nature portfolio

Peer Review File

Eliminating malaria transmission requires targeting immature and mature gametocytes through lipoidal uptake of antimalarials

Corresponding Author: Professor Lyn-Marie Birkholtz

This file contains all reviewer reports in order by version, followed by all author rebuttals in order by version.

Version 0:

Reviewer comments:

Reviewer #1

(Remarks to the Author)
This manuscript presents extensive work characterising antimalarial molecules against both asexual stage parasites that cause disease pathology and gametocytes that are responsibly for transmission to the mosquito. It focuses on comparing the differential activity of these molecules across these life stages and tests some hypotheses which could account for why some molecules are more active on a particular stage than another. This is then linked to the implication that some future antimalarial treatments might be optimal for curing the disease but not effective against transmission.
The authors report an impressive body of data, however in places I struggled to understand the reasoning and conclusions that the authors draw from their work. In places, the text would benefit from further explanations and clearly described data which would aid interpretation for a non-expert reader. There are exciting novel data presented, however there are a lot of assumptions made in places that need explaining further and justified.
General comments:
- Figure legends need to be much clearer in explaining the figures. Abbreviations are not always explained.
- P values are given for “significance” using statistical tests that I am not always sure are appropriate for the data.
- The methods state that unless indicated, all data is the mean of at least 3 independent experiments. I would prefer each figure legend to show the precise number of independent experiments for each data shown as it is not clear which experiments had fewer or more.
- Drug codes/names. Please check the manuscript for consistency. For example, it is confusing to refer to Cabamiquine in the text and M5717 in the figures. Please select one and use it throughout.
- In the text and figure legends, frequently sentences are extremely long (I counted one sentence that covered 11 lines!). It would aid understanding if these could be broken up.
- Throughout the manuscript the terminology for which parasite stage is being tested is ambiguous. “Mature” gametocytes are those that have completed development and are fertile for mosquito transmission. The authors describe “Mature” gametocytes as mixed populations containing mostly (80-90%) stage V gametocytes and the remainder stage IV gametocytes. It would be better to describe these as “late” stage gametocytes in line with the conventions published in many comparable papers in the last decade.

Specific comments:
Figure 1
This figure shows the clustered activity profiles of a variety of antimalarial molecules derived from previously published data. Please make the figure legend more understandable – it is an extremely long sentence. I am also confused as to why the graph in Figure 1 plots average IC50 of each cluster group. If the clusters are made based upon fold change between activity of parasite stage, surely this is independent of IC50. Therefore molecules in a cluster could have widely variable IC50 values as long as the FC pattern is similar?
Figure 2
The authors then select a panel of molecules to test in their own assays rather than rely on published data to perform a similar analysis (reducing confounding factors). It is not clear here how the k means clustering has been performed. Is the clustering based upon IC50 against each parasite stage? If so, how does this handle molecules with similar activity profiles but different starting IC50 values against asexuals? Also, the blue/white heatmap colors in Fig 2 refer to IC50, whilst in Fig 1 they refer to p-value. This is quite confusing.

