

Hyper-diverse antigenic variation and resilience to transmission-reducing intervention in falciparum malaria



Open Access This file is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. In the cases where the authors are anonymous, such as is the case for the reports of anonymous peer reviewers, author attribution should be to 'Anonymous Referee' followed by a clear attribution to the source work. The images or other third party material in this file are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

REVIEWER COMMENTS

Reviewer #1 (Remarks to the Author):

In their manuscript, 'Hyper-diverse antigenic variation and resilience to transmission-reducing intervention in falciparum malaria', the authors use agent-based modelling to assess how antigenic variation can cause rebounds in malaria transmission after a period during which malaria transmission has been controlled. The control tool considered here is indoor-residual spraying (IRS), which is a highly effective malaria control tool. However, it is relatively expensive and labour-intensive to apply. This means that, in regions where it has been widely used, it can be difficult to know when IRS can be withdrawn. For this reason, although the manuscript is highly technical, I believe that the findings presented here could be of interest to the malaria community. However, I have a number of questions about the details of the modelling, and the presentation of the results that I would like to see addressed before the article can be published. However, I stress that, in terms of the bigger picture, I am positive about the approach that the authors have taken in this piece of work.

- There are two different periods of IRS use that the authors describe: (i) A two-year period of IRS, after which IRS is withdrawn, and malaria rebounds; (ii) A decade-long period of IRS, where withdrawing IRS is not considered. I was a bit surprised to read that malaria rebounded in the latter case, where IRS was not withdrawn. On page 4, the authors explain that this is due to an increased duration of infection after IRS is introduced. I think this means that, due to decreased levels of immunity, it takes longer for humans to successively clear all of the var variants (or at least, all those not seen recently) present in an individual infection, is this right? I suspect that this increased duration of infection is heightened by not considering types of immune response that are not purely specific to individual var antigens. As the authors mention in the Discussion, immune responses that transcend var surface antigens could play a role. For example, looking at Figure 1 in Molineaux et al. (Parasitology 122, 379-391 [2001]), which shows data from malaria therapy patients, it is clear that parasite densities decrease with time since infection. This suggests to me that other types of immune response (var transcending, or cross-reactive responses) are playing an important role (this assumes that the successive peaks of parasitaemia are due to switching to var variants not yet seen in the infection). I do respect the approach that the authors have taken, and I'm not suggesting that extra work (e.g. sensitivity analyses) are necessary. But it seems clear to me that the rate at which malaria rebounds will be highly sensitive to changes in assumptions on the overall 'makeup' of the human immune response (how much of it is var variant specific? How much is cross-reactive? How much of it transcends specific variants?). I think this should be stated somewhere (perhaps in the penultimate paragraph of the Discussion, as this is where the authors mention the lack of VSA-transcending immunity) i.e. that the rebound rate (at a given intervention coverage) will be sensitive to assumptions about the make up of the immune response.

- The Agent-based model does not include drug treatment, as I understand it. How do the authors think the inclusion of drug treatment would change their results? I also think it's worth mentioning that the model doesn't include the type of age structure commonly seen in malaria transmission models (e.g. J. Griffin et al. Plos Med 2010), and agents don't appear to die and be replaced with infants with low/no immunity. This is not necessarily a problem, but this simplification should be clearly stated somewhere.

- The Shannon Diversity Index is mentioned several times. It would be useful to define this or, at the very least, provide a reference

- Figure 2B + 2D: some error bars seem to have gone missing here.

- Figure 3D: I'd recommend fixing the range on the y axis to be identical in each panel. This would make it clearer that the blue p.m.f. is (I think?) the same in each panel.

- Figure 6A: Can you add a legend for the colours here? The caption suggests that they might represent different values for the migration rates of genomes being introduced into the parasite population, but I wouldn't be 100% confident about this

- Figure 6: according to the caption, this figure shows results from scenarios in which IRS is applied for two years, and then withdrawn. Under each panel in part A is written '(Sustained IRS)'. Can this be removed? This suggests to me that the 10-year IRS campaign has been used in these instances.

