of doublets, triplets et al. The cliques demonstrate evolving stable *cooperative relationship* among mixed strains, similar to *stable clusters* in the previous (limited-pool) case. We show (Figure 15) persistent 5-cliques of the core-node (#184), and compare it to 'non-core' node (#9). In both cases, the resulting networks consist of several single-step trees, headed by #184 or #9, with linked branches (shared strains). The highly connected 'cooperative' node #184 has 9 branching clusters, interlinked by the fellow core members (##32,68,132, 124,158, 172) (Table 5), some of them linking multiple clusters. On the contrary, clusters of #9 are mostly disconnected, with a single link by the core node #32.

Thus our SP-evolution leads to formation of long-lasting network, made of a few dominant *cooperating nodes* with high CR-linkage, and multiple weekly-linked tree clusters attached to the dominant nodes. It qualitatively resembles observed serum-dominance networks [26], and numeric model simulations [27, 30].

Comments, conclusions

Immune selection and transmission intensity. The key drivers of parasite evolution, immune clearing and transmission intensity, have opposite effect. The former is expected to drive cross-reacting clones apart into disjunct clusters (host subpopulations), hence reduced CR (see e.g. [7-9, 18]). The latter should maintain them together, or expand (increased CR), via horizontal mosquito mixing. The net balance depends on relative strength of 'immune segregation' vs. 'transmission spread and mixing'. Two examples of SP-evolution for confined strain pool demonstrated the effect of such balance. High transmission rate limited immune effects by maintaining diverse strain pool and high CR-connectivity. Low intensity produced greater strain segregation, clustering with reduced diversity and lower CR. The immune effect was limited in the former case, due to shorter infection duration, while the latter allowed more persistent 'segregation'. We also observe organized network structures that arise under 'weak mixing' and 'strong segregation', dominated by few core nodes, and weakly linked clusters attached to the core.

The open-source SP allowed us to sample the complete strain repertoire. It produced merkedly different selection patterns, where 'cooperative' behavior proved more important, than 'competitive fitness'.

Model limitations. The present study was confined to naïve hosts. It includes individual histories, host ensembles, SP lines, et al. The 'naïve-host' assumption can be justified under special conditions, e.g. relatively short host turnover compared to infection duration (so each host can experience at most a single episode), or the bulk of transmission carried over by naïve hosts (children). Neither one is particularly realistic. Here we adopted 'naive-host' hypothesis for technical reasons, and the future work will extend the scope of model and analysis to fully coupled systems.

Future model development; control implications. The next step in model development and applications will include long term (life-long) infection histories for demographically structured host populations, and multiple transmission cycles. Furthermore, one can add parasite recombination in the system, during mosquito sexual reproduction stage, as well as ectopic (haploid) recombination within-human. In such systems parasite diversity is no constrained by any prescribed strain pool, but can evolve depending on transmission intensity, host makeup (immune status) and possible interventions (drug treatment, ITN). Next step would to set up spatially (geographically) distributed host communities, and mosquito environment. Similar to [42], one can explore links between mosquito distribution and transmission

intensity and parasite diversity (multiplicity of infection) for individual hosts, risk groups and communities.

Another possible application of multi-strain models is to track geographic spread of infection in distributed host communities, using clones as bio-markers.

There are many ways to expand our model and make it consistent with the available molecular datasets on var-gene makeup and diversity (e.g. [27]). Such calibrated (validated) multi-strain systems can have practical applications for malaria control. A case in point is putative anti-malaria vaccine. Var-genes don't look a suitable vaccine target due to their large potential diversity. However, if basic patterns of strain selection observed in analysis (few dominant clones), is confirmed for the real field data, one could ask for partially protective vaccine that targets selected clones.

Acknowledgement

The work was initiated during DG sabbatical leave at NIMBioS (University of Tennessee, Knoxville), which we thank for hospitality. We also thank LJ. Gross, V. Ganusov and S. Karl for productive discussions on topics related to the work.

Tables

Table 1: TP outcomes for a 3-strain system (left column) in 3 contests: (i) Single (uncontested) shows near identical fitness score, (ii) synchronous initialization maintains TP2 (strain 2), but significantly reduces TP1 and TP3; (iii) delayed inoculum of #3 (by 4 cycles) vs. ##1-2, has marginal effect on TP2, lowers TP1 and drops dramatically TP3.

	Single	Sync	Delayed
{10, 9, 19, 20, 1}	57.7	45.8	49.6
{7, 19, 0, 12, 14}	57.7	57.9	56.8
{3, 10, 16, 20, 9}	57.7	40.9	26.4

Table 2: Fair contests for mixed infections (single, double, triple) for random strain repertoire of 200. In mixed cases (double, triple) each strain 'plays' against equal number of its partners (8 in double contest, and 45 in triple contest)

	Single	Double	Triple
Total contests	200	800	3000
#contests/strain	1	8	45

Table 3: Selected 95% quantile of the TP1-TP2-TP3 primary contest

1	18	5	10	16	15
2	17	13	20	10	7
3	18	13	11	16	4
4	15	20	19	17	11
5	13	16	20	18	0
6	20	12	13	15	5
7	17	20	15	3	4
8	19	16	0	20	10
9	20	18	13	19	3
10	20	13	9	17	10

Table 4: Mean 'CR/node' for 3 core strains ##(6,8,10) in the complete residual complex ##(4,5,9,6,8,10) of Figure 10; 3 dominant clusters, and the core triplet (6,8,10)

	Total	CL3	CL2	CL1	*Core*
6	1.6	1.3	1.4	1.3	1.
8	1.6	1.3	1.6	1.5	1.
10	1.8	1.5	1.6	1.5	1.3

Table 5: 11 common-core clones shared by double-triple- quadruple,-quintuple cooperative clusters in simulated SP-ensemble with open source. In each row, number on the left is clone's label in the 200 repertoire, the remaining 5 digits give its var-gene makeup

32	20	13	9	17	10
44	19	13	1	10	2
68	19	16	0	20	10
105	20	10	1	15	9
124	18	14	4	1	8
132	19	2	8	11	3
158	19	1	17	0	10
172	20	12	13	15	5
179	20	2	5	18	7
180	19	12	1	16	10
184	20	0	8	15	10