Line 163 – Be more specific – PfPKG inhibitors prevent the initiation of gametogenesis rather than gametocyte rounding up and gamete formation.
Line 189 – How can a cell be morphologically viable? What is morphologically viable? I would suggest “has morphology indistinguishable from untreated cells” would be more appropriate.
Figure 3
The authors note that previous reports have shown that sub-optimal doses of some antimalarial drugs can actually enhance gametocyte commitment. They then use a variety of dosing schedules and readouts to determine the effect of selected molecules throughout gametocyte development. I do not understand the authors selecting 1xIC50 against asexual parasites as the dose that they tested in this figure. Clearly at this concentration only 50% of the asexual parasites would be killed. This would not be an appropriate dose for treating a patient and of course it will only have a partial/no effect downstream on gametocytes which are less sensitive – which is obviously what the authors observe. I realise that a fixed concentration needed to be selected for comparison between molecules, but why not use a higher concentration? Why not the IC99? Why not 10xIC50? Antimalarial drugs will be administered to patients presumably in doses more than adequate to kill all asexual parasites (at least in a combination therapy). Therefore, does it really matter if gametocytes show several-fold less activity? I fear that due to selecting low concentrations, this figure demonstrates a much more worrying lack of activity against gametocytes than is really necessary.
Line 277 – “Together with their potential to induce gametocytogenesis, this raises concerns that compounds with such profiles could not prevent gametocyte differentiation, with the resultant immature gametocytes not effectively killed.” This is only a concern if you are treating at 1x(asexual)IC50. I would argue that this is less of a concern if you treated at a higher (more physiologically relevant) concentration.
Line 288 – p-value stated but no reference to the statistical test used (this is a general comment as it is found throughout the text). The p-value is rather meaningless without information about the statistical test used.
Lines 284-289 – This needs clarification – Fig 2 D shows immature gametocytes treated with 1x(iGc)IC50. Surely this should mean for each drug tested, 50% of the gametocytes should die and so at the endpoint, the maximum number of gametocytes present at maturity can be no more than 50% of the original starting value? Perhaps if you take the mean gametocytemia at Day 13 of all treatments then you get about 50%? In this case how can the reduction by M5717 be deemed “significant”. Is this not just random spread? This is all the more confusing if compared to panel E. If the immature gametocytes were treated at 1x(iGc)IC50 and then allowed to mature, surely there can only be at most 50% gametocytes left at maturity to form gametes? How can, for example, MMV533 have around 60% male gametogenesis compared to the control and M5717 be 80%?
Line 298 – “Gametes could be produced from gametocytes treated by all the compounds, with a significant reduction in gamete formation only seen for Cabamiquine (>80 % reduction) and MMV253 against female gametocytes ($P < 0.01$, Figure 3E).” I find this interpretation concerning as I cannot see how it could have been derived from the data presented. Firstly, the control is normalised to 100% therefore it is impossible to perform a student’s t-test to compare for significant reductions. Secondly, Fig 3E does not even show this. The statistical test appears to be comparing within a drug treatment, effectively the sex ratio of female and male gametogenesis. Please check this for errors and correct.
Line 300 – I don’t understand how the authors have decided that compounds with 5-fold less activity against immature gametocytes compared to asexual parasites are good whilst >5-fold are bad. It is mentioned several times but the reasoning is not really explained at all.

Figure 4
I found this figure particularly confusing to interpret. What is it showing? What is the point it is trying to make? What is the pattern? The title of this section implies that there is a correlation between target abundance and activity profiles but then on line 333 the authors state that there is no linear relationship? Is there a relationship or not? I think that the authors are trying to argue that above a certain threshold there is a relationship and below there is not. But how was this threshold determined? What it randomly chosen or experimentally derived?
Line 308 – It is very confusing trying to understand how transcriptomic data and proteomic data was incorporated into this analysis. It needs more explanation about methods used.

Figure 5
The authors used machine learning to correlate physicochemical properties of the molecules with stage-specific activity. They identified several factors that might be correlated including topological polar surface area – linking stage-specific activity potentially with the ability of a compound to access the parasite.
Line 365 – Typo? Divers/drivers?
Figure 6
The authors tested the hypothesis that stage-specific activity of test molecules could be linked to the ability for the molecule to enter the cell of that particular stage. This is an interesting hypothesis and they cleverly used analogues of the same molecule with the same mode of action but different physicochemical activity to investigate this. The data is highly novel and does appear to demonstrate that gametocytes and asexuals have different propensities to accumulate molecules. To my knowledge this hypothesis has not been proposed or tested before and is potentially highly significant.
Line 490 – Sentence is 11 lines long(!)
Line 497 – To me, the scenario presented is not very realistic. This statement would need significant modelling to determine if it was a real concern. The antimalarial dose administered and its relationship to the actual asexual IC50 needs to be considered. If the antimalarial is being given substantially higher than the IC50 (which is likely), why is it a concern if gametocytes are 10-fold less sensitive?
Line 507 – I disagree entirely. There is currently no clear clinical evidence for the need of a drug that targets immature gametocytes. I would argue that the authors provide evidence to contradict their own statement. In Fig 1 and 2 of this manuscript, virtually all molecules presented show that mature (“late”) gametocytes are the least sensitive to killing. Therefore, if you want to block transmission, you will need to treat at the concentration that kills the mature (“late”)