- Supplementary Methods, page 2: In the first paragraph on this page, the authors discuss the question of the probability that a bite from an infectious mosquito results in an infected human host. The authors state that this parameter doesn't have to be explicitly stated, as it can be absorbed as a rescaling. Normally, such a rescaling involves rescaling time, is that the case here? This is fine, although it may have implications for the assumed duration of the IRS campaign, and the details used for malaria seasonality. I think this should be stated somewhere (if I am correct about this). It's fine to state in the Supplementary Methods, I appreciate that this is a pedantic point.

That concludes my report. I commend the authors for an interesting piece of work.

Reviewer #1 (Remarks on code availability):

Honestly, reading the 63 page document was enough for me. Plus, the code was supplied to me at a very late stage.

Reviewer #2 (Remarks to the Author):

Zhan et al. use an agent-based model to investigate how different and plausible epidemiological systems respond short and long-term to transmission-reducing interventions. The authors use the simulation framework to (i) characterise rebound probability and transition states and (ii) identify molecular markers of those states. Finally, the identified markers (or indicators) are evaluated against a longitudinal dataset from Ghana, where the population was sampled before, during, and after indoor residual spraying (IRS).

Measuring an intervention's effectiveness in near real-time is crucial to maximising the chance of success; thus, the manuscript has clear potential for tangible impact. Point in case, albeit through the lens of the proposed stochastic model, the authors present the case of an IRS intervention in Ghana close to a transition phase and where intensifying or sustaining the intervention could have yielded a more sustainable path to regional elimination. In general, this type of information is critical for data-driven decisions on how to tailor interventions, maximise their impact, and guarantee their long-term sustainability.

Overall, the methods and analysis proposed are solid and well substantiated, the paper is information-dense yet very clear and well-written, and the conclusions (specifically, the existence of a transition regime and the quantitative use of molecular indicators to measure its proximity) are reasonable within the context of the proposed framework. My comments are either minor or general.

Presentation

Could the display items be made more direct and the message sharper? Specifically, most figures are heavily multi-panelled (with insets within the panels), and it gets quite hard to recall which is which. I fully appreciate the intent of being transparent and comprehensive. However, in this case, the information overload might produce the opposite effect of burying meaningful knowledge in a sea of data. At the very least, the authors might consider whether all figures are necessary as main display items and possibly only maintain those that are critical for the narration. Even better would be to re-elaborate some of the items to make the message more straightforward. For example, in fig. 3, 4, and 5, the key message is the dynamic of the distributions (namely, how they change over time). If that is the case, then I wonder whether this change can be visualised directly (e.g. distance between distributions, change in the first/second momentum, etc.).

As an aside, I wonder whether it is possible to generate something akin to the phase transition diagrams commonly used in physics, i.e. representing for a given scenario and a set of molecular indicators readings (for the sake of argument median MOI and PTS), what state the system is likely to be in. Although this might not be visually possible, it leads to my next comment.

Applicability

The application of this framework to real data from Ghana is condensed into ~ten lines, and, in my view, it does not provide sufficient details to justify, other than at an intuitive level, the statement that the intervention has brought the population close to a transition state. While I appreciate that this work aims to explore the behaviour of the stochastic model, its ultimate validity and usefulness rely on the ability to interpret readings from real-world data (of course, with all the relevant caveats). While I am not necessarily advocating for a full-fetched decision framework (although an approximate Bayesian computation approach might be a potentially helpful future direction), at the very least, the authors should describe in more quantitative detail the process that led to the conclusion and/or provide some guidelines for interpretation: how much reduction in the PTS distribution is expected? How big of a shift in MOI distribution?

Molecular indicators

Do the authors expect the two molecular indicators to only be valid in a dynamic context (i.e. to measure relative changes), or could they be calibrated to a steady-state readout, e.g. to infer transmission rate in a population at a given time or to compare two distinct settings? It could be a point for discussion.

Related to that, I would imagine that the background genetic diversity of the population might play a role, but perhaps the var evolutionary rate is so high that it only captures very recent events and thus is only marginally affected by "ancient" migration and isolation?

