nature portfolio

Peer Review File

Impact of intensive control on malaria population genomics under elimination settings in Southeast Asia

Corresponding Author: Dr Timothy Anderson

Version 0:

Reviewer comments:

Reviewer #1

(Remarks to the Author)
This manuscript reports a notable population genomic dataset that captures malaria parasite clone dynamics in a region of Thailand, Kayin State, that experiences low disease transmission, and where transmission was further reduced during the study sampling period through mass drug administration (MDA). The most important conclusion of this manuscript is that MDA did not generate evidence of selection for drug resistance in local parasite populations, informing future malaria intervention policy in this region and other pre-elimination locations. Other important findings that will be of interest to a broad audience include the usefulness of effective population size (as calculated from LD between unlinked markers) as a measure of changes in transmission in settings where transmission is already very low.

This manuscript is well-written, the figures are clear, and the analyses are well-chosen and have been conducted effectively. The manuscript could be further improved through attention to a modest number of issues.

Most importantly, the conclusion that MDA does not cause drug resistance selection should be fortified by more thoroughly extending kelch13-focused analyses to the loci known to mediate partner drug (piperaquine) resistance. The authors report that pfcrt mutations known to confer resistance in other regions were not observed in the Kayin State sampling sites. Were other pfcrt polymorphisms of unknown phenotypic impact observed? Were there any other discernible signals of selection (for example locally enhanced relatedness for a chromosomal position) in the parasite genomes post-MDA that could indicate resistance selection on other loci? The lack of attention to plasmepsin copy numbers should be mentioned as an explicit caveat to the conclusion that MDA does not drive drug resistance, if it cannot be addressed empirically. For example, could a subset of samples be analyzed with qRTPCR to determine if MDA impacted copy number of plasmepsin II/III?

Minor Comments:

Line 95: Rg not defined in the Results; reported in Methods, but not interpretable for readers linearly reading the manuscript

Line 155: Is it possible to infer the directionality of long-distance connectivity, given the relative frequency of clones in different places, knowledge of human movement patterns, or patterns of IBD (eg evidence of a common clone from 1 site that has recombined with a locally common clone in a sink population.)

Line 223: K13 diversity dropped by half in 2020, perhaps as a result of MDA conducted that year, but did not show drops in previous years when MDA was conducted. Can the authors speculate on why all of the MDA events did not have similar effects on parasite populations?

Line 242: Text is a bit confusing as worded: ‘We predicted that MDA would reduce relatedness of pre and post MDA parasite populations’; should this read ‘. . . reduce relatedness of post MDA parasite populations relative to pre-MDA’ or similar? The subsequent sentence suggesting founder effects would influence structure post-MDA structure would enhance relatedness. Guessing that this paragraph should actually read ‘… would reduce relatedness BETWEEN pre and post MDA parasite populations…’, with founder effects increasing relatedness within the post-MDA pops.

Line 260: In reporting the relatedness and clonal structure of parasite populations in the METF study relative to other geographical regions, it would be helpful to limit the cited statistics to the pre-MDA populations to exclude the immediate impact of the MDA intervention.

Line 328: define ‘complexity of infection’

Figure 3B: Can the authors speculate further in the text on why the autocorrelation coefficient for parasite populations at

intermediate distances (~50 km) is negative, rather than close to zero, as it is for comparisons of more distant populations (>= 100 km). What would cause this statistic to become so sharply negative for intermediate distances? Could this effect be driven by one particular genetically distinct population sharing a common distance of ~50 km from many of the other sampled populations?

Figure 4A: It would be informative to simultaneously plot incidence on these graphs to enable readers to visually evaluate the association of these genetic metrics with incidence.

(Remarks on code availability)

Reviewer #2

(Remarks to the Author)
Li et al describe an interesting malaria genomic epidemiology in the Kayin state of Myanmar. They show that the genomic epidemiology of the parasites in this region are consistent with previous reports that show that the parasite population in SE Asia are highly clonal and fragmented due to very low transmission. A unique aspect of this study was the use of IBD-related statistics to identify distinct, but sometimes genetically related, clonal families within Hpapun. This raises an interesting epidemiological situation characterized by clonal outbreaks/isolation with just enough mixing to occasionally create a new familial lineage that then goes on and creates its own clonal cluster. This raises an interesting possibility that that IBD-based genealogical inference might be useful to study the movement of different transmission lineages in extremely low transmission settings where the distribution of infections are extremely structured/fragmented to limit outcrossing opportunities.
Of particular interest is the reported resurgence of kelch13-561H in 2020, which they hypothesize is not necessarily the result of a new selective sweep, but the result of relative reduction in diversity due to a sharp reduction in population size. This highlights the complex relationship between genetics and epidemiology that is particularly relevant in (very) low transmission settings where drift and structured transmission heterogeneity are so important. The discussion that Ne may be a useful proxy for incidence, particularly in very low transmission settings, is interesting as well.
Overall, I find the paper is well presented and only have a few minor comments.

Major
1. Ne as an informative metric for incidence. If I may, I would like to offer a slightly more nuanced interpretation of the SMRU and Kayin incidence correlation results. It is curious that the correlation coefficient between Ne and incidence in Kayin is much higher for Kayin (r2 = 0.90) than in SMRU (r2 = 0.40). Closer inspection of the table 5B suggests that the multi-infection and Rg statistics are informative for Kayin, but still informative for SMRU. This could suggest that the relative usefulness of Ne , multi-infection, and Rg depend on incidence and that the “incidence boundaries” where Ne is informative and overtakes multi-infection and Rg useful lies somewhere between the incidence of SMRU and Kayin.
2.
Minor
1. Clarify definition of R_G on line 95 <- from context it is clear but spelling it out would be helpful
2. Figure 1: There is a lot of visual noise in these figures, which makes it difficult to read at a glance. The important information (health posts and their treatment, sample density, etc etc) requires a bit too much effort (in my opinion) to disentangle from the lesser features of the map (road network, rivers).
a. Hpapun marker: having a distinct marker on the map for where Hpapun is might be useful
b. Clarify Kayin state boundaries to be more visually obvious: suggest adding a thicker border around the boundaries – it is currently hard to see where Kayin state is because the yellow shading merges with the light blue shading (and also because yellow and blue area also used for roads and rivers)
c. Road/River networks: Suggest muting color scheme reduce the amount of visual noise while still preserving the salient features of the map. Ex: changing road networks to a light grey or similarly neutral color
d. Figure 1C: it might be worth considering censoring the topological map so that only the topology of the relevant study region is shown, similar to how 1A and 1B is done.

(Remarks on code availability)
The supporting code base is clean and easy to follow.

Reviewer #3

(Remarks to the Author)
In this manuscript, the authors present a comprehensive analysis of a large Plasmodium falciparum whole genome sequence data set from an area that has undergone intensive malaria elimination using a combination of community-based diagnosis and treatment and mass drug administration. The authors identify patterns consistent with a steep decline in malaria incidence, including reduced effective population size and clonal propagation, consistent with genetic drift. Importantly, they do not find evidence for selection of drug resistance alleles following MDA, and they show that effective population size may correlate better with reductions in transmission intensity in low malaria incidence settings compared to approaches based on infection complexity. The major limitation of this study is that there is a mismatch in sampling time for many of the between location comparisons, with many of the isolate sequences from other study sites being represented by older samples and the isolates analyzed from Kayin being more recent. This mismatch and its potential impact on study inferences were not discussed in depth in the manuscript.