gametocytes and by very nature, this will also kill the less-sensitive immature gametocytes. Whilst it is important to characterise killing of immature gametocytes, there is no need to design molecules to target them.
Line 514 – Again, see point above. There is little rationale for focusing on clearance of immature gametocytes if they are less sensitive than mature gametocytes.
Line 524 – This correlation is confusing. How was it determined?
Line 537 – See point above. This statement seems to directly contradict it. Is there or is there not a correlation between target abundance and activity?
Line 555 – This is confusing – asexual parasites have an IMC as well.
Line 601 – Typp - /– 5
Line 659 – 80% stage V is not “mature” gametocytes

Reviewer #2

(Remarks to the Author)
The authors have provided a framework of activity profiles of antimalarial candidates active against ABS parasites and different stages of gametocytes. This work is of general interest to the readers however certain sections of this work are unclear.

Please find below my comments regarding the same.

Naïve Bayes classifier assume that features are conditionally independent. It is not clear how the authors are utilizing the results obtained from PCA and SHAP analysis for machine learning model building, was the model retrained on the important and uncorrelated features identified from PCA and SHAP analysis? How was correlation between logD and logP (both are related to lipophilicity) handled?

It is not clear why authors utilized Naïve bayes classifier instead of PLS-DA or other algorithms. Fig 5b shows that PC1 and PC2 are able to capture 85.6% variability in the data which suggest that the data may be linear hence PLS or other algorithms may work as well.

It seems like the authors have utilized a single random split (80:20) to generate the training and test dataset for their machine learning model based on their github code. One of the drawbacks of a single random data split is that the performance of the model dependent on the random state (seed) value. It is highly recommended that the authors follow some kind of cross validation technique to increase confidence in their model prediction and understand the inherent variance within the model as well.

Was the dataset used for building the classifier balanced? Did the authors use some kind of oversampling/undersampling technique to ensure a balanced dataset? If not, please specify why the authors did not feel the need to do so.

Additionally, it is recommended to include some information/plots regarding the underlying distribution of the test and training dataset.

The authors have utilized a t-test to perform statistical significance analysis but one of the major assumptions regarding t-tests is that the distribution of the population must follow normal distribution, did the authors ensure that the distribution of dataset used in the work follows a normal distribution? If not, it may be better to use non-parametric tests for this case.

Line 477: Authors have "highlighted that the extent of loss of activity of compounds from ABS parasites to mature gametocytes is more pronounced than previously perceived". Did authors perform any statistical significance analysis to compare the results of the different analysis?

Minor comments:
750: "compounds belonged to. A total of 2013". Perhaps should be "compounds belonged to a total of 2013"

Line 751: Github link doesn’t work but the code can be accessed using Github link on line 822

Figure S2 a) X Axis labels should be ABS, iGc, mGc instead of ABS, II/III, V. Same applies to Figure S3e.

Supplementary: Data analysis section "Statistical significance was determined either with unpaired Student’s t-tests using GraphPad Prism (version 9) program" is there a typo here or do authors utilize another approach to determine statistical significance

Fig S6a perhaps a violin plot may be better to represent this data.

Reviewer #3

(Remarks to the Author)
In this manuscript, the authors evaluate the anti-plasmodial activity of several antimalarial compounds on Plasmodium falciparum asexual and sexual stages (gametocytes). By performing a systematic meta-analysis of published IC50 data from 118 compounds currently used and in clinical development, they conclude that most compounds exhibit a loss of activity between asexual and sexual parasites. To circumvent the variations between different assay platforms and conditions used