Stochastic agent-based model

The model presented depends on a relatively large number of parameters, and the authors have done a thorough job of justifying their choices within the limits of current knowledge. It would be good to provide an intuition, if not actual data, of which parameters are more critical for the validity of the conclusions. By "critical," I specifically mean those instances where a relatively small change to the parameter would yield significantly different results and interpretations. The utility is two-fold. On the one hand, it can inform and steer future research (e.g. generate more data to estimate those parameters accurately). At the same time, it also allows users to assess the robustness of their hypothesis under a range of values.

Roberto Amato

Response to REVIEWER COMMENTS

(Our responses are in bold below).

Reviewer #1 (Remarks to the Author):

In their manuscript, 'Hyper-diverse antigenic variation and resilience to transmission-reducing intervention in falciparum malaria', the authors use agent-based modelling to assess how antigenic variation can cause rebounds in malaria transmission after a period during which malaria transmission has been controlled. The control tool considered here is indoor-residual spraying (IRS), which is a highly effective malaria control tool. However, it is relatively expensive and labour-intensive to apply. This means that, in regions where it has been widely used, it can be difficult to know when IRS can be withdrawn. For this reason, although the manuscript is highly technical, I believe that the findings presented here could be of interest to the malaria community. However, I have a number of questions about the details of the modelling, and the presentation of the results that I would like to see addressed before the article can be published. However, I stress that, in terms of the bigger picture, I am positive about the approach that the authors have taken in this piece of work.

R: We thank the referee for the positive evaluation of our work.

- There are two different periods of IRS use that the authors describe: (i) A two-year period of IRS, after which IRS is withdrawn, and malaria rebounds; (ii) A decade-long period of IRS, where withdrawing IRS is not considered. I was a bit surprised to read that malaria rebounded in the latter case, where IRS was not withdrawn. On page 4, the authors explain that this is due to an increased duration of infection after IRS is introduced. I think this means that, due to decreased levels of immunity, it takes longer for humans to successively clear all of the var variants (or at least, all those not seen recently) present in an individual infection, is this right? I suspect that this increased duration of infection is heightened by not considering types of immune response that are not purely specific to individual var antigens. As the authors mention in the Discussion, immune responses that transcend var surface antigens could play a role. For example, looking at Figure 1 in Molineaux et al. (Parasitology 122, 379-391 [2001]), which shows data from malaria therapy patients, it is clear that parasite densities decrease with time since infection. This suggests to me that other types of immune response (var transcending, or cross-reactive responses) are playing an important role (this assumes that the successive peaks of parasitaemia are due to switching to var variants not yet seen in the infection). I do respect the approach that the authors have taken, and I'm not suggesting that extra work (e.g. sensitivity analyses) are necessary. But it seems clear to me that the rate at which malaria rebounds will be highly sensitive to changes in assumptions on the overall 'makeup' of the human immune response (how much of it is var variant specific? How much is cross-reactive? How much of it transcends specific variants?). I think this should be stated somewhere (perhaps in the penultimate paragraph of the Discussion, as this is where the authors mention the lack

of VSA-transcending immunity) i.e. that the rebound rate (at a given intervention coverage) will be sensitive to assumptions about the make up of the immune response.

R: We thank the reviewer for this insightful comment. The reviewer is indeed correct on the reason why malaria rebounds even under sustained interventions when IRS is not discontinued. As the time into the intervention progresses, the immunity loss due to a reduction in transmission increases the susceptibility of individual human hosts and consequently, their infection duration. Because of these trends, although transmission remains at lower levels than those of pre-IRS, prevalence still rebounds. The rate at which prevalence rebounds depends on the rate of immunity loss. Different processes underlie the different types of immunity emphasized by the reviewer (i.e., variant-specific immunity, variant-transcending immunity or cross-reactive responses) and these can potentially exhibit different immunity loss rates. Therefore, the immune makeup of the human host (in terms of the relative strength of variant-specific immunity versus variant-transcending immunity) will impact the specific rate at which the rebound in prevalence happens. We added this point to the paragraph in the Discussion where we described the lack of variant-transcending immunity in the current implementation of our agent-based model and its potential extensions for future directions (lines 424-431).