Figures

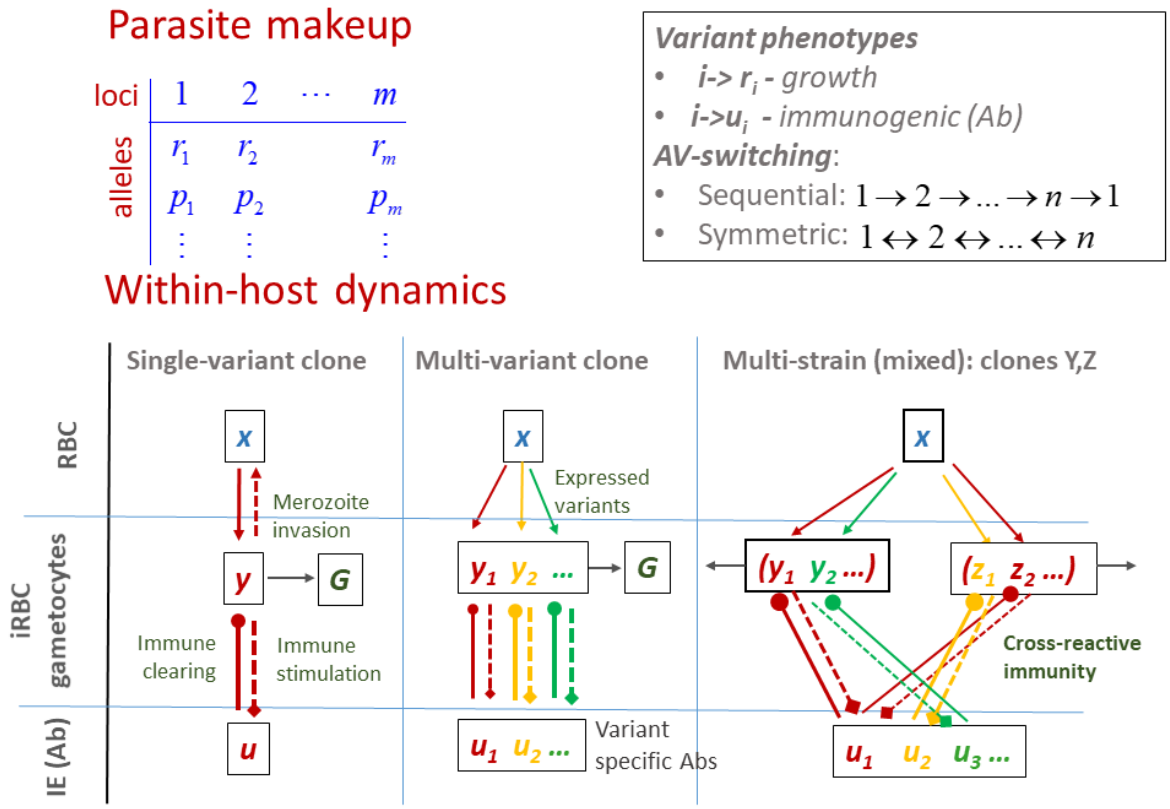


Figure 1: Schematic view of in-host model. Each strain is a combination of var-genes with specific growth and immunogenic profiles (antibodies), selected from large var-gene pool. The model features target cells (RBC), infected iRBC, that express different variants of a given strain, and stimulated specific immune effectors (IE). Shared variants (marked in different colors) create cross-reactivity patterns in mixed-strain infections. Different pattern of antigenic switching (AV) can be implemented in such setup.

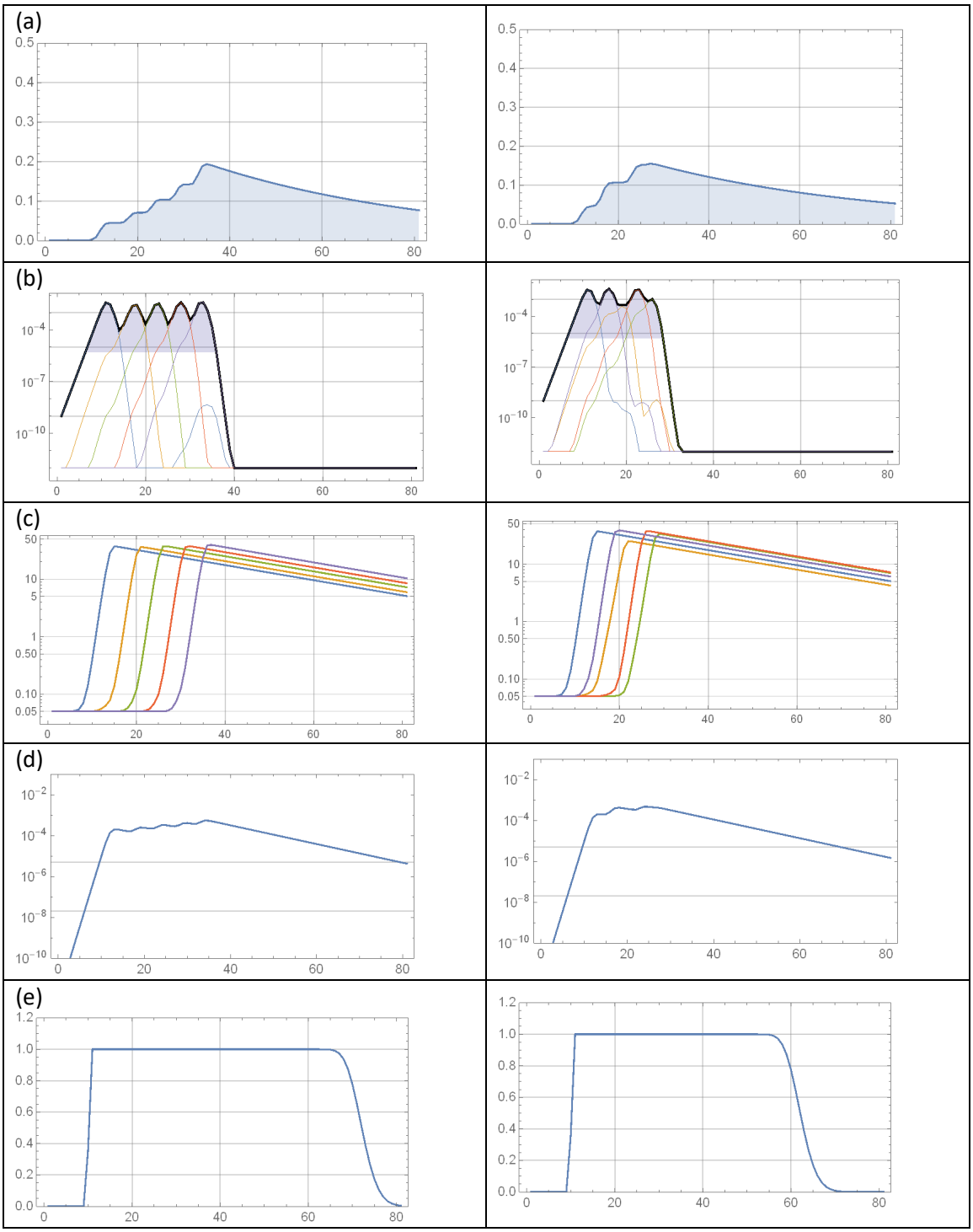


Figure 2: Naïve infection histories for 5V strain of clone $(8, 4, 9, 13, 22)$. *Two columns correspond to different AV switching patterns: (left) one-way sequential; (right) two-way sequential. Panel rows (top to bottom) represent a) RBC depletion; b) log-Parasitemia (iRBC) – total (solid black) and expressed variants (thin); c) variant-specific immune effectors; d) log-Gametocytemia; e) probability of mosquito infection. Two-way switching (right column) would shorten infection history, lower virulence (RBC depletion), and reduce transmissibility range.*