in different laboratories to calculate the IC50s, they confirm these findings on 10 compounds using the same assay and the same luciferase reporter line. They further investigate the confounding factors contributing to the reduced drug activity against gametocytes. Their data show that for most drugs, the activity profiles correlate with the drug target abundance between the stages, but this is not the case for compounds losing >10-fold of their activity in gametocytes. Using cheminformatics and machine learning, they identified physicochemical properties of the drug (topological polar surface area, solubility, and lipophilicity) as the most critical chemical features that contribute to compound efficacy. Accordingly, they provide evidence that PSAC and aquaglyceroporin, which are absent in mature gametocytes, contribute to compound uptake in asexual stages. Finally, they show that the different lipid composition of the infected erythrocyte in asexual parasites compared to gametocytes results in reduced diffusion of compounds into gametocytes. They conclude that the loss in activity of compounds against gametocytes correlates to specific requirements in the physicochemical properties to allow for effective diffusion across membranes.
The authors address an important question that will likely attract interest in the scientific community and that must be considered for development of new antimalarials. Although several conclusions made by the authors only confirm what was already known in the field, there are several novel aspects to these findings. The experiments performed and reported were convincing. In most cases, the data provide the evidence needed for the authors to make their conclusions.
Major comments
- The results reported lines 399-401 do not correspond to the data shown in figure 6D: there is not significant changes for MMV048 at all stages, nor for DHA and Blasticidine at immature gametocyte stages.
- This is surprising that inhibition of PSAC with furosemide decreases the activity for MMV048 and DHA in asexual stages, since PSAC is not active in ring stages, which represent 95% of the asexual blood stage parasites used in this study.
- All the statistical analyses are done with unpaired Student's t-tests. However, ANOVA should be more appropriate in several figures (for example in Figure 1, 2 and 6 where several conditions are compared).
- The authors address how the abundance of parasite protein targets correlates with the activity profiles of the drugs; however, they do not address or at least discuss how the different concentration of hemoglobin in the red blood cell between asexual (ring stages) and gametocytes contributes to this phenomenon. It is an important point since the mechanism of action of several antimalarials, including chloroquine or artemisinin derivatives, relies on the degradation of hemoglobin.
Minor comments
- the legend of Figure 1 is not clear
- Figure 2B: Primer location should be shown on the schema
- l 240: The transgenic reporter line expressing tdTomato red fluorescence under the Pfef1α promotor is not described in the reference 36 (Portugalizia et al, 2019).
- Supp Fig S4C: All gametocytes shown in this figure look more like stage IV gametocytes than mature stages V. Also, the morphology of Cabamaquine-treated parasites does not appear different than the other ones (at least on the shown pictures).
- l 452: the sentence "compounds that prefer ABS parasites are also found outside this preferred absorption model" should be rephrase, since only 6 out of 31 compounds are outside of the ellipse.
- l 486: replace gametogenesis by gametocytogenesis
- in the discussion, l 514 – 517, the authors discuss the caveat that compounds would need to be able to access erythroblastic islands where gametocyte sequester. They should also discuss that the activity profiles of several drugs could be modified if gametocytes develop within erythroblasts, as it has been suggested in PMID: 25009232 and PMID: 32589714.

Version 1:

Reviewer comments:

Reviewer #1

(Remarks to the Author)
This revised manuscript is substantially improved both in clarity and data presentation. Most of my previous issues have been satisfactorily addressed. Below are some outstanding issues, most of which are relatively minor and could be addressed with simple clarification of sentence structure.

Line 112: "As expected" – why was this expected? Also low dose primaquine is an effective anti-gametocyte treatment. However it needs liver metabolism to the active form which is missing from in vitro assays. Probably not the best antimalarial to be discussing in this context.

Line 117: Confusing sentence structure – as read, it appears that CHX is a 4AQ or and 8AQ.

Line 152L/Fig 2A: Please add ML10 to the list of compounds in cluster 6 seeing as the dose response curve is included in Fig 2B.

Lines 225-230: "Under these conditions, gametocytogenesis still occurred, similar to the sexual commitment rates for most of the compounds evaluated…." What does "similar" mean? Does this statement mean that there was no statistically significant difference? If so, it would be good to state this as it would strengthen argument.

Lines 233-235: Are the authors demonstrating that a 130-fold loss in activity results in only a doubling of immature gametocytes? Whilst any increase is clearly undesirable, how concerning actually is this? If it takes >100-fold loss of activity to increase the population by 100% this does not seem that much to me.