Comments:

For several of the comparisons there is a mismatch in the sampling time between comparator populations. For example,

comparisons of genetic richness (lines 95-100) between countries/regions are comparing more recent isolates from Kayin to isolates sampled earlier in other countries. As a result, some of the differences in genetic richness could reflect sampling of infections prior to declines in P. falciparum observed in some of the comparator countries (e.g. Cambodia). Likewise, comparisons between K13 prevalence between Kayin and SMRU clinics (lines 203-205) and presence of specific K13 haplotypes (lines 232-233) would also likely be impacted by this temporal mismatch. I believe there is publicly available WGS data from more recent isolates from some of these comparator populations that could be examined to compare patterns in areas where malaria has also recently declined. At the very least, the implications of this temporal mismatch should be discussed in the context of the results.

Lines 351-355 – There have been prior studies showing limited genetic relatedness and different drug resistance alleles between parasite populations in East and West SE Asia. Please see: Genomic structure and diversity of Plasmodium falciparum in Southeast Asia reveal recent parasite migration patterns - PubMed (Shetty et al.) and The origins of malaria artemisinin resistance defined by a genetic and transcriptomic background - PubMed (Zhu et al.). As noted above this comparison would also be impacted by the temporal mismatch.

Supplementary Figure 4 – Why were different Fws threshold used for Kayin isolates versus isolates from other geographic areas? Although I find Fws difficult to interpret as a quantitative variable, could the authors comment on why the distribution of Fws is much wider (i.e., wider range of Fws values) for Kayin than for the other SE Asian sites?

Line 662 – How was IBD estimated? IsoRelate is mentioned in later sections of the methods to compare populations (line 702), but after the discussion of pruning highly related isolates, as defined by the proportion of the genome shared IBD. If IsoRelate was used, please discuss how use of this algorithm over others may have impacted IBD estimates. In addition, can the authors please discuss how the threshold of IBD to define clonal clusters was selected and whether the results were robust to the threshold used?

Minor comments:

There are several instances where additional context is required in the results section to explain things that are discussed in the methods section, since the methods section comes after the results section. For example, in line 95 – the authors need to define “RG”, which should also be defined in the Figure 2 legend. Likewise, beginning in lines 119-120, Hpapun township is mentioned, but it is not clear until line 179 that most of the sequenced samples come from this township. I think this statement should occur somewhere at the beginning of the results section, so the reader is not wondering why this particular township is being discussed over others.

Figure 1C – third flow chart box from the top should read “2744 samples excluded, too little parasite DNA” rather that “too less parasite DNA”

Figure 2 legend – I think the second sentence should read: “Pairwise parasite relatedness (r) was measured as the proportion of the genome shared identical by decent (IBD) between pairs of samples”.

Lines 148-149 – Sentence has extra words, please revise.

Line 177-178 – Sentence should read “The malaria incidence decreased significantly in both regions over the duration of the study”, or something to that effect, rather than “over studying time”.

Figure 5B – check for typos

Lines 701-702 – not sure of the meaning of “We measured the proportion of pairs IBD across the genome within populations following the scripts in isoRelate”. Do the authors mean the proportion of the genome IBD between pairs of isolates?

(Remarks on code availability)

Decision Letter:

14th August 2025

*Please ensure you delete the link to your author homepage in this e-mail if you wish to forward it to your co-authors.

Dear Tim,

Thank you for your patience while your manuscript "Impact of intensive control on malaria population genomics in Eastern Myanmar" was under peer-review at Nature Microbiology. It has now been seen by 3 referees, whose expertise and comments you will find at the end of this email. Although they find your work of some potential interest, they have raised a number of concerns that will need to be addressed before we can consider publication of the work in Nature Microbiology.

In particular, referee #1 says that the conclusion that MDA does not cause drug resistance selection should be fortified by more thoroughly extending kelch13-focused analyses to the loci known to mediate partner drug (piperaquine) resistance. The referee also asks to caveat the lack of analysis of plasmepsin copy numbers. Referee #2 suggests to have a more nuanced interpretation of the SMRU and Kayin incidence correlation results. Referee #3 says the major limitation of this study is that there

is a mismatch in sampling time for many of the between location comparisons, with many of the isolate sequences from other study sites being represented by older samples and the isolates analyzed from Kayin being more recent. The referee notes that at the very least, the implications of this temporal mismatch should be discussed. Furthermore, this referee asks why different Fws threshold were used for Kayin isolates versus isolates from other geographic areas. Referee #3 also asks how IBD was estimated.

Should further experimental data allow you to address these criticisms, we would be happy to look at a revised manuscript.

We are committed to providing a fair and constructive peer-review process. Please do not hesitate to contact us if there are specific requests from the reviewers that you believe are technically impossible or unlikely to yield a meaningful outcome.

We strongly support public availability of data. Please place the data used in your paper into a public data repository, if one exists, or alternatively, present the data as Source Data or Supplementary Information. If data can only be shared on request, please explain why in your Data Availability Statement, and also in the correspondence with your editor. For some data types, deposition in a public repository is mandatory - more information on our data deposition policies and available repositories can be found at https://www.nature.com/nature-research/editorial-policies/reporting-standards#availability-of-data.

Please include a data availability statement as a separate section after Methods but before references, under the heading "Data Availability". This section should inform readers about the availability of the data used to support the conclusions of your study. This information includes accession codes to public repositories (data banks for protein, DNA or RNA sequences, microarray, proteomics data etc…), references to source data published alongside the paper, unique identifiers such as URLs to data repository entries, or data set DOIs, and any other statement about data availability. At a minimum, you should include the following statement: "The data that support the findings of this study are available from the corresponding author upon request", mentioning any restrictions on availability. If DOIs are provided, we also strongly encourage including these in the Reference list (authors, title, publisher (repository name), identifier, year). For more guidance on how to write this section please see: http://www.nature.com/authors/policies/data/data-availability-statements-data-citations.pdf

When revising your manuscript:

* Include a "Response to referees" document detailing, point-by-point, how you addressed each referee comment. If no action was taken to address a point, you must provide a compelling argument. This response will be sent back to the referees along with the revised manuscript.

* If you have not done so already we suggest that you begin to revise your manuscript so that it conforms to our Article format instructions at http://www.nature.com/nmicrobiol/info/final-submission. Refer also to any guidelines provided in this letter.

* Include a revised version of any required reporting checklist. It will be available to referees (and, potentially, statisticians) to aid in their evaluation if the manuscript goes back for peer review. A revised checklist is essential for re-review of the paper.

When submitting the revised version of your manuscript, please pay close attention to our href="https://www.nature.com/nature-portfolio/editorial-policies/image-integrity">Digital Image Integrity Guidelines. and to the following points below:

-- that unprocessed scans are clearly labelled and match the gels and western blots presented in figures.
-- that control panels for gels and western blots are appropriately described as loading on sample processing controls
-- all images in the paper are checked for duplication of panels and for splicing of gel lanes.