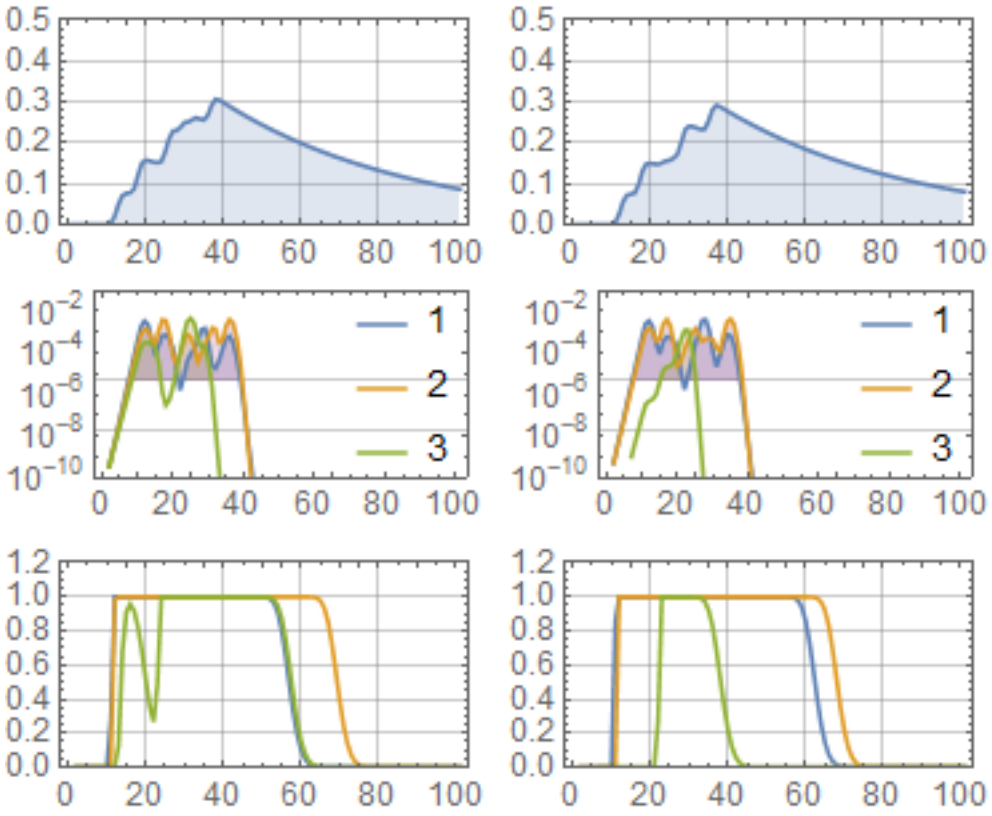


Figure 3: 3-strain contest for triplet of Table 1. Panels in each column show (i) RBC depletion, (ii) strain parasitemia with MS-detectable level marked by shading, (iii) probability of mosquito infection. Left column has all 3 strains initialized at t=0; right column has #3 initialized 4 time units after ##1-2. It results in a marked drop of mosquito infection curve and its cumulative TP for strain #3.

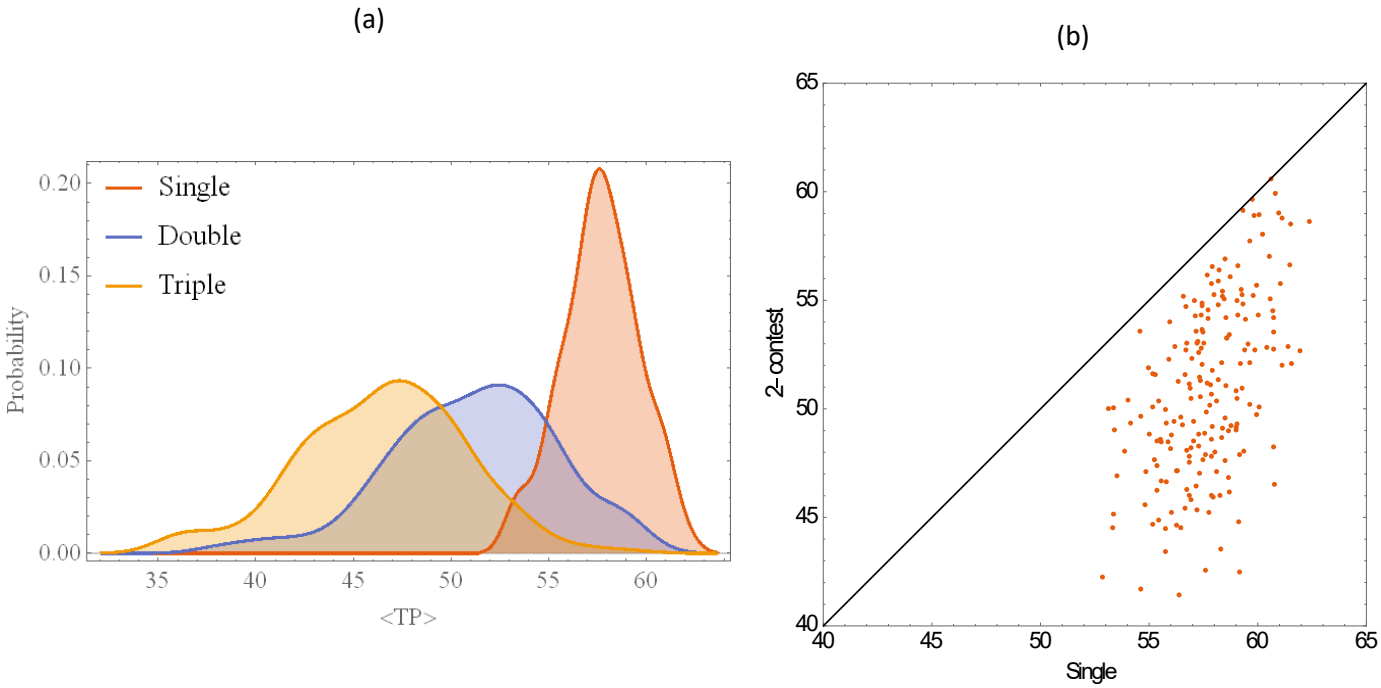


Figure 4: (a) Mean TP distribution for single, double, triple contest; (b) scatter plot of single TP vs. double-mean TP. The overall pattern is loss of TP-fitness in mixed competition

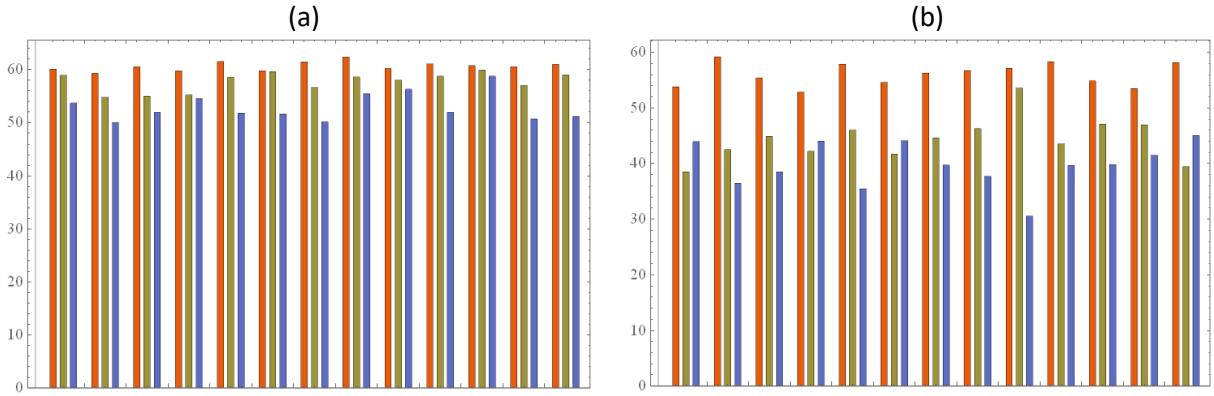


Figure 5: Mean TP drop for single (red), double (gray) triple (blue) contests. Panel (a) shows 13 best selected types (based on 80-th quantiles); panel (b) 13 lowest types.