As the referee rightly pointed out, the elegant model of Molineaux et al. (Parasitology 122, 379-391 [2001]) for the intra-host dynamics of parasitemia did make a case for a role of variant specific immunity in regulating the successive peaks of parasitemia and ultimately the duration of infection, although it also included effects of acquired strain-transcending immunity. Although strain-transcending immunity from exposure to various conserved antigens does reduce parasitemia and clinical symptoms (e.g., merozoite opsonization) (1), it does not necessarily prevent re-infection (2, 3, 4). In our study, because of our focus on a high-transmission endemic region, we considered the asymptomatic pool of infection which is known to contribute significantly to local transmission (5, 6). Even though the majority of hosts have gained generalized immunity, they still harbor chronic falciparum infections detectable as microscopic- or PCR-positive (2, 3, 4). Therefore, we assume in our model that the duration of infection is mainly determined by variant-specific immunity. We have highlighted this point in the same paragraph of the Discussion to help clarify the motivation and focus of our work (lines 432-438), while acknowledging this aspect of the model that would benefit from further consideration in future work.

- The Agent-based model does not include drug treatment, as I understand it. How do the authors think the inclusion of drug treatment would change their results? I also think it's worth mentioning that the model doesn't include the type of age structure commonly seen in malaria transmission models (e.g. J. Griffin et al. Plos Med 2010), and agents don't appear to die and be replaced with infants with low/no immunity. This is not necessarily a problem, but this simplification should be clearly stated somewhere.

R: We thank the reviewer for another insightful comment. Indeed, drug treatment can be common across varying transmission intensity settings. There are two main types of drug treatments to consider with different implications for our results. First, in the context of a local or regional intervention with IRS for example, individual human hosts can seek antimalarial curative treatments in response to their symptoms or their perception of transmission risk. Such curative drug treatments can potentially impact the infection status of treated individuals, as well as their MOI and pairwise-type sharing score or PTS between isolates. These treated individuals can be excluded from the analysis when calculating the distributions for these quantities, which would reduce sample size. Although the fraction of individuals who have pursued such treatment can be high, reaching ~25-50% across different age groups for example in our study site at the end of wet/high-transmission season for the pre-IRS phase (4, 7), the relatively large sample size would still leave us with enough individuals. During and immediately after the IRS intervention though, the fraction of individuals reporting they received an antimalarial drug treatment was found to be much lower, ~5-20% for our site (4, 7). We have now added an analysis that removes treated individuals whose status we know from the field questionnaire that accompanied the surveys (lines 276-283, and 727-741). We show that the same evaluation of the impact of the three-round transient IRS holds.

The second type concerns seasonal malaria chemoprevention (SMC), typically administered to all children between the ages of 3-59 months (i.e., < 5 years) to protect them mainly from mortality and severe disease manifestations. It cannot reduce however the overall transmission for the remaining population, which constitutes to the majority of the transmission reservoir. The overall population-wise prevalence throughout such intervention would remain high and largely unchanged (7), with no need to evaluate approaching transition. This kind of intervention was not considered here for our field study as it was rolled out later in the Upper East Region, including Bongo District, from 2016.

In our agent-based model, individual human hosts die and are replaced with infants with no immunity. The age structure of the human host population follows a truncated exponential distribution with a mean age of 30 years and a maximum age of 80 years. We apologize for not stating this information explicitly in the previous version of the manuscript and we added it to Methods (lines 504-506).

- The Shannon Diversity Index is mentioned several times. It would be useful to define this or, at the very least, provide a reference

R: We added two references, provided the definition and formula of the Shannon Diversity Index in Methods (lines 742-751), and described briefly why we reported both the richness, i.e., the number of circulating *var* genes, and this Index when quantifying systems *var* gene diversity.