Lines 239-253/Fig3: Did not reproduce well in the pdf and some of the data is cut off from Fig2C and E so it is not possible to comment on this section.

Line 271: “Taken together, only compounds that cannot effectively target immature gametocytes will not necessarily prevent onward transmission of the parasites.” This is a very confusing sentence structure – please clarify. Does it mean, “Taken together, compounds that are ineffective at targeting immature gametocytes may also be ineffective at preventing onward parasites transmission to the mosquito”?

Line 378: Are you sure PfAQP is absent from mature gametocytes? This needs a citation. In Pb, genetically tagged AQP is present in gametocytes (but higher expression in asexuals): https://doi.org/10.1016/j.ijpara.2012.10.006

Reviewer #2

(Remarks to the Author)
Authors have addressed my comments.

Reviewer #3

(Remarks to the Author)
All my concerns have been adequately addressed by the authors

Reviewers feedback

Reviewer #1 (Remarks to the Author):

This manuscript presents extensive work characterising antimalarial molecules against both asexual stage parasites that cause disease pathology and gametocytes that are responsibly for transmission to the mosquito. It focuses on comparing the differential activity of these molecules across these life stages and tests some hypotheses which could account for why some molecules are more active on a particular stage than another. This is then linked to the implication that some future antimalarial treatments might be optimal for curing the disease but not effective against transmission.

The authors report an impressive body of data, however in places I struggled to understand the reasoning and conclusions that the authors draw from their work. In places, the text would benefit from further explanations and clearly described data which would aid interpretation for a non-expert reader. There are exciting novel data presented, however there are a lot of assumptions made in places that need explaining further and justified.

General comments:
- Figure legends need to be much clearer in explaining the figures. Abbreviations are not always explained.
Response: This was corrected throughout all the figure legends to make sure the legend is as comprehensive as possible to allow in-depth explanation of the data presented in the figures.

- P values are given for "significance" using statistical tests that I am not always sure are appropriate for the data.
Response: Thank you for highlighting this. All statistical tests were re-checked and, where needed, replaced with a different test. We confirmed whether each dataset is normally distributed and where comparison was done either with 2 groups or with more than 2 groups requiring t-test or ANOVA comparisons.

- The methods state that unless indicated, all data is the mean of at least 3 independent experiments. I would prefer each figure legend to show the precise number of independent experiments for each data shown as it is not clear which experiments had fewer or more.
Response: We corrected this and the precise number of independent experiments were added for each dataset in each of the figure legends.

- Drug codes/names. Please check the manuscript for consistency. For example, it is confusing to refer to Cabamiquine in the text and M5717 in the figures. Please select one and use it throughout.
Response: Corrected as requested.

- In the text and figure legends, frequently sentences are extremely long (I counted one sentence that covered 11 lines!). It would aid understanding if these could be broken up.
Response: We checked all sentence constructions and corrected long sentences throughout the manuscript.

- Throughout the manuscript the terminology for which parasite stage is being tested is ambiguous. "Mature" gametocytes are those that have completed development and are fertile for mosquito transmission. The authors describe "Mature" gametocytes as mixed populations containing mostly (80-90%) stage V gametocytes and the remainder stage IV gametocytes. It would be better to describe these as "late" stage gametocytes in line with the conventions published in many comparable papers in the last decade.
Response: This is correct, mature gametocytes are stage V gametocytes. In our protocol, we refer to the mature population as having >90% stage V gametocytes, as reported before in Reader et al., 2015, 2022 from our lab, comparable to other papers where mature gametocytes were produced for gametocytocidal assays as in Plouffe et al., 2015; Lucantoni et al., 2015 and Lelievre et al., 2012. In all cases where we refer to mature gametocytes, the population consisted of >90% stage V, with <~8% consisting of stage IV and no earlier stages present. Leaving the gametocytes post-day 13 on our protocol results in a dramatic reduction of stage V gametocyte viability such that we cannot comfortably

use these populations for a further 48 h drug assay. We are confident in our enrichment of stage V gametocytes in our populations as we see a clear (and expected) drop in activity of a large group of compounds between our late-stage gametocyte population (where we have >10% stage IV parasites present). We do however recognise the confusion where we referred to >80% stage V gametocytes in the transgenic NF54^{nap} line that may have been construed as mature and changed reference to this to late-stage gametocytes as suggested.