EXTENDED DATA FIGURES

When re-submitting your manuscript, please ensure that any supplementary figures and tables that are crucial to the manuscript's conclusions are converted into Extended Data figures and tables to increase visibility of these data. Extended Data figures and tables are online-only (present in the online PDF and full-text HTML versions of the paper), peer-reviewed display items that provide essential background to the article but are not included in the main article due to space constraints. A maximum of ten Extended Data display items (figures and tables) is permitted.

Finally, please ensure that you retain unprocessed data and metadata files after publication, ideally archiving data in perpetuity, as these may be requested during the peer review and production process or after publication if any issues arise.

Please use the link below to submit a revised paper:

Link Redacted

Note: This url links to your confidential homepage and associated information about manuscripts you may have submitted or be reviewing for us. If you wish to forward this e-mail to co-authors, please delete this link to your homepage first.

Nature Microbiology is committed to improving transparency in authorship. As part of our efforts in this direction, we are now requesting that all authors identified as 'corresponding author' on published papers create and link their Open Researcher and Contributor Identifier (ORCID) with their account on the Manuscript Tracking System (MTS), prior to acceptance. This applies to primary research papers only. ORCID helps the scientific community achieve unambiguous attribution of all scholarly

contributions. You can create and link your ORCID from the home page of the MTS by clicking on ‘Modify my Springer Nature account’. For more information please visit please visit www.springernature.com/orcid.

If you wish to submit a suitably revised manuscript we would hope to receive it within 6 months. If you cannot send it within this time, please let us know.

In the meantime we hope that you find our referees' comments helpful.

Yours sincerely,

****

Reviewer Expertise:

Referee #1: Malaria evolutionary genomics
Referee #2: Malaria drug resistance, malaria transmissio
Referee #3: Malaria, epidemiology, global health, drug resistance

Reviewer Comments:

Reviewer #1 (Remarks to the Author):

This manuscript reports a notable population genomic dataset that captures malaria parasite clone dynamics in a region of Thailand, Kayin State, that experiences low disease transmission, and where transmission was further reduced during the study sampling period through mass drug administration (MDA). The most important conclusion of this manuscript is that MDA did not generate evidence of selection for drug resistance in local parasite populations, informing future malaria intervention policy in this region and other pre-elimination locations. Other important findings that will be of interest to a broad audience include the usefulness of effective population size (as calculated from LD between unlinked markers) as a measure of changes in transmission in settings where transmission is already very low.

This manuscript is well-written, the figures are clear, and the analyses are well-chosen and have been conducted effectively. The manuscript could be further improved through attention to a modest number of issues.

Most importantly, the conclusion that MDA does not cause drug resistance selection should be fortified by more thoroughly extending kelch13-focused analyses to the loci known to mediate partner drug (piperaquine) resistance. The authors report that pfcrt mutations known to confer resistance in other regions were not observed in the Kayin State sampling sites. Were other pfcrt polymorphisms of unknown phenotypic impact observed? Were there any other discernible signals of selection (for example locally enhanced relatedness for a chromosomal position) in the parasite genomes post-MDA that could indicate resistance selection on other loci? The lack of attention to plasmepsin copy numbers should be mentioned as an explicit caveat to the conclusion that MDA does not drive drug resistance, if it cannot be addressed empirically. For example, could a subset of samples be analyzed with qRTPCR to determine if MDA impacted copy number of plasmepsin II/III?

Minor Comments:

Line 95: Rg not defined in the Results; reported in Methods, but not interpretable for readers linearly reading the manuscript

Line 155: Is it possible to infer the directionality of long-distance connectivity, given the relative frequency of clones in different places, knowledge of human movement patterns, or patterns of IBD (eg evidence of a common clone from 1 site that has recombined with a locally common clone in a sink population.)

Line 223: K13 diversity dropped by half in 2020, perhaps as a result of MDA conducted that year, but did not show drops in previous years when MDA was conducted. Can the authors speculate on why all of the MDA events did not have similar effects on parasite populations?

Line 242: Text is a bit confusing as worded: ‘We predicted that MDA would reduce relatedness of pre and post MDA parasite populations’; should this read ‘. . . reduce relatedness of post MDA parasite populations relative to pre-MDA’ or similar? The subsequent sentence suggesting founder effects would influence structure post-MDA structure would enhance relatedness. Guessing that this paragraph should actually read ‘… would reduce relatedness BETWEEN pre and post MDA parasite populations…’, with founder effects increasing relatedness within the post-MDA pops.

Line 260: In reporting the relatedness and clonal structure of parasite populations in the METF study relative to other geographical regions, it would be helpful to limit the cited statistics to the pre-MDA populations to exclude the immediate impact of the MDA intervention.

Line 328: define ‘complexity of infection’

Figure 3B: Can the authors speculate further in the text on why the autocorrelation coefficient for parasite populations at intermediate distances (~50 km) is negative, rather than close to zero, as it is for comparisons of more distant populations (>= 100 km). What would cause this statistic to become so sharply negative for intermediate distances? Could this effect be driven

by one particular genetically distinct population sharing a common distance of ~50 km from many of the other sampled populations?

Figure 4A: It would be informative to simultaneously plot incidence on these graphs to enable readers to visually evaluate the association of these genetic metrics with incidence.

Reviewer #2 (Remarks to the Author):

Li et al describe an interesting malaria genomic epidemiology in the Kayin state of Myanmar. They show that the genomic epidemiology of the parasites in this region are consistent with previous reports that show that the parasite population in SE Asia are highly clonal and fragmented due to very low transmission. A unique aspect of this study was the use of IBD-related statistics to identify distinct, but sometimes genetically related, clonal families within Hpapun. This raises an interesting epidemiological situation characterized by clonal outbreaks/isolation with just enough mixing to occasionally create a new familial lineage that then goes on and creates its own clonal cluster. This raises an interesting possibility that that IBD-based genealogical inference might be useful to study the movement of different transmission lineages in extremely low transmission settings where the distribution of infections are extremely structured/fragmented to limit outcrossing opportunities.
Of particular interest is the reported resurgence of kelch13-561H in 2020, which they hypothesize is not necessarily the result of a new selective sweep, but the result of relative reduction in diversity due to a sharp reduction in population size. This highlights the complex relationship between genetics and epidemiology that is particularly relevant in (very) low transmission settings where drift and structured transmission heterogeneity are so important. The discussion that Ne may be a useful proxy for incidence, particularly in very low transmission settings, is interesting as well.
Overall, I find the paper is well presented and only have a few minor comments.

Major
1. Ne as an informative metric for incidence. If I may, I would like to offer a slightly more nuanced interpretation of the SMRU and Kayin incidence correlation results. It is curious that the correlation coefficient between Ne and incidence in Kayin is much higher for Kayin ($r^2 = 0.90$) than in SMRU ($r^2 = 0.40$). Closer inspection of the table 5B suggests that the multi-infection and Rg statistics are informative for Kayin, but still informative for SMRU. This could suggest that the relative usefulness of Ne , multi-infection, and Rg depend on incidence and that the “incidence boundaries” where Ne is informative and overtakes multi-infection and Rg useful lies somewhere between the incidence of SMRU and Kayin.
2.
Minor
1. Clarify definition of R_G on line 95 <- from context it is clear but spelling it out would be helpful
2. Figure 1: There is a lot of visual noise in these figures, which makes it difficult to read at a glance. The important information (health posts and their treatment, sample density, etc etc) requires a bit too much effort (in my opinion) to disentangle from the lesser features of the map (road network, rivers).
a. Hpapun marker: having a distinct marker on the map for where Hpapun is might be useful
b. Clarify Kayin state boundaries to be more visually obvious: suggest adding a thicker border around the boundaries – it is currently hard to see where Kayin state is because the yellow shading merges with the light blue shading (and also because yellow and blue area also used for roads and rivers)
c. Road/River networks: Suggest muting color scheme reduce the amount of visual noise while still preserving the salient features of the map. Ex: changing road networks to a light grey or similarly neutral color
d. Figure 1C: it might be worth considering censoring the topological map so that only the topology of the relevant study region is shown, similar to how 1A and 1B is done.