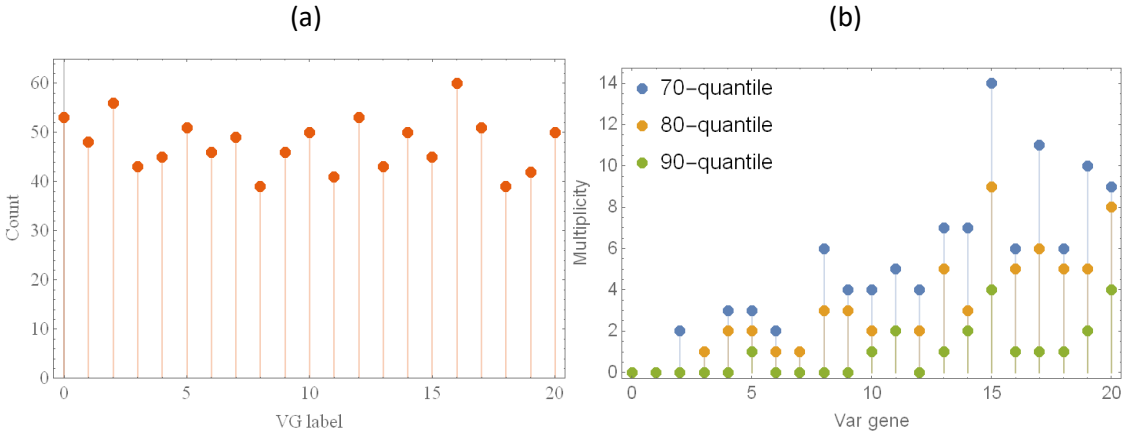


Figure 6: Var-gene multiplicity in the original random repertoire (a); and the best selected TP1-TP2-TP3 pool for high quantiles p=70-80-90% . High growth vars [10-20] dominate in contest selection.

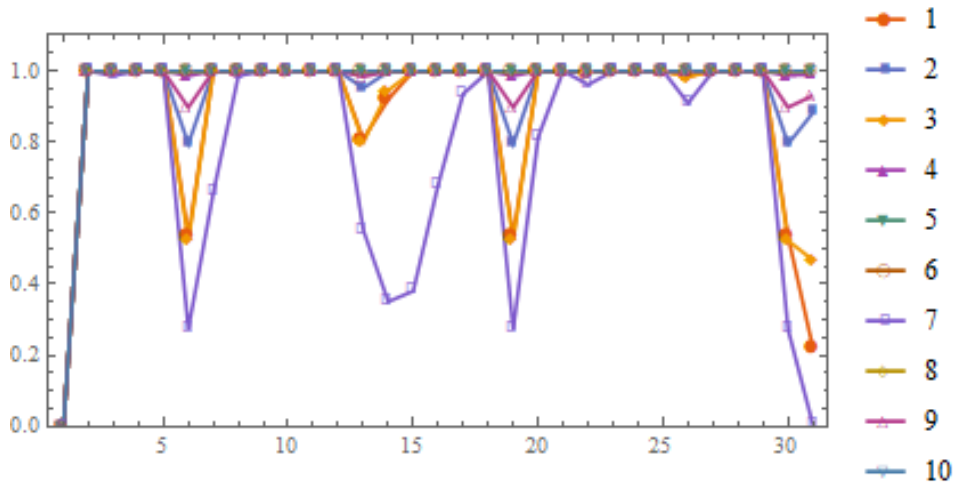
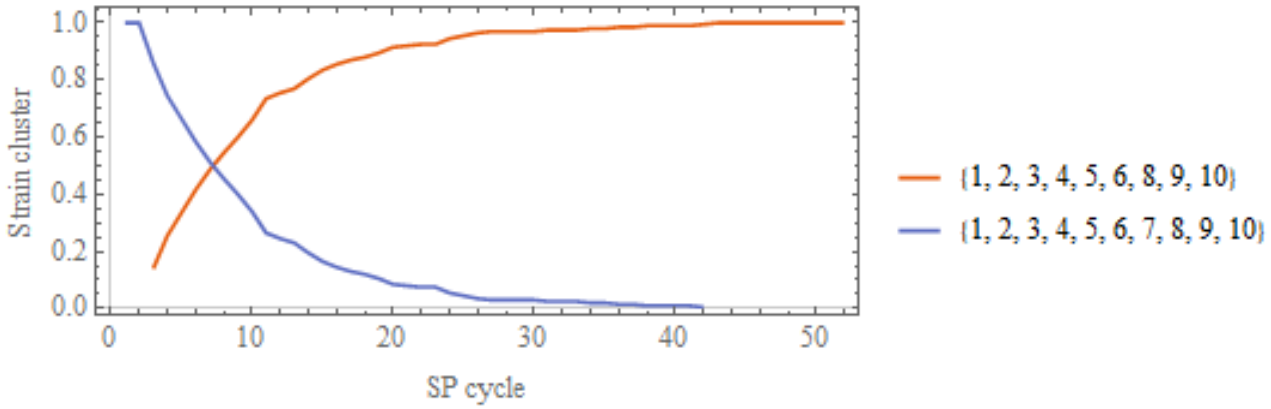
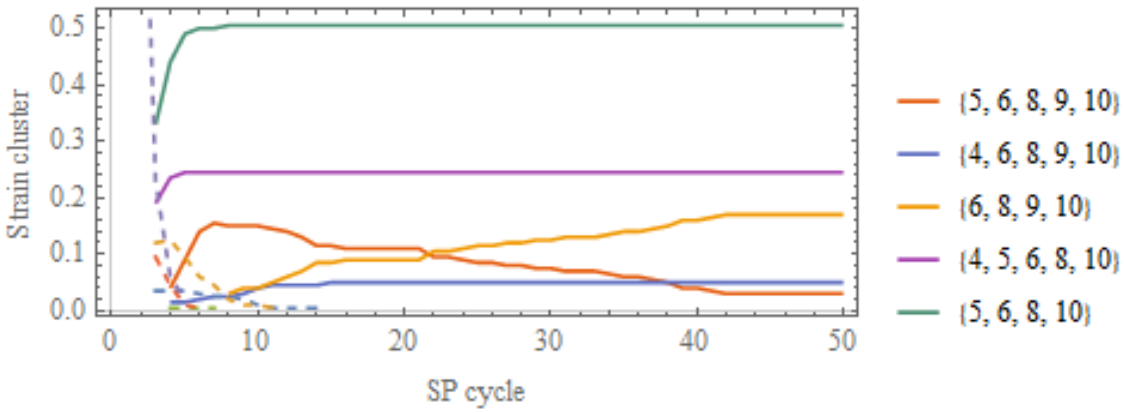


Figure 7: A typical SP-line of 10-starin system (Table 3) over 30 cycles, with random $d_P = 15 - 30$ *. On each cycle we measure strain PT (probability of mosquito transmission) by the end of cycle. Most strains persist over the entire span, but some (#7) get depleted by the end (30th cycle)*

(a)



(b)



(c)