Specific comments:
Figure 1 This figure shows the clustered activity profiles of a variety of antimalarial molecules derived from previously published data. Please make the figure legend more understandable – it is an extremely long sentence. I am also confused as to why the graph in Figure 1 plots average IC50 of each cluster group. If the clusters are made based upon fold change between activity of parasite stage, surely this is independent of IC50. Therefore molecules in a cluster could have widely variable IC50 values as long as the FC pattern is similar?
Response: We changed the figure legend to include details and allow a clearer explanation of the data provided. We also shortened all sentences and inserted reference to A and B elements in the figure and in the legend. The K-means clustering was based on IC50s and not on fold change. Therefore, we are considering variable IC50s as they cluster together and only thereafter were fold changes calculated.

Figure 2 The authors then select a panel of molecules to test in their own assays rather than rely on published data to perform a similar analysis (reducing confounding factors). It is not clear here how the k means clustering has been performed. Is the clustering based upon IC50 against each parasite stage? If so, how does this handle molecules with similar activity profiles but different starting IC50 values against asexuals?
Response: The clustering was based on IC_{50}s against each parasite stage. The clustering therefore groups compounds together that share similar activity profiles against the different stages. Depending on the number of clusters used (based on elbow plot analysis) compounds with differing IC_{50} values against ABS stages can either be clustered together due to their similar activity profiles or be clustered into 2 separate clusters with similar activity profiles but then separated from each other due to the different IC50 values against ABS parasites. Based on elbow plot analysis, 8 clusters were the optimum number of clusters for the dataset in Figure 2A. In Figure 2A, clusters 6 and 4 showed similar activity profiles but separated because the IC50s against the ABS parasites between these 2 clusters are very different. Therefore, the clustering addressed these differing IC50s against different parasite stages.

Also, the blue/white heatmap colors in Fig 2 refer to IC50, whilst in Fig 1 they refer to p-value. This is quite confusing.
Response: We changed the colour in Figure 1 for p-values were changed from blue shades to grey shades to prevent any confusion.

Line 163 – Be more specific – PfPKG inhibitors prevent the initiation of gametogenesis rather than gametocyte rounding up and gamete formation.
Response: Corrected as suggested.

Line 189 – How can a cell be morphologically viable? What is morphologically viable? I would suggest "has morphology indistinguishable from untreated cells" would be more appropriate.
Response: Changed as suggested.

Figure 3 The authors note that previous reports have shown that sub-optimal doses of some antimalarial drugs can actually enhance gametocyte commitment. They then use a variety of dosing schedules and readouts to determine the effect of selected molecules throughout gametocyte development. I do not understand the authors selecting 1xIC50 against asexual parasites as the dose that they tested in this figure. Clearly at this concentration only 50% of the asexual parasites would be killed. This would not be an appropriate dose for treating a patient and of course it will only have a partial/no effect downstream on gametocytes which are less sensitive – which is obviously what the authors observe. I realise that a fixed concentration needed to be selected for comparison between molecules, but why not use a higher concentration? Why not the IC99? Why not 10xIC50? Antimalarial drugs will be administered to patients presumably in doses more than adequate to kill all asexual parasites (at least in a combination therapy). Therefore, does it really matter if gametocytes show several-fold less activity? I fear that due to selecting low concentrations, this figure demonstrates a much more worrying lack of activity against gametocytes than is really necessary.

Response: Thank you for highlighting this. For the commitment studies (Figure 3B), we initially used the IC_{50} as determined against the ABS parasites to have comparable data to previous papers, including those selected based on Thommen 2022, which observed an increase in commitment around the ABS IC_{50} of compounds. However, we have repeated this experiment at IC_{90} and at concentrations corresponding to the effective human dose with no significant increase in gametocyte commitment rates observed, although gametocytes could still be formed under IC_{90} treatment conditions. We similarly repeated all the other experiments on immature gametocytes, gametocyte maturation and gamete formation at the therapeutically relevant concentrations.