Reviewer #2 (Remarks on code availability):

The supporting code base is clean and easy to follow.

Reviewer #3 (Remarks to the Author):

In this manuscript, the authors present a comprehensive analysis of a large Plasmodium falciparum whole genome sequence data set from an area that has undergone intensive malaria elimination using a combination of community-based diagnosis and treatment and mass drug administration. The authors identify patterns consistent with a steep decline in malaria incidence, including reduced effective population size and clonal propagation, consistent with genetic drift. Importantly, they do not find evidence for selection of drug resistance alleles following MDA, and they show that effective population size may correlate better with reductions in transmission intensity in low malaria incidence settings compared to approaches based on infection complexity. The major limitation of this study is that there is a mismatch in sampling time for many of the between location comparisons, with many of the isolate sequences from other study sites being represented by older samples and the isolates analyzed from Kayin being more recent. This mismatch and its potential impact on study inferences were not discussed in depth in the manuscript.

Comments:

For several of the comparisons there is a mismatch in the sampling time between comparator populations. For example,

comparisons of genetic richness (lines 95-100) between countries/regions are comparing more recent isolates from Kayin to isolates sampled earlier in other countries. As a result, some of the differences in genetic richness could reflect sampling of infections prior to declines in P. falciparum observed in some of the comparator countries (e.g. Cambodia). Likewise, comparisons between K13 prevalence between Kayin and SMRU clinics (lines 203-205) and presence of specific K13 haplotypes (lines 232-233) would also likely be impacted by this temporal mismatch. I believe there is publicly available WGS data from more recent isolates from some of these comparator populations that could be examined to compare patterns in areas where malaria has also recently declined. At the very least, the implications of this temporal mismatch should be discussed in the context of the results.

Lines 351-355 – There have been prior studies showing limited genetic relatedness and different drug resistance alleles between parasite populations in East and West SE Asia. Please see: Genomic structure and diversity of Plasmodium falciparum in Southeast Asia reveal recent parasite migration patterns - PubMed (Shetty et al.) and The origins of malaria artemisinin resistance defined by a genetic and transcriptomic background - PubMed (Zhu et al.). As noted above this comparison would also be impacted by the temporal mismatch.

Supplementary Figure 4 – Why were different Fws threshold used for Kayin isolates versus isolates from other geographic areas? Although I find Fws difficult to interpret as a quantitative variable, could the authors comment on why the distribution of Fws is much wider (i.e., wider range of Fws values) for Kayin than for the other SE Asian sites?

Line 662 – How was IBD estimated? IsoRelate is mentioned in later sections of the methods to compare populations (line 702), but after the discussion of pruning highly related isolates, as defined by the proportion of the genome shared IBD. If IsoRelate was used, please discuss how use of this algorithm over others may have impacted IBD estimates. In addition, can the authors please discuss how the threshold of IBD to define clonal clusters was selected and whether the results were robust to the threshold used?

Minor comments:

There are several instances where additional context is required in the results section to explain things that are discussed in the methods section, since the methods section comes after the results section. For example, in line 95 – the authors need to define “RG”, which should also be defined in the Figure 2 legend. Likewise, beginning in lines 119-120, Hpapun township is mentioned, but it is not clear until line 179 that most of the sequenced samples come from this township. I think this statement should occur somewhere at the beginning of the results section, so the reader is not wondering why this particular township is being discussed over others.

Figure 1C – third flow chart box from the top should read “2744 samples excluded, too little parasite DNA” rather that “too less parasite DNA”

Figure 2 legend – I think the second sentence should read: “Pairwise parasite relatedness (r) was measured as the proportion of the genome shared identical by decent (IBD) between pairs of samples”.

Lines 148-149 – Sentence has extra words, please revise.

Line 177-178 – Sentence should read “The malaria incidence decreased significantly in both regions over the duration of the study”, or something to that effect, rather than “over studying time”.

Figure 5B – check for typos

Lines 701-702 – not sure of the meaning of “We measured the proportion of pairs IBD across the genome within populations following the scripts in isoRelate”. Do the authors mean the proportion of the genome IBD between pairs of isolates?

****

Version 1:

Reviewer comments:

Reviewer #1

(Remarks to the Author)
I am satisfied with the authors' thorough response to my comments and those of the other reviewers.

(Remarks on code availability)

Reviewer #3

(Remarks to the Author)

I have reviewed the authors' responses to the prior critique, as well as the revised manuscript, and feel that they have addressed all my comments.

(Remarks on code availability)

Decision Letter:

Our ref: NMICROBIOL-25062099A

4th February 2026

Dear Tim,

Thank you for submitting your revised manuscript "Impact of intensive control on malaria population genomics in Eastern Myanmar" (NMICROBIOL-25062099A). It has now been seen by the original referees and their comments are below. The reviewers find that the paper has improved in revision, and therefore we'll be happy in principle to publish it in Nature Microbiology, pending minor revisions to comply with our editorial and formatting guidelines.

We are now performing detailed checks on your paper and will send you a checklist detailing our editorial and formatting requirements in about a week. Please do not upload the final materials and make any revisions until you receive this additional information from us.

Thank you again for your interest in Nature Microbiology Please do not hesitate to contact me if you have any questions.

Sincerely,

Reviewer #1 (Remarks to the Author):

I am satisfied with the authors' thorough response to my comments and those of the other reviewers.

Reviewer #3 (Remarks to the Author):

I have reviewed the authors' responses to the prior critique, as well as the revised manuscript, and feel that they have addressed all my comments.

Version 2:

Decision Letter:

16th March 2026

Dear Tim,

I am pleased to accept your Article "Impact of intensive control on malaria population genomics under elimination settings in Southeast Asia" for publication in Nature Microbiology. Thank you for having chosen to submit your work to us and many congratulations.

Over the next few weeks, your paper will be copyedited to ensure that it conforms to Nature Microbiology style. We look particularly carefully at the titles of all papers to ensure that they are relatively brief and understandable.

Once your paper is typeset, you will receive an email with a link to choose the appropriate publishing options for your paper and our Author Services team will be in touch regarding any additional information that may be required. Once your paper has been scheduled for online publication, the Nature press office will be in touch to confirm the details.