- Figure 2B + 2D: some error bars seem to have gone missing here

R: For some IRS runs, the variation of the prevalence and the Shannon Diversity Index across replicates is small. We chose a large size for the points (the mean value), and the error bars appeared to have gone missing. We adjusted the format of the figure to better reflect the small variation across different replicates of those IRS runs.

- Figure 3D: I'd recommend fixing the range on the y axis to be identical in each panel. This would make it clearer that the blue p.m.f. is (I think?) the same in each panel.

R: The blue probability mass functions in each panel are from different replicate runs of the pre-IRS phase with the same parameter combination. These pre-IRS replicate runs were followed by the series of IRS interventions which reduced systems' transmission log-linearly along the gradient of transmission intensity. Therefore, the blue probability mass functions are not identical but very similar to each other. Following the reviewer's suggestion, we fixed the range on the y axis to be identical in each panel because it improves the visualization.

- Figure 6A: Can you add a legend for the colours here? The caption suggests that they might represent different values for the migration rates of genomes being introduced into the parasite population, but I wouldn't be 100% confident about this

R: Different panels of Figure 6A correspond to different simulated scenarios (seasonal or constant transmission, different spatial configurations); we added a title to each to make these scenarios clear. Different colors indicate the level of the series of transient IRS interventions which reduce system transmission log-linearly along the gradient of transmission intensity. We added a legend for the colors and a sentence clarifying their meaning in the caption. We also realized that the previous version of Figure 6A was too dense. Thus, we further simplified Figure 6A by moving the non-seasonal cases into the supplementary section (figs. S25).

- Figure 6: according to the caption, this figure shows results from scenarios in which IRS is applied for two years, and then withdrawn. Under each panel in part A is written '(Sustained IRS)'. Can this be removed? This suggests to me that the 10-year IRS campaign has been used in these instances.

R: The reviewer is indeed correct about these IRS interventions being transient and lasting only two years. We have fixed this error by replacing "(Sustained IRS)" with the correct notation "(Transient IRS)".

- Supplementary Methods, page 2: In the first paragraph on this page, the authors discuss the question of the probability that a bite from an infectious mosquito results in an infected human host. The authors state that this parameter doesn't have to be explicitly stated, as it can be absorbed as a rescaling. Normally, such a rescaling involves rescaling time, is that the case here? This is fine, although it may have implications for the assumed duration of the IRS campaign, and the details used for

malaria seasonality. I think this should be stated somewhere (if I am correct about this). It's fine to state in the Supplementary Methods, I appreciate that this is a pedantic point.

R: We thank the reviewer for the thoughtful comment. Force of infection (FOI), the number of new infections acquired by an individual human host over a given time interval (or the number of effective infections which advance to the blood stage), defines the time scale of malaria transmission. There are other surrogate variables for transmission intensity used in field studies, including the entomological inoculation rate or EIR, defined as the number of infectious bites received by an individual human host over a given time period. Because FOI refers directly to detectable blood-stage infections whereas EIR concerns human-vector contact rates, the difference between these quantities is mediated primarily by immunity and within-host dynamics, as well as measurement bias. The probability of transmission from an infectious mosquito bite is commonly used in mathematical models to mediate the difference between these two transmission quantities, as a general parameter encapsulating these processes. Our effective contact rate is related to EIR (even though we do not explicitly model the mosquito population), but it includes the probability implicitly. Thus, we would need to increase the value of the contact rate to achieve comparable values of FOI if we were to incorporate a transmission probability factor explicitly (i.e., the probability that a bite from an infectious mosquito results in an infected human host, noted by the reviewer), and the amount of this increment would depend on the specific choice of values for this factor. The goal is to match the emergent FOI values in the simulations to the reported ones in the field studies. For our simulated pre-IRS dynamics, annual FOI values are within the range of 9-15, directly comparable to the directly measured and indirectly estimated field values for high-transmission endemic regions in sub-Saharan Africa. We have stated this matter in more detail in the Methods (lines 1299-1306).

That concludes my report. I commend the authors for an interesting piece of work.

R: Thank you again.