Line 277 – "Together with their potential to induce gametocytogenesis, this raises concerns that compounds with such profiles could not prevent gametocyte differentiation, with the resultant immature gametocytes not effectively killed." This is only a concern if you are treating at 1x(asexual)IC50. I would argue that this is less of a concern if you treated at a higher (more physiologically relevant) concentration.
Response: As per the comment above, re-evaluation was performed at higher concentrations. Although the risk for increased induction of gametocytogenesis was diminished, it was indeed still present.

Line 288 – p-value stated but no reference to the statistical test used (this is a general comment as it is found throughout the text). The p-value is rather meaningless without information about the statistical test used.
Response: Corrected as requested.

Lines 284-289 – This needs clarification – Fig 2 D shows immature gametocytes treated with 1x(iGc)IC50. Surely this should mean for each drug tested, 50% of the gametocytes should die and so at the endpoint, the maximum number of gametocytes present at maturity can be no more than 50% of the original starting value? Perhaps if you take the mean gametocytemia at Day 13 of all treatments then you get about 50%? In this case how can the reduction by M5717 be deemed "significant". Is this not just random spread? This is all the more confusing if compared to panel E. If the immature gametocytes were treated at 1x(iGc)IC50 and then allowed to mature, surely there can only be at most 50% gametocytes left at maturity to form gametes? How can, for example, MMV533 have around 60% male gametogenesis compared to the control and M5717 be 80%?
Response: This is correct. At the higher concentrations in our re-test, we did see the expected decrease in gametocytemia within the first 48 h (decrease to 0.5% gametocytemia from the initial starting gametocytemia of 5%) for most compounds (panel D). MMV390048 treatment however did not show this full extent, possibly due to the slow-acting nature of this compound (Paquet et al, 2017), as well as its sex-specificity as reflected in panel E.

Line 298 – "Gametes could be produced from gametocytes treated by all the compounds, with a significant reduction in gamete formation only seen for Cabamiquine (>80 % reduction) and MMV253 against female gametocytes ($P < 0.01$, Figure 3E)." I find this interpretation concerning as I cannot see how it could have been derived from the data presented. Firstly, the control is normalised to 100% therefore it is impossible to perform a student's t-test to compare for significant reductions. Secondly, Fig 3E does not even show this. The statistical test appears to be comparing within a drug treatment, effectively the sex ratio of female and male gametogenesis. Please check this for errors and correct.
Response: Thank you for this observation. Indeed, in the re-test at higher concentrations, the reduction in mature gametocytemia's correlated to a reduced ability to produce gametes in all instances evaluated. We also corrected the statistical analysis.

Line 300 – I don't understand how the authors have decided that compounds with 5-fold less activity against immature gametocytes compared to asexual parasites are good whilst >5-fold are bad. It is mentioned several times but the reasoning is not really explained at all.
Response: Thank you for highlighting this. The 5-fold loss in activity was based on Figure 3C where compounds that showed 3-7 fold loss in activity inhibited gametocytes to the same extent as inhibition of ABS parasites. However, to avoid confusion, we removed this statement and only indicated the requirement of compounds to effectively target immature gametocytes, to prevent seeding maturation of gametocytes and onwards transmission.

Figure 4
I found this figure particularly confusing to interpret. What is it showing? What is the point it is trying to make? What is the pattern? The title of this section implies that there is a correlation between target

abundance and activity profiles but then on line 333 the authors state that there is no linear relationship? Is there a relationship or not? I think that the authors are trying to argue that above a certain threshold there is a relationship and below there is not. But how was this threshold determined? What it randomly chosen or experimentally derived?
Response: We realised that the figure legend in Figure 4 lacks clarity and does not clearly explain the content; we apologise. We changed the figure legend to fully explain the data. The figure essentially indicates the correlation between the loss of a compounds activity between ABS and either immature or mature gametocytes, and the abundance levels of the protein target of this compound. Indeed, there is no general correlation between the target abundance of a compound and its activity profile, except for a few compounds where there is a 1:1 relationship, where the activity of the compound is explained by the presence of the protein target. In the majority of cases, the abundance of the protein target is not predictive of the activity of the compound. This was clarified in the text.