You may wish to make your media relations office aware of your accepted publication, in case they consider it appropriate to organize some internal or external publicity. Once your paper has been scheduled you will receive an email confirming the publication details. This is normally 3-4 working days in advance of publication. If you need additional notice of the date and time of publication, please let the production team know when you receive the proof of your article to ensure there is sufficient time to coordinate. Further information on our embargo policies can be found here:
https://www.nature.com/authors/policies/embargo.html

After the grant of rights is completed, you will receive a link to your electronic proof via email with a request to make any

corrections within 48 hours. If, when you receive your proof, you cannot meet this deadline, please inform us at rjsproduction@springernature.com immediately. You will not receive your proofs until the publishing agreement has been received through our system

Due to the importance of these deadlines, we ask you please us know now whether you will be difficult to contact over the next month. If this is the case, we ask you provide us with the contact information (email, phone and fax) of someone who will be able to check the proofs on your behalf, and who will be available to address any last-minute problems.

Acceptance of your manuscript is conditional on all authors' agreement with our publication policies (see https://www.nature.com/nmicrobiol/editorial-policies). In particular your manuscript must not be published elsewhere.

Authors may need to take specific actions to achieve compliance with funder and institutional open access mandates. If your research is supported by a funder that requires immediate open access (e.g. according to Plan S principles or the NIH public access policy) then you should select the gold OA route, and we will direct you to the compliant route where possible. Because authors warrant under our subscription licensing terms that they haven't committed to licensing any version of their article under a licence inconsistent with the terms of our agreement – including the applicable embargo period – publication under the subscription model isn't suitable for authors whose funders require no embargo.

If you have any questions about our publishing options, costs, Open Access requirements, or our legal forms, please contact ASJournals@springernature.com

An online order form for reprints of your paper is available at https://www.nature.com/reprints/author-reprints.html. All co-authors, authors' institutions and authors' funding agencies can order reprints using the form appropriate to their geographical region.

We welcome the submission of potential cover material (including a short caption of around 40 words) related to your manuscript; suggestions should be sent to Nature Microbiology as electronic files (the image should be 300 dpi at 210 x 297 mm in either TIFF or JPEG format). Please note that such pictures should be selected more for their aesthetic appeal than for their scientific content, and that colour images work better than black and white or grayscale images. Please do not try to design a cover with the Nature Microbiology logo etc., and please do not submit composites of images related to your work. I am sure you will understand that we cannot make any promise as to whether any of your suggestions might be selected for the cover of the journal.

You can now use a single sign-on for all your accounts, view the status of all your manuscript submissions and reviews, access usage statistics for your published articles and download a record of your refereeing activity for the Nature journals.

To assist our authors in disseminating their research to the broader community, our SharedIt initiative provides you with a unique shareable link that will allow anyone (with or without a subscription) to read the published article. Recipients of the link with a subscription will also be able to download and print the PDF.

As soon as your article is published, you will receive an automated email with your shareable link.

Congratulations once again and I look forward to seeing the article published.

With kind regards,

P.S. Click on the following link if you would like to recommend Nature Microbiology to your librarian http://www.nature.com/subscriptions/recommend.html#forms

** Visit the Springer Nature Editorial and Publishing website at www.springernature.com/editorial-and-publishing-jobs for more information about our career opportunities. If you have any questions please click here.**

Response to reviewers

We thank the reviewers and editor for the comprehensive and insightful comments on our manuscript. We have addressed each point in detail (blue text), incorporating several new analyses and revising the manuscript accordingly. We hope these revisions have addressed all the concerns and that the manuscript is now suitable for publication.

Reviewer #1 (Remarks to the Author):

This manuscript reports a notable population genomic dataset that captures malaria parasite clone dynamics in a region of Thailand, Kayin State, that experiences low disease transmission, and where transmission was further reduced during the study sampling period through mass drug administration (MDA). The most important conclusion of this manuscript is that MDA did not generate evidence of selection for drug resistance in local parasite populations, informing future malaria intervention policy in this region and other pre-elimination locations. Other important findings that will be of interest to a broad audience include the usefulness of effective population size (as calculated from LD between unlinked markers) as a measure of changes in transmission in settings where transmission is already very low.

This manuscript is well-written, the figures are clear, and the analyses are well-chosen and have been conducted effectively. The manuscript could be further improved through attention to a modest number of issues.

Most importantly, the conclusion that MDA does not cause drug resistance selection should be fortified by more thoroughly extending kelch13-focused analyses to the loci known to mediate partner drug (piperaquine) resistance. The authors report that pfcrt mutations known to confer resistance in other regions were not observed in the Kayin State sampling sites. Were other pfcrt polymorphisms of unknown phenotypic impact observed? Were there any other discernible signals of selection (for example locally enhanced relatedness for a chromosomal position) in the parasite genomes post-MDA that could indicate resistance selection on other loci? The lack of attention to plasmepsin copy numbers should be mentioned as an explicit caveat to the conclusion that MDA does not drive drug resistance, if it cannot be addressed empirically. For example, could a subset of samples be analyzed with qRTPCR to determine if MDA impacted copy number of plasmepsin II/III?

We thank the reviewer for these constructive suggestions. In response, we have expanded our analyses to include *pfcrt* polymorphisms and plasmepsin II/III (*pm2/3*) copy number variation (CNV), as summarized in the new Supplementary Table 4.

- *pfcrt SNP*: We did not detect any known *pfcrt* single nucleotide polymorphisms (SNPs) associated with piperaquine (PPQ) resistance. All *pfcrt* polymorphisms identified in Kayin samples have been previously described and are predominantly background alleles linked to chloroquine (CQ) resistance.

- *pm2/3* copy number: We conducted copy number variation (CNV) analysis using Illumina short-read sequencing data. CNV regions were defined based on the established *pm2/3* amplification boundaries reported in MalariaGEN Pf8 and Kane *et al.*, 2024. Read coverage across the CNV region was compared with that of 10 kb upstream and downstream flanking regions. Coverage fold-change was then calculated as (mean coverage within the CNV region) / (mean coverage in the flanking regions). Samples were classified as putative *pm2/3* amplifications if the coverage fold-change > 1.5. Among the 281 samples (58

collected post-MDA) with sufficient sequencing depth (>20× coverage across both CNV and flanking regions), three samples had a coverage fold-change greater than 1.5. However, manual inspection using Integrative Genomics Viewer (IGV) revealed no clear CNV boundaries (i.e., no step-like increase in read depth). Signals from these three samples were likely artifacts resulting from selective whole-genome amplification. Therefore, we conclude that there is no evidence of *pm2/3* amplification in the analyzed samples.

- *pfmdr1 SNPs and copy number:* We did not detect any *pfmdr1* gene amplification, which is a known marker of lumefantrine resistance, in the Kayin dataset. The predominant *pfmdr1* haplotypes were the wild type (N86–Y184–D1246; 81.1%) and NFD (N86–184F–D1246; 18.8%). The frequency of the NFD haplotype remained relatively stable between 2016 and 2019 (13.8–36.8%), but was not detected in 2020. The F1226Y allele, which is commonly found in Southeast Asia, was also observed at a high frequency (50.1%) in Kayin; however, this is unlikely due to recent selection. Overall, these results indicate no evidence of *pfmdr1*-mediated selection in the samples analyzed.