Reviewer #1 (Remarks on code availability):

Honestly, reading the 63 page document was enough for me. Plus, the code was supplied to me at a very late stage.

Reviewer #2 (Remarks to the Author):

Zhan et al. use an agent-based model to investigate how different and plausible epidemiological systems respond short and long-term to transmission-reducing interventions. The authors use the simulation framework to (i) characterise rebound probability and transition states and (ii) identify molecular markers of those states.

Finally, the identified markers (or indicators) are evaluated against a longitudinal dataset from Ghana, where the population was sampled before, during, and after indoor residual spraying (IRS).

Measuring an intervention's effectiveness in near real-time is crucial to maximising the chance of success; thus, the manuscript has clear potential for tangible impact. Point in case, albeit through the lens of the proposed stochastic model, the authors present the case of an IRS intervention in Ghana close to a transition phase and where intensifying or sustaining the intervention could have yielded a more sustainable path to regional elimination. In general, this type of information is critical for data-driven decisions on how to tailor interventions, maximise their impact, and guarantee their long-term sustainability.

Overall, the methods and analysis proposed are solid and well substantiated, the paper is information-dense yet very clear and well-written, and the conclusions (specifically, the existence of a transition regime and the quantitative use of molecular indicators to measure its proximity) are reasonable within the context of the proposed framework. My comments are either minor or general.

R: Thank you for the appreciation of the work and the constructive comments.

Presentation

Could the display items be made more direct and the message sharper? Specifically, most figures are heavily multi-panelled (with insets within the panels), and it gets quite hard to recall which is which. I fully appreciate the intent of being transparent and comprehensive. However, in this case, the information overload might produce the opposite effect of burying meaningful knowledge in a sea of data. At the very least, the authors might consider whether all figures are necessary as main display items and possibly only maintain those that are critical for the narration. Even better would be to re-elaborate some of the items to make the message more straightforward. For example, in fig. 3, 4, and 5, the key message is the dynamic of the distributions (namely, how they change over time). If that is the case, then I wonder whether this change can be visualised directly (e.g. distance between distributions, change in the first/second momentum, etc.).

R: We simplified some figures in the main document and some supplementary figures by reducing the number of panels, and only displaying the critical ones for the narration.

As an aside, I wonder whether it is possible to generate something akin to the phase transition diagrams commonly used in physics, i.e. representing for a given scenario and a set of molecular indicators readings (for the sake of argument median MOI and PTS), what state the system is likely to be in. Although this might not be visually possible, it leads to my next comment.

R: We thank the reviewer for this insightful comment. We addressed it together with the next comment, as the two are related.

Applicability

The application of this framework to real data from Ghana is condensed into ~ten lines, and, in my view, it does not provide sufficient details to justify, other than at an intuitive level, the statement that the intervention has brought the population close to a transition state. While I appreciate that this work aims to explore the behaviour of the stochastic model, its ultimate validity and usefulness rely on the ability to interpret readings from real-world data (of course, with all the relevant caveats). While I am not necessarily advocating for a full-fetched decision framework (although an approximate Bayesian computation approach might be a potentially helpful future direction), at the very least, the authors should describe in more quantitative detail the process that led to the conclusion and/or provide some guidelines for interpretation: how much reduction in the PTS distribution is expected? How big of a shift in MOI distribution?

R: Prompted by this valuable comment, we investigated a more quantitative approach to evaluating the comparison of PTS distributions across the age group of children (1-10 years old) and that of adults (>20 years old), we performed additional analysis by calculating the difference in the corresponding PTS values of the two age groups at consecutive quantiles within the range of 0-0.40 for both the simulated output and the Ghana surveys. We chose this range because it encompasses the lower end of PTS values whose change reflects the rate of outcrossing. The difference at those quantiles exhibits two distinct patterns before and close to the transition regime in the simulated output. We also considered a more quantitative assessment of the MOI distribution, with a threshold value for the fraction of the distribution with 1 or 2 infections, that is necessary but not sufficient. Taken together with the PTS indicator, we can assess the change. On this basis, the pattern for the Ghana data supports our previous assessment that the system was pushed into approaching the transition regime. We included this additional analysis in the Results (lines 256-283).