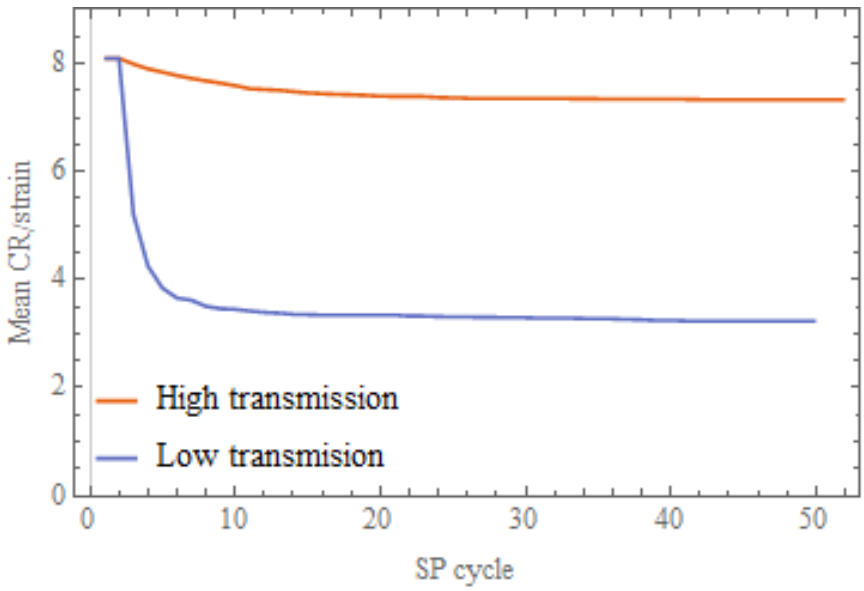


Figure 8: Long term evolution (frequencies of strain clusters) in SP-ensembles initialized with 10-strain mixture (Table 3) and different intensity of transmission: (a) $d_P = 15-30$ *; (b)* $d_P = 25-40$ *. Panel (c) shows evolution of mean CR/strain in 2 cases*

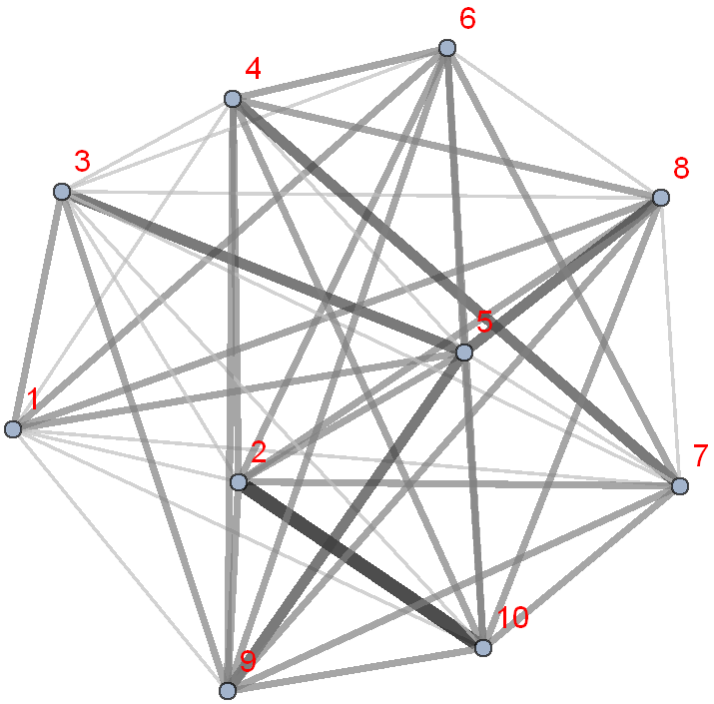


Figure 9: Tightly linked CR network of 10 selected clones (Table 3). Link weights (1,2,…) are marked by thickness and shade of gray, e.g. weight (1-9) =1, weight (2-10) = 4 (4 shared var genes)

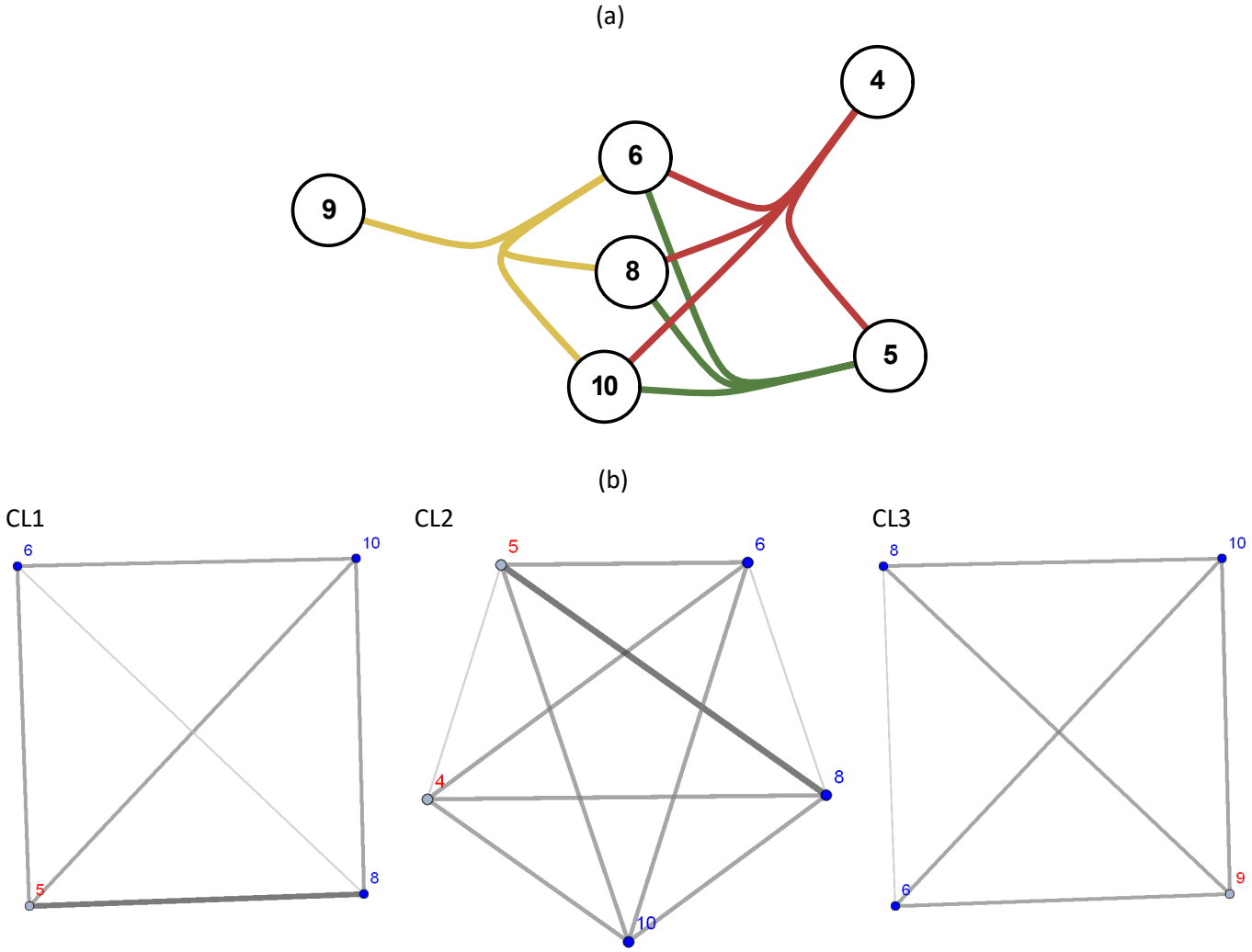
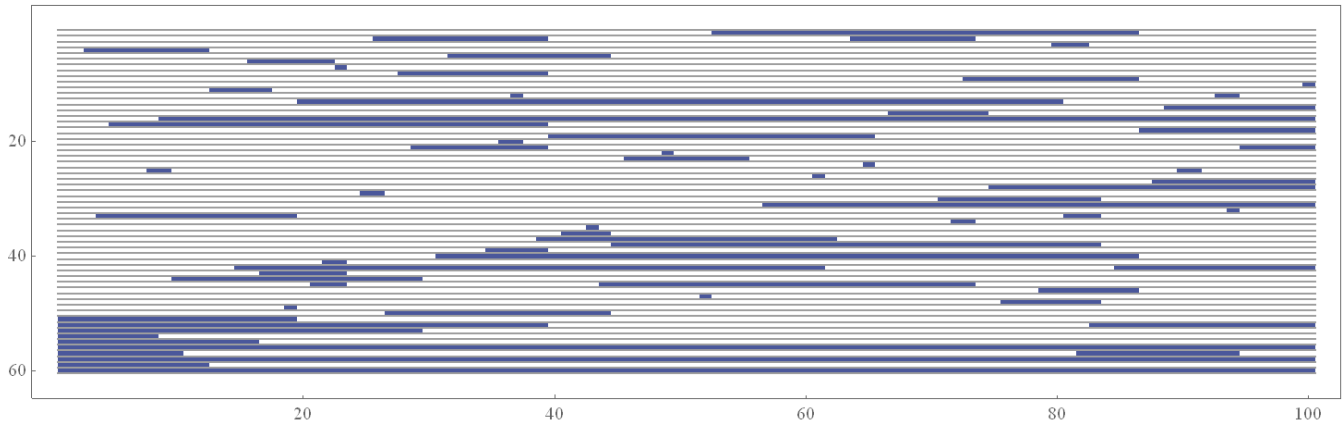
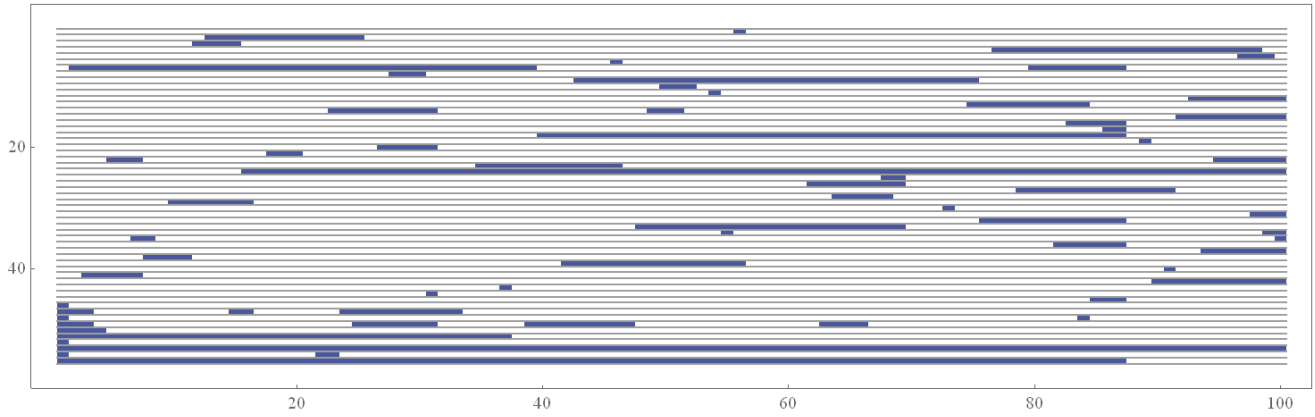