- Additional drug-resistance markers: To further characterize the drug-resistance landscape in the Kayin population, we examined additional loci, including *aat1*, *dhfr*, and *dhps*, as well as CNVs in *gch1*, and deletions in *hrp2* and *hrp3*. For *gch1* CNV analyses, we applied the same approach used for *pm2/3* CNV detection, defining CNV boundaries based on MalariaGEN Pf8. To identify potential *hrp2* or *hrp3* deletions, we visually inspected samples showing read coverage <3× across the *hrp2* or *hrp3* loci, while maintaining >20× coverage in the *β-tubulin* gene region (40kb flanking) as an internal control. Samples with read coverage greater than 3× across the *hrp2* and *hrp3* loci were considered to have intact *hrp2* and *hrp3* genes. No partial or complete *hrp2* or *hrp3* deletions were observed. The predominant *aat1* mutations in the Kayin population were K541N (53.65%), F313S (79.65%), G308E (10.25%), and S258L (45.26%). A high prevalence (>90%) of *pfdhfr* and *pfdhps* mutations confirming sulfadoxine-pyrimethamine resistance was detected, as expected. *gch1* amplification was observed in more than 80% of samples.

The analysis results are summarized in Supplementary Table 4.

Genome-wide selection scans

Beyond known resistance loci, we performed genome-wide scans to identify potential signals of selection following MDA (Figure S17). We compared IBD relatedness across the genome by sampling year and pre- vs post-MDA. No strong signal of recent selection was detected.

The manuscript has been revised accordingly in the Methods (lines 736-750) and Results (lines 273–282).

References

1. Malaria Genomic Epidemiology, N. *et al.* Pf8: an open dataset of Plasmodium falciparum genome variation in 33,325 worldwide samples. *Wellcome Open Res* **10**, 325 (2025).
2. Kane, J. *et al.* A Plasmodium falciparum genetic cross reveals the contributions of pfcrt and plasmepsin II/III to piperaquine drug resistance. *mBio* **15**, e0080524 (2024).

Minor Comments:

Line 95: Rg not defined in the Results; reported in Methods, but not interpretable for readers linearly reading the manuscript

We have now defined R_G in the Results section (Line 95-96) when it first appears to ensure clarity for readers who are following the manuscript sequentially.

Line 155: Is it possible to infer the directionality of long-distance connectivity, given the relative frequency of clones in different places, knowledge of human movement patterns, or patterns of IBD (eg evidence of a common clone from 1 site that has recombined with a locally common clone in a sink population.)

Thanks for this suggestion. Human movement along the Thailand–Myanmar border is complex. Kayin State (particularly Hpapun, where most of our samples were collected) is located deep in mountainous terrain, whereas the SMRU clinics are situated near Mae Sot, a major cross-border travel hub. These differences in geography and mobility likely influence parasite flow between the two regions.

To explore the directionality of long-distance connectivity, we examined patterns of IBD between the two parasite populations. As no identical clones were shared between Kayin and SMRU, we used closely related parasite pairs (IBD ≥ 0.35) as indicators of recent gene flow, following the approach used in Shetty et al., 2019. We inferred directionality by comparing the relative frequency of each closely related parasite group across sites. When a parasite group (IBD ≥ 0.35) was more common in one location and rare in the other, we interpreted the more prevalent site as the likely source.

Most closely related parasite groups were restricted to a single site, suggesting predominantly local transmission. However, we identified several cases indicating bidirectional long-distance connectivity. For example, two parasites from Kayin had their closest relatives in SMRU, while five parasites from SMRU were more closely related to parasites from Kayin than to those within SMRU. These patterns suggest that parasite flow occurs in both directions.

We note, however, that malaria incidence is now extremely low in the SMRU catchment area, whereas transmission persists at relatively higher levels in Kayin State. Thus, differences in recent epidemiological dynamics may weaken the precision of directional inference, as the Kayin parasite population continues to evolve while very few SMRU samples were collected after 2014.

Overall, our analysis supports the presence of bidirectional long-distance connectivity, although the strength and detectability of directional flow are influenced by current epidemiological conditions. We have included these results in Supplementary figure 9 legend.

Reference

Shetty, A. C. *et al.* Genomic structure and diversity of Plasmodium falciparum in Southeast Asia reveal recent parasite migration patterns. *Nat Commun* **10**, 2665 (2019).

Line 223: K13 diversity dropped by half in 2020, perhaps as a result of MDA conducted that year, but did not show drops in previous years when MDA was conducted. Can the authors speculate on why all of the MDA events did not have similar effects on parasite populations?

It is important to note that two different control approaches - mass drug administration (MDA) and intensive case management through onsite malaria posts - were implemented concurrently. In combination, these two approaches resulted in both efficient case detection and treatment, leading to a reduction in malaria incidence. By 2020, most sampling sites were located in areas

of very low transmission, where only a few locations continued to report infections. The gradual reduction in population size – and effective population size (*Ne*) – over the study period resulted in severe population bottleneck in 2020, and contributed to the marked loss of *kelch13* allelic diversity.

We have added this explanation to the lines 240-244.

Line 242: Text is a bit confusing as worded: 'We predicted that MDA would reduce relatedness of pre and post MDA parasite populations'; should this read '. . . reduce relatedness of post MDA parasite populations relative to pre-MDA' or similar? The subsequent sentence suggesting founder effects would influence structure post-MDA structure would enhance relatedness. Guessing that this paragraph should actually read '… would reduce relatedness BETWEEN pre and post MDA parasite populations…', with founder effects increasing relatedness within the post-MDA pops.

We have revised the sentence for clarity. The updated wording is (line 264-264): "We predicted that MDA would reduce relatedness between pre- and post-MDA parasite populations, by clearing local infections and enabling repopulation by new parasite genotypes, leading to founder effects."

Line 260: In reporting the relatedness and clonal structure of parasite populations in the METF study relative to other geographical regions, it would be helpful to limit the cited statistics to the pre-MDA populations to exclude the immediate impact of the MDA intervention.

We have recalculated relatedness using only pre-MDA samples and those collected from regions where no MDA was implemented (lines 295–297). This analysis identified 153 genotype clusters (defined as pairs sharing ≥90% of the genome IBD) among 1408 single-genotype samples, resulting in an updated R_G value of 0.11.

Line 328: define 'complexity of infection'

We have now added a precise definition in line 364-365.
Specifically, "complexity of infection (COI)" refers to the estimates of the number of genetically distinct parasite strains co-infecting an individual (Schaffner et al., 2023; Wong et al., 2024). COI can be used to distinguish single clone infections (COI = 1) from poly-genomic infections (COI > 1).

References:
1. Schaffner, S. F. *et al.* Malaria surveillance reveals parasite relatedness, signatures of selection, and correlates of transmission across Senegal. *Nat Commun* **14**, 7268 (2023).
2. Wong, W. *et al.* Evaluating the performance of Plasmodium falciparum genetic metrics for inferring National Malaria Control Programme reported incidence in Senegal. *Malar J* **23**, 68 (2024).

Figure 3B: Can the authors speculate further in the text on why the autocorrelation coefficient for parasite populations at intermediate distances (~50 km) is negative, rather than close to zero, as it is for comparisons of more distant populations (>= 100 km). What would cause this statistic to become so sharply negative for intermediate distances? Could this effect be driven by one particular genetically distinct population sharing a common distance of ~50 km from many of the other sampled populations?

We thank the reviewer for the opportunity to clarify the interpretation of the autocorrelation coefficient. We have expanded the text to explain why negative values were observed at ~50 km.