Molecular indicators

Do the authors expect the two molecular indicators to only be valid in a dynamic context (i.e. to measure relative changes), or could they be calibrated to a steady-state readout, e.g. to infer transmission rate in a population at a given time or to compare two distinct settings? It could be a point for discussion.

R: We thank the reviewer for this insightful comment. The two molecular indicators are qualitatively valid for steady-state readout. The age-dependent pattern of PTS distribution results from immune selection (3, 8). In high-transmission endemic regions, children have been exposed to fewer infections than adults, and their immunity level against the circulating diversity is lower than that of adults. In other words, children's susceptibility against the circulating diversity is higher than that of adults. In general, any circulating *var* gene can infect more children than adults. Or it can be found in more children than adults.

There is thus a higher overlap in *var* genes of parasites that infect children than the ones that infect adults. The PTS of parasites infecting adults exhibit a mode at around 0, much more so than that of the ones infecting children. In contrast, in low-transmission regions, all age groups experience low exposure with either low or high diversity (9). Different age groups are similarly susceptible to the circulating diversity. Here, many persons may reach adult age without having built protective immunity and are thus susceptible to the disease, including severe and fatal illness (10, 11). The signature of immune selection is much weaker and correspondingly the difference in the PTS distributions across the two age groups is smaller or does not exist. In general, the age-dependent pattern in PTS is associated with transmission intensity, which is a proxy for the intensity of immune selection. Similarly, changes in the MOI distribution have been identified as an indicator of parasite transmission, including those of the mean or the proportion of monoclonal infections.

However, we acknowledge that the sensitivity of the two molecular indicators when used for calibrating a steady-state readout and inferring transmission rate remains to be investigated. On-going work focuses on this future direction. We are using quantitative information on the population-wise PTS distribution, the difference in the PTS distributions across age groups, and the MOI distribution for calibration purposes, including inference of the transmission rate for both stationary and transient stages, i.e., pre-IRS intervention and during IRS intervention. We have added a paragraph to the Discussion on this matter (lines 462-477).

Related to that, I would imagine that the background genetic diversity of the population might play a role, but perhaps the *var* evolutionary rate is so high that it only captures very recent events and thus is only marginally affected by "ancient" migration and isolation?

R: Another thoughtful comment. Although the *var* evolutionary rate is extremely high, we note that the fixation probability of new genes generated from ectopic recombination and mutation (i.e., the local innovation events) is much lower than that of the new genes gained from migration. The difference between the two probabilities is similar to the contrast between hard versus soft sweeps in population genetics (12). "Hard sweeps" refer to the scenario when only a single copy carrying the adaptive mutation rises to high frequency, whereas "soft sweeps" refer to the case of multiple copies of the same adaptive allele rising to high frequency in the population. In the context of malaria transmission, local innovation events give rise to single copies of new *var* genes, which have a low probability of surviving the genetic drift barrier and would mostly go extinct. In contrast, migrant genes can get introduced to the local parasite population repeatedly, and are more likely to survive the genetic drift barrier and rise to higher frequencies. In addition, negative frequency-dependent selection (NFDS) as a form of balancing selection, can maintain genetic variation for longer than expected by random chance (13), including the variation introduced to the population by migration a long time ago. Thus, the overall diversity represents a

long-term malaria burden for different regions. Therefore, *var* genomic data should contain signatures of both local evolutionary events and recent or more ancient migration events. We added a short paragraph on this matter in the Discussion (lines 411-420).

Stochastic agent-based model

The model presented depends on a relatively large number of parameters, and the authors have done a thorough job of justifying their choices within the limits of current knowledge. It would be good to provide an intuition, if not actual data, of which parameters are more critical for the validity of the conclusions. By "critical," I specifically mean those instances where a relatively small change to the parameter would yield significantly different results and interpretations. The utility is two-fold. On the one hand, it can inform and steer future research (e.g. generate more data to estimate those parameters accurately). At the same time, it also allows users to assess the robustness of their hypothesis under a range of values.