Figure 10: (a) Three dominant clusters of Figure 8(b): CL1= (5, 6,8,10), CL2=(4,5,6,8,10), CL3=(6,8,9,10) , share common core ##6,8,10. (b) their CR networks; the core nodes marked in blue

(a)



(b)



(c)

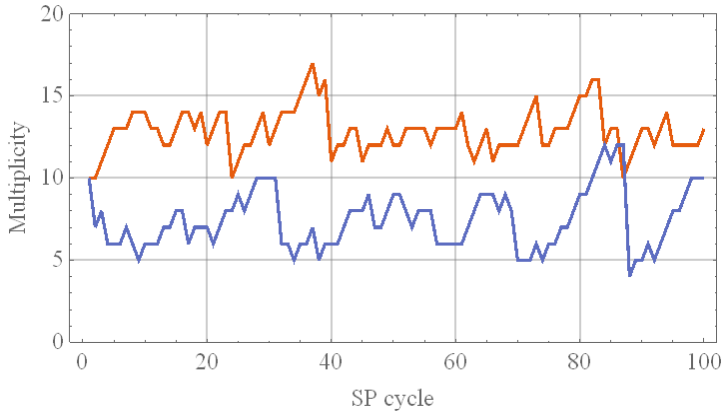


Figure 11: Typical SP –lines of 100 generations with random inoculae in 2 cases: (a) high intensity $d_P = 15-30$*, (b) low intensity* $d_P = 15-30$*. Each row marks a particular clone appearance and duration (shaded). History (a) contained 60 clones altogether, while (b) had 55 (reduced diversity). The clone ordering in (a)-(b) is not the original 200-repertoire, so identical numbers could be different types, except the low left corner (initial inoculum) made up of the original 10 best-fit selection of the previous section. Panel (c) shows multiplicity (strain diversity) of histories (a)-(b).*

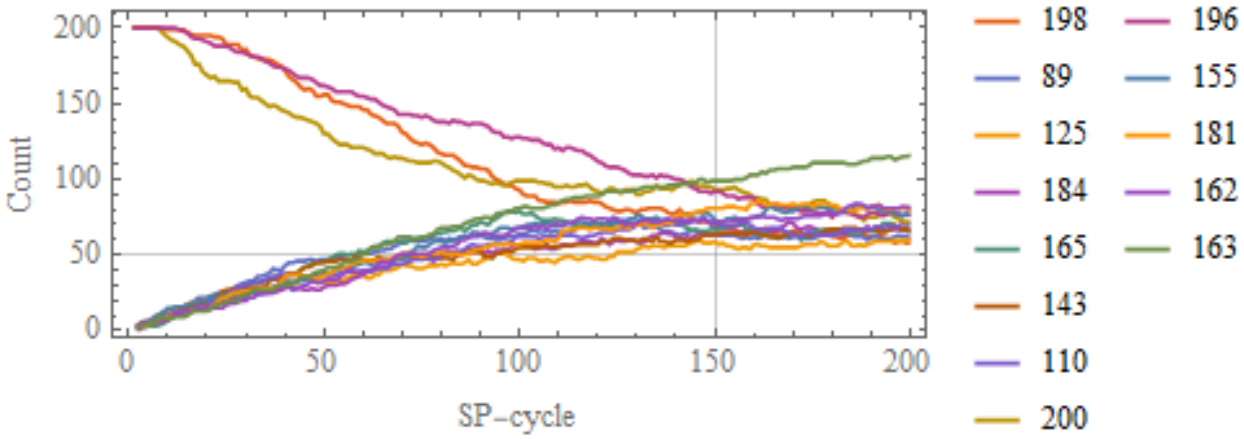


Figure 12: Evolutionary path of high-frequency clones of the SP-ensemble of 200 lines. Each curves shows population frequency of selected clones (number of SP-lines that carry the clone). Selection was based on frequency > 50/200, attained after 150 generations.