Autocorrelation coefficients range from +1 (more similar than expected) to –1 (more different than expected), with 0 indicating random similarity. Negative values at intermediate distances suggest that nearby populations are sometimes more genetically distinct than expected under isolation-by-distance. This switching from positive to negative is a common feature of correlograms in ecological and genetic studies (Legendre & Fortin, 1989; Fortin & Dale, 2005), indicative of strong spatial clustering. In our dataset, the strong spatial clustering of clonal parasite genotypes (see Figure 4A) results in the patterns observed. The distance where correlation changes sign (≈25 km) gives an empirical measure of the spatial scale of local transmission.

We have added this interpretation to the Results (lines 128–132) and expanded relevant sections of the Methods (lines 784–788).

References

1. Legendre, P., & Fortin, M.-J. (1989). *Spatial pattern and ecological analysis*. Vegetatio, 80(2), 107–138.
2. Fortin, M.-J., & Dale, M. R. T. (2005). *Spatial Analysis: A Guide for Ecologists*. Cambridge University Press.

Figure 4A: It would be informative to simultaneously plot incidence on these graphs to enable readers to visually evaluate the association of these genetic metrics with incidence.

We have now incorporated malaria incidence data into Figure 4A as suggested.

Reviewer #2 (Remarks to the Author):

Li et al describe an interesting malaria genomic epidemiology in the Kayin state of Myanmar. They show that the genomic epidemiology of the parasites in this region are consistent with previous reports that show that the parasite population in SE Asia are highly clonal and fragmented due to very low transmission. A unique aspect of this study was the use of IBD-related statistics to identify distinct, but sometimes genetically related, clonal families within Hpapun. This raises an interesting epidemiological situation characterized by clonal outbreaks/isolation with just enough mixing to occasionally create a new familial lineage that then goes on and creates its own clonal cluster. This raises an interesting possibility that that IBD-based genealogical inference might be useful to study the movement of different transmission lineages in extremely low transmission settings where the distribution of infections are extremely structured/fragmented to limit outcrossing opportunities.
Of particular interest is the reported resurgence of kelch13-561H in 2020, which they hypothesize is not necessarily the result of a new selective sweep, but the result of relative reduction in diversity due to a sharp reduction in population size. This highlights the complex relationship between genetics and epidemiology that is particularly relevant in (very) low transmission settings where drift and structured transmission heterogeneity are so important. The discussion that Ne may be a useful proxy for incidence, particularly in very low transmission settings, is interesting as well.
Overall, I find the paper is well presented and only have a few minor comments.

Major
1. Ne as an informative metric for incidence. If I may, I would like to offer a slightly more nuanced interpretation of the SMRU and Kayin incidence correlation results. It is curious that the correlation coefficient between Ne and incidence in Kayin is much higher for Kayin (r2 = 0.90) than in SMRU (r2 = 0.40). Closer inspection of the table 5B suggests that the multi-infection and Rg statistics are informative for Kayin, but still informative for SMRU. This could suggest that the relative usefulness of Ne, multi-infection, and Rg depend on incidence and that the “incidence boundaries” where Ne is informative and overtakes multi-infection and Rg useful lies somewhere between the incidence of SMRU and Kayin.

We agree that the relationship among genetic metrics (N_e, multi-infection and R_G) likely depends on transmission intensity.

Following the reviewer’s suggestion, we recalculated the correlations using an incidence boundary of 100 cases per 1,000 population per year to further explore this relationship (Wong et al., 2024).

Our updated analysis supports the reviewer’s interpretation. *Ne* remains strongly correlated with incidence in lower-transmission settings for both Kayin and SMRU regions. Multiple genotype infection is also correlated with incidence in low-transmission settings, but this pattern is primarily observed in the SMRU region. In higher-incidence areas, none of the genetic metrics show a strong correlation with incidence, although the absolute values of these metrics are markedly elevated compared to low-transmission sites.

These findings suggest that *Ne* serves as a more reliable genetic indicator for incidence below approximately 100 cases per 1,000 population per year. We have incorporated this clarification into lines 181–211.

Reference

Wong, W. *et al.* Evaluating the performance of Plasmodium falciparum genetic metrics for inferring National Malaria Control Programme reported incidence in Senegal. *Malar J* **23**, 68 (2024).

Minor
1. Clarify definition of R_G on line 95 <- from context it is clear but spelling it out would be helpful

We have now clarified the definition of genetic richness (R_G) at its first mention in line 95-96. Specifically, we define R_G as the number of unique genomes divided by the total number of parasites.

2. Figure 1: There is a lot of visual noise in these figures, which makes it difficult to read at a glance. The important information (health posts and their treatment, sample density, etc etc) requires a bit too much effort (in my opinion) to disentangle from the lesser features of the map (road network, rivers).
a. Hpapun marker: having a distinct marker on the map for where Hpapun is might be useful
b. Clarify Kayin state boundaries to be more visually obvious: suggest adding a thicker border around the boundaries – it is currently hard to see where Kayin state is because the yellow shading merges with the light blue shading (and also because yellow and blue area also used for roads and rivers)
c. Road/River networks: Suggest muting color scheme reduce the amount of visual noise while still preserving the salient features of the map. Ex: changing road networks to a light grey or similarly neutral color
d. Figure 1C: it might be worth considering censoring the topological map so that only the topology of the relevant study region is shown, similar to how 1A and 1B is done.

Thank you for this suggestion. We have revised Figure 1, following these suggestions, to improve visual readability and emphasize the key geographic and epidemiological features.

Reviewer #2 (Remarks on code availability):

The supporting code base is clean and easy to follow.

We thank the reviewer for this positive feedback. We are delighted to hear that the code was clean and easy to follow.

Reviewer #3 (Remarks to the Author):

In this manuscript, the authors present a comprehensive analysis of a large Plasmodium falciparum whole genome sequence data set from an area that has undergone intensive malaria elimination using a combination of community-based diagnosis and treatment and mass drug administration. The authors identify patterns consistent with a steep decline in malaria incidence, including reduced effective population size and clonal propagation, consistent with genetic drift. Importantly, they do not find evidence for selection of drug resistance alleles following MDA, and they show that effective population size may correlate better with reductions in transmission intensity in low malaria incidence settings compared to approaches based on infection complexity. The major limitation of this study is that there is a mismatch in sampling time for many of the between location comparisons, with many of the isolate sequences from other study sites being represented by older samples and the isolates analyzed from Kayin being more recent. This mismatch and its potential impact on study inferences were not discussed in depth in the manuscript.

Comments:

For several of the comparisons there is a mismatch in the sampling time between comparator populations. For example, comparisons of genetic richness (lines 95-100) between countries/regions are comparing more recent isolates from Kayin to isolates sampled earlier in other countries. As a result, some of the differences in genetic richness could reflect sampling of infections prior to declines in P. falciparum observed in some of the comparator countries (e.g. Cambodia). Likewise, comparisons between K13 prevalence between Kayin and SMRU clinics (lines 203-205) and presence of specific K13 haplotypes (lines 232-233) would also likely be impacted by this temporal mismatch. I believe there is publicly available WGS data from more recent isolates from some of these comparator populations that could be examined to compare patterns in areas where malaria has also recently declined. At the very least, the implications of this temporal mismatch should be discussed in the context of the results.

We agree that temporal differences in sampling could influence comparisons of genetic richness, *K13* prevalence and *K13* haplotype composition, particularly because several comparator regions (e.g., Cambodia) experienced major transmission declines during and after earlier sampling periods.