R: We thank the reviewer for this valuable comment. We conducted test runs to determine which subset of parameters the population dynamics of our stochastic ABM model are more sensitive to. We provided our choices of those parameters in both the main and supplementary Methods in detail (lines 598-702, 1299-1301, 1380-1440). Parameters pertaining to the specification of the regional pool for regionally-open systems are key to population dynamics. We circumvent the sensitivity question by conducting simulation runs with a wide range of these parameters, which cover from the extreme scenario of regionally-open systems composed of a few local parasite populations, to the one in which they are composed of more than ten such populations. Our results hold for the two extremes. Transmission intensity is key to population dynamics, but our results are meant to guide control and elimination efforts in high-transmission endemic regions. We conduct simulation runs with transmission intensity and an emergent force of infection within the range defined by values measured directly in the field or estimated indirectly for empirical high-transmission endemic regions from sub-Saharan Africa. Another key parameter is the immunity loss rate. Our results hold for the estimated value based on the historical datasets of treatment of neurosyphilis patients with malaria infections. We have added a summary paragraph on this matter to the Discussion to make the message clearer (lines 448-461).

Roberto Amato

References and Notes

1. Hill, D. L. et al. Merozoite antigens of *Plasmodium falciparum* elicit strain-transcending opsonizing immunity. *Infection and immunity*, 84(8), 2175-2184 (2016).
2. Tiedje, K. E. et al. Seasonal Variation in the Epidemiology of Asymptomatic *Plasmodium falciparum* Infections across Two Catchment Areas in Bongo District, Ghana. *Am J Trop Med Hyg*, **97**(1), 199-212 (2017).

3. Ruybal-Pesántez, S. et al. Age-specific patterns of DBL α var diversity can explain why residents of high malaria transmission areas remain susceptible to *Plasmodium falciparum* blood stage infection throughout life. *International Journal for Parasitology*, **52**(11), 721-731 (2022).
4. Tiedje, K. E. et al. Indoor residual spraying with a non-pyrethroid insecticide reduces the reservoir of *Plasmodium falciparum* in a high-transmission area in northern Ghana. *PLOS Glob Public Health* **2**(5) (2022).
5. Andolina, C. et al. Sources of persistent malaria transmission in a setting with effective malaria control in eastern Uganda: a longitudinal, observational cohort study. *The Lancet. Infectious diseases*, **21**, 1568–1578 (2021).
6. Lindblade, K. A., Steinhardt, L., Samuels, A., Kachur, S. P. & Slutsker, L. The silent threat: asymptomatic parasitemia and malaria transmission. *Expert Rev Anti Infect Ther.*, **11**(6), 623-639 (2013).
7. Tiedje, K. E. et al. Measuring changes in *Plasmodium falciparum* census population size in response to sequential malaria control interventions. *eLife*, **12** (2023).
8. He, Q. et al. Networks of genetic similarity reveal non-neutral processes shape strain structure in *Plasmodium falciparum*. *Nat. Commun.*, **9**, 1817 (2018).
9. Ruybal-Pesántez, S. et al. Population genomics of virulence genes of *Plasmodium falciparum* in clinical isolates from Uganda. *Sci. Rep.*, **7**, 11810 (2017).
10. Teresa M. et al. Severe malaria in an adult patient from low-endemic area in Flores Island, East Nusa Tenggara. *Case. Rep. Med.* (2023).
11. White, N. J. Severe malaria. *Malar J* **21**, 284 (2022).
12. Hermisson, J. & Pennings, P. S. Soft Sweeps: Molecular Population Genetics of Adaptation From Standing Genetic Variation. *Genetics*, **169**(4) (2005).
13. King, R. C., Stansfield, W. D., & Mulligan, P. K. A dictionary of genetics. Oxford University Press. (2006).

REVIEWERS' COMMENTS

Reviewer #1 (Remarks to the Author):

I am happy that the authors have addressed my comments

Reviewer #2 (Remarks to the Author):

The authors have adequately addressed my comments in the revised version of the manuscript.