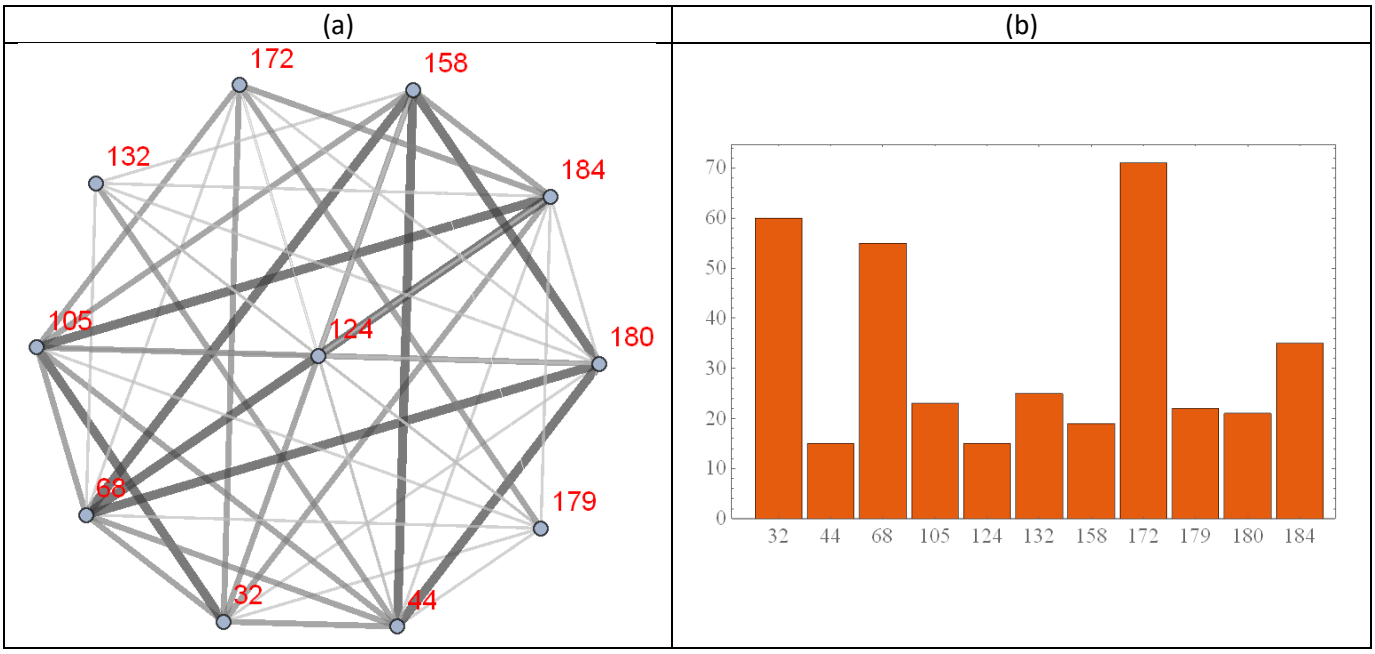
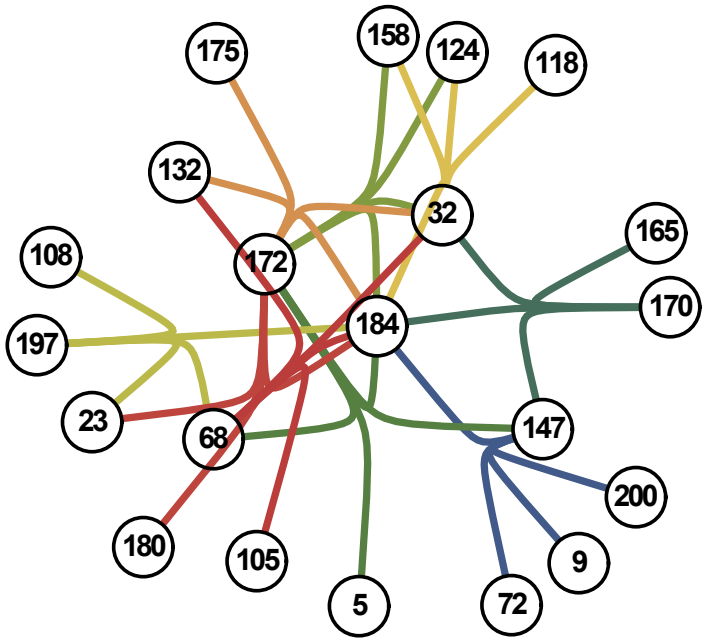


Figure 13: (a) CR network of the core-strain selection of Table 5; (b) their frequencies in the SP-ensemble of 200 lines over the time-range [50-150].

(a)



(b)

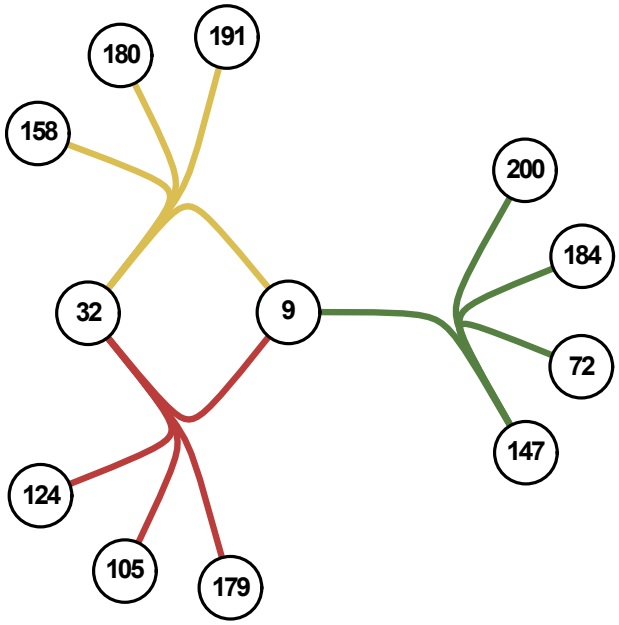


Figure 14: Persistent 5-cliques of the core node # 184 (a), vs. non-code #9 (b). In each case the graph is made of linked '4-trees' headed by #184 or #9.

References

[1] F. E. McKenzie, D. L. Smith, W. P. O'Meara, and E. M. Riley, "Strain theory of malaria: the first 50 years," *Advances in Parasitology,* vol. 66, pp. 1-46, 2008.

[2] V. V. Ganusov and R. Antia, "Trade-offs and the evolution of virulence of microparasites: do details matter?," *Theor Popul Biol,* vol. 64, pp. 211-20, Sep 2003.

[3] R. Antia, S. S. Pilyugin, and R. Ahmed, "Models of immune memory: on the role of cross-reactive stimulation, competition, and homeostasis in maintaining immune memory," *Proc Natl Acad Sci U S A,* vol. 95, pp. 14926-31, Dec 8 1998.

[4] R. Antia and M. Lipsitch, "Mathematical models of parasite responses to host immune defences," *Parasitology,* vol. 115 Suppl, pp. S155-67, 1997.

[5] R. Antia, M. A. Nowak, and R. M. Anderson, "Antigenic variation and the within-host dynamics of parasites," *Proc Natl Acad Sci U S A,* vol. 93, pp. 985-9, Feb 6 1996.

[6] M. A. Gilchrist and D. Coombs, "Evolution of virulence: interdependence, constraints, and selection using nested models," *Theor Popul Biol,* vol. 69, pp. 145-53, Mar 2006.

[7] M. Recker and S. Gupta, "A model for pathogen population structure with cross-protection depending on the extent of overlap in antigenic variant repertoires," *J Theor Biol,* vol. 232, pp. 363-73, Feb 7 2005.

[8] S. Gupta and R. Anderson, "Population structure of pathogens: the role of immune selection," *Parasitology Today,* vol. 15, pp. 497-501, 1999.

[9] S. Gupta, M. C. Maiden, I. M. Feavers, S. Nee, R. M. May, and R. M. Anderson, "The maintenance of strain structure in populations of recombining infectious agents," *Nature medicine,* vol. 2, p. 437, 1996.

[10] S. Frank and P. Schmid-Hempel, "Mechanisms of pathogenesis and the evolution of parasite virulence," *J Evol Biol,* vol. 21, pp. 396-404, Mar 2008.

[11] G. Macdonald, *The Epidemiology and Control of Malaria*: London Oxford University Press, 1957.