To address this, we incorporated the most recent publicly available WGS dataset (MalariaGEN Pf8, Wellcome Sanger Institute, 2025 release), which includes recently collected isolates from both Asia and Africa and is more closely aligned with the Kayin sampling period. We updated all relevant comparative analyses (genetic richness, K13 allele frequency, and haplotype networks) using this expanded dataset, and revised the Methods, Results, Discussion and Supplementary Figures/Tables to clearly document sampling windows by dataset and country.

- For genetic richness comparisons (Figure 2B, Table S3, lines 97-105), we now stratify the analysis by both country/region and year. The updated results show a marked decline in genetic richness over time across multiple Southeast Asian populations, whereas genetic richness in African populations remains stable. The Kayin region has R_G values that are comparable to recent population samples from Vietnam and Laos, but considerably lower than those observed in other populations. This new analysis indicates that the Kayin populations are representative of other low transmission populations in SE Asia.

- *K13* prevalence and haplotype comparisons between Kayin and SMRU clinics: There were very few *P. falciparum* malaria cases at SMRU clinics after 2014 (only 12 sequenced samples collected between 2015 and 2017). Hence comparison of contemporary samples from both Kayin and SMRU clinics was not possible. We therefore retained the original analysis and explicitly acknowledge this limitation in the revised text (lines 218-220, line 257-259).

Lines 351-355 – There have been prior studies showing limited genetic relatedness and different drug resistance alleles between parasite populations in East and West SE Asia. Please see: Genomic structure and diversity of Plasmodium falciparum in Southeast Asia reveal recent parasite migration patterns - PubMed (Shetty et al.) and The origins of malaria artemisinin resistance defined by a genetic and transcriptomic background - PubMed (Zhu et al.). As noted above this comparison would also be impacted by the temporal mismatch.

We thank the reviewer for pointing us to these important references. We have now cited both studies (Shetty et al., 2019; Zhu et al., 2018) in the revised Discussion (lines 389–392) and clarified that our observed population differentiation is consistent with these findings.

Supplementary Figure 4 – Why were different Fws threshold used for Kayin isolates versus isolates from other geographic areas? Although I find Fws difficult to interpret as a quantitative variable, could the authors comment on why the distribution of Fws is much wider (i.e., wider range of Fws values) for Kayin than for the other SE Asian sites?

We thank the reviewer for this comment.

We have now recalculated F_{WS} in the Kayin dataset using pipelines kindly provided by the authors (Drs. Nina White and Richard Pearson) of the MalariaGEN Pf8 dataset. The Kayin samples now show a plateau in the cumulative frequency distribution for F_{WS} between 0.95 and 1, comparable to the other malaria populations. The lower plateau in our original submission resulted from differences in the methods used for data filtering, prior to implementation of the F_{WS} pipeline.

We have updated methods and analysis the analysis pipeline for F_{WS} accordingly. We have also updated Supplementary Figure 4 (now Supplementary Figure 3) and Figure 5 accounting for the corrected F_{WS} analyses and revised the numbers of single clone infections (1726 to 1781) and associated analyses in the manuscript. The revised analyses did not result in changes to our conclusions.

Line 662 – How was IBD estimated? IsoRelate is mentioned in later sections of the methods to compare populations (line 702), but after the discussion of pruning highly related isolates, as defined by the proportion of the genome shared IBD. If IsoRelate was used, please discuss how use of this algorithm over others may have impacted IBD estimates. In addition, can the authors please discuss how the threshold of IBD to define clonal clusters was selected and whether the results were robust to the threshold used?

IBD estimation:
We now clarify that two IBD methods were used for complementary purposes (lines 720-734). *hmmIBD* was used to estimate pairwise whole-genome IBD across the entire sample set, which enabled us to identify clonal parasites and quantify genetic relatedness (*r*) between isolates. In contrast, *IsoRelate* was used for genome-wide scans to detect regions with elevated shared IBD, which we interpreted as signatures of recent selection. The revised Methods section now

explicitly states when each algorithm was applied and why both were necessary (lines 720-734). This distinction is now made at the first mention of IBD, rather than later in the manuscript.

Clonal clustering threshold and robustness:
To determine an appropriate IBD threshold for defining clonal clusters, we examined cloned parasites from three experimental genetic crosses (Supplementary Figure 4). Based on these empirical data, we selected an IBD cutoff of ≥0.9. Although truly clonal parasites are expected to show IBD = 1, a slightly lower threshold provides a conservative buffer for genotyping errors, which are more common in datasets generated from finger-prick blood samples and whole-genome amplification (WGA) (Supplementary Figure 4D).

Minor comments:

There are several instances where additional context is required in the results section to explain things that are discussed in the methods section, since the methods section comes after the results section. For example, in line 95 – the authors need to define “RG”, which should also be defined in the Figure 2 legend. Likewise, beginning in lines 119-120, Hpapun township is mentioned, but it is not clear until line 179 that most of the sequenced samples come from this township. I think this statement should occur somewhere at the beginning of the results section, so the reader is not wondering why this particular township is being discussed over others.

We have now defined R_G (genetic richness) at its first mention in the Results section (line 95) and added this definition to the Figure 2 legend.

We have also moved the statement that most sequenced samples come from Hpapun township to line 124, so readers understand why this location is discussed in detail later.

Figure 1C – third flow chart box from the top should read “2744 samples excluded, too little parasite DNA” rather that “too less parasite DNA”

Revised as suggested. The text in the third flow-chart box of Figure 1C now reads: “2744 samples excluded, too little parasite DNA.”

Figure 2 legend – I think the second sentence should read: “Pairwise parasite relatedness (r) was measured as the proportion of the genome shared identical by decent (IBD) between pairs of samples”.

Thank you for pointing this out. The sentence now reads: “Pairwise parasite relatedness (*r*) was measured as the proportion of the genome shared identical by descent (IBD) between pairs of samples.”

Lines 148-149 – Sentence has extra words, please revise.
Revised the sentence to remove extra words and improve clarity (line 156).

Line 177-178 – Sentence should read “The malaria incidence decreased significantly in both regions over the duration of the study”, or something to that effect, rather than “over studying time”.

We revised the sentence as suggested to: “Malaria incidence declined sharply in both regions over the study period: from 273.9 to 2.2 cases per 1,000 person-years between 2001 and 2014

in the SMRU clinics, and from 58.8 to 5.3 cases per 1,000 person-years between 2016 and 2020 in Hpapun Township in northern Kayin State, where >96% of sequenced samples originated (Table S3).” (lines 184-188).

Figure 5B – check for typos

We have reviewed Figure 5B and corrected the typos.

Lines 701-702 – not sure of the meaning of “We measured the proportion of pairs IBD across the genome within populations following the scripts in isoRelate”. Do the authors mean the proportion of the genome IBD between pairs of isolates?

We thank the reviewer for pointing out this ambiguity. *IsoRelate* was used to perform genome-wide scans to quantify the proportion of the genome shared IBD between pairs of isolates and to identify regions with elevated IBD, which we interpreted as signatures of recent selection. We have revised the sentence for clarity to avoid confusion about what was measured. Additional clarification on the distinction between our IBD analyses and their purposes is also provided in the response above.