[12] N. J. T. Bailey, *The Biomathematics of Malaria*. London: Griffin, 1982.

[13] D. Gurarie, S. Karl, P. A. Zimmerman, C. H. King, T. G. St Pierre, and T. M. Davis, "Mathematical modeling of malaria infection with innate and adaptive immunity in individuals and agent-based communities," *PLoS One,* vol. 7, p. e34040, 2012.

[14] P. Eckhoff, "P. falciparum Infection Durations and Infectiousness Are Shaped by Antigenic Variation and Innate and Adaptive Host Immunity in a Mathematical Model," *PLoS One,* vol. 7, p. e44950, 2012.

[15] I. M. Hastings, S. Paget-McNicol, and A. Saul, "Can mutation and selection explain virulence in human P. falciparum infections?," *Malar J,* vol. 3, p. 2, Mar 2 2004.

[16] L. Molineaux and K. Dietz, "Review of intra-host models of malaria," *Parassitologia,* vol. 41, pp. 221-31, Sep 1999.

[17] F. E. McKenzie and W. H. Bossert, "An integrated model of Plasmodium falciparum dynamics," *J Theor Biol,* vol. 232, pp. 411-26, Feb 7 2005.

[18] M. Recker, C. O. Buckee, A. Serazin, S. Kyes, R. Pinches, Z. Christodoulou*, et al.*, "Antigenic Variation in Plasmodium falciparum Malaria Involves a Highly Structured Switching Pattern," *Plos Pathogens,* vol. 7, Mar 2011.

[19] K. Dietz, G. Raddatz, and L. Molineaux, "Mathematical model of the first wave of Plasmodium falciparum asexual parasitemia in non-immune and vaccinated individuals," *Am J Trop Med Hyg,* vol. 75, pp. 46-55, Aug 2006.

[20] T. Day, A. L. Graham, and A. F. Read, "Evolution of parasite virulence when host responses cause disease," *Proceedings of the Royal Society B: Biological Sciences,* vol. 274, pp. 2685-2692, 2007.

[21] J. C. de Roode, M. E. Helinski, M. A. Anwar, and A. F. Read, "Dynamics of multiple infection and within-host competition in genetically diverse malaria infections," *Am Nat,* vol. 166, pp. 531-42, Nov 2005.

[22] M. J. Mackinnon and A. F. Read, "Immunity promotes virulence evolution in a malaria model," *PLoS Biology,* vol. 2, p. e230, 2004.

[23] M. J. Mackinnon and A. F. Read, "Virulence in malaria: an evolutionary viewpoint," *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences,* vol. 359, pp. 965-986, 2004.

[24] A. W. F. Edwards, "Mathematizing Darwin," *Behavioral Ecology and Sociobiology,* vol. 65, pp. 421-430, March 01 2011.

[25] S. A. Kauffman, *The origins of order: Self-organization and selection in evolution*: OUP USA, 1993.

[26] C. O. Buckee, P. C. Bull, and S. Gupta, "Inferring malaria parasite population structure from serological networks," *Proceedings of the Royal Society of London B: Biological Sciences,* vol. 276, pp. 477-485, 2009.

[27] K. P. Day, Y. Artzy-Randrup, K. E. Tiedje, V. Rougeron, D. S. Chen, T. S. Rask*, et al.*, "Evidence of strain structure in Plasmodium falciparum var gene repertoires in children from Gabon, West Africa," *Proceedings of the National Academy of Sciences,* vol. 114, pp. E4103-E4111, 2017.

[28] Y. Artzy-Randrup, M. M. Rorick, K. Day, D. Chen, A. P. Dobson, and M. Pascual, "Population structuring of multi-copy, antigen-encoding genes in Plasmodium falciparum," *Elife,* vol. 1, p. e00093, 2012.

[29] P. C. Bull, C. O. Buckee, S. Kyes, M. M. Kortok, V. Thathy, B. Guyah*, et al.*, "Plasmodium falciparum antigenic variation. Mapping mosaic var gene sequences onto a network of shared, highly polymorphic sequence blocks," *Molecular Microbiology,* vol. 68, pp. 1519-1534, 2008.

[30] Q. He, S. Pilosof, K. E. Tiedje, S. Ruybal-Pesántez, Y. Artzy-Randrup, E. B. Baskerville*, et al.*, "Networks of genetic similarity reveal non-neutral processes shape strain structure in Plasmodium falciparum," *Nature Communications,* vol. 9, p. 1817, 2018/05/08 2018.

[31] M. Recker, S. Nee, P. C. Bull, S. Kinyanjui, K. Marsh, C. Newbold*, et al.*, "Transient cross-reactive immune responses can orchestrate antigenic variation in malaria," *Nature,* vol. 429, pp. 555-8, Jun 3 2004.

[32] S. A. Frank, *Immunology and evolution of infectious disease*: Princeton University Press, 2002.

[33] A. Scherf, J. J. Lopez-Rubio, and L. Riviere, "Antigenic variation in Plasmodium falciparum," *Annu Rev Microbiol,* vol. 62, pp. 445-70, 2008.

[34] M. Recker, S. Nee, P. Bull, S. Kinyanjui, K. Marsh, C. Newbold*, et al.*, "Transient cross-reactive immune responses can maintain antigenic variation in malaria," *Nature,* vol. 429, 2004.

[35] M. Severins, D. Klinkenberg, and H. Heesterbeek, "How selection forces dictate the variant surface antigens used by malaria parasites," *Journal of The Royal Society Interface,* 2011.

[36] M. Pinkevych, J. Petravic, K. Chelimo, J. W. Kazura, A. M. Moormann, and M. P. Davenport, "The dynamics of naturally acquired immunity to Plasmodium falciparum infection," *PLoS computational biology,* vol. 8, p. e1002729, 2012.

[37] W. E. Collins and G. M. Jeffery, "A retrospective examination of the patterns of recrudescence in patients infected with Plasmodium falciparum," *Am J Trop Med Hyg,* vol. 61, pp. 44-8, Jul 1999.

[38] L. Molineaux, H. H. Diebner, M. Eichner, W. E. Collins, G. M. Jeffery, and K. Dietz, "Plasmodium falciparum parasitaemia described by a new mathematical model," *Parasitology,* vol. 122, pp. 379-91, Apr 2001.

[39] T. S. Churcher, T. Bousema, M. Walker, C. Drakeley, P. Schneider, A. L. Ouédraogo*, et al.*, "Predicting mosquito infection from Plasmodium falciparum gametocyte density and estimating the reservoir of infection," *Elife,* vol. 2, p. e00626, 2013.

[40] P. Schneider, J. T. Bousema, L. C. Gouagna, S. Otieno, M. Van De Vegte-Bolmer, S. A. Omar*, et al.*, "Submicroscopic Plasmodium falciparum gametocyte densities frequently result in mosquito infection," *The American journal of tropical medicine and hygiene,* vol. 76, pp. 470-474, 2007.

[41] M. Eichner, H. H. Diebner, L. Molineaux, W. E. Collins, G. M. Jeffery, and K. Dietz, "Genesis, sequestration and survival of Plasmodium falciparum gametocytes: parameter estimates from fitting a model to malariatherapy data," *Trans R Soc Trop Med Hyg,* vol. 95, pp. 497-501, Sep-Oct 2001.

[42] S. Karl, M. T. White, G. J. Milne, D. Gurarie, S. I. Hay, A. E. Barry*, et al.*, "Spatial Effects on the Multiplicity of Plasmodium falciparum Infections," *PLoS One,* vol. 11, p. e0164054, 2016.