Line 308 – It is very confusing trying to understand how transcriptomic data and proteomic data was incorporated into this analysis. It needs more explanation about methods used.
Response: The proteomic data was used for Figure 4 and a separate analysis was performed using transcriptomic data, which was shown in Supplementary Figure S5 (a and b) and not in Figure 4. We clarified this in the text.

Line 365 – Typo? Divers/drivers?
Response: Corrected

Figure 6
The authors tested the hypothesis that stage-specific activity of test molecules could be linked to the ability for the molecule to enter the cell of that particular stage. This is an interesting hypothesis and they cleverly used analogues of the same molecule with the same mode of action but different physicochemical activity to investigate this. The data is highly novel and does appear to demonstrate that gametocytes and asexuals have different propensities to accumulate molecules. To my knowledge this hypothesis has not been proposed or tested before and is potentially highly significant.

Line 490 – Sentence is 11 lines long(!)
Response: Apologies for this. This was corrected.

Line 497 – To me, the scenario presented is not very realistic. This statement would need significant modelling to determine if it was a real concern. The antimalarial dose administered and its relationship to the actual asexual IC50 needs to be considered. If the antimalarial is being given substantially higher than the IC50 (which is likely), why is it a concern if gametocytes are 10-fold less sensitive?
Response: This statement was re-evaluated given the above concerns, and also in light of the re-evaluation of the efficacy of compounds at more therapeutically relevant concentrations as presented in Figure 3. Indeed, therapeutically relevant concentrations are higher than the IC_{50} against ABS parasites, and we used this concentration to show if the compounds can then effectively kill gametocytes. In most instances, a substantial increase in efficacy was seen, but gametocytes were not completely eliminated. We argue that equipotency remains favourable, and compounds with significant decreased efficacy against gametocytes will still support gametocyte maturation to numbers that would sustain transmission.

Line 507 – I disagree entirely. There is currently no clear clinical evidence for the need of a drug that targets immature gametocytes. I would argue that the authors provide evidence to contradict their own statement. In Fig 1 and 2 of this manuscript, virtually all molecules presented show that mature ("late") gametocytes are the least sensitive to killing. Therefore, if you want to block transmission, you will need to treat at the concentration that kills the mature ("late") gametocytes and by very nature, this will also kill the less-sensitive immature gametocytes. Whilst it is important to characterise killing of immature gametocytes, there is no need to design molecules to target them. Line 514 – Again, see point above. There is little rationale for focusing on clearance of immature gametocytes if they are less sensitive than mature gametocytes.
Response: Thank you for highlighting this. Our argument was based on the advantage that compounds would have if they could target immature gametocytes, and we did not mean to suggest that new compounds be designed to specifically target only gametocytes. Indeed, if compounds with activity against multiple stages are used at therapeutically relevant concentrations, they should be able to target immature gametocytes similarly. Indeed, this was the case in our re-evaluated data. We, therefore,

modified this statement in the discussion to focus on the requirement of compounds having dual activity against ABS and mature gametocytes as a minimum requirement.

Line 524 – This correlation is confusing. How was it determined?
Response: This information was explained better in the section related to Figure 4 and its figure legend as requested above. The statement here in the discussion was clarified to indicate that protein abundance is not predictive of gametocytocidal activity.

Line 537 – See point above. This statement seems to directly contradict it. Is there or is there not a correlation between target abundance and activity?
Response: This was corrected as above for clarity.

Line 555 – This is confusing – asexual parasites have an IMC as well.
Response: Correct, this was corrected in the text.

Line 601 – Typp - /– 5
Response: Corrected, thank you.

Line 659 – 80% stage V is not "mature" gametocytes
Response: Correct, this was changed to late-stage gametocytes.

Reviewer #2 (Remarks to the Author):
The authors have provided a framework of activity profiles of antimalarial candidates active against ABS parasites and different stages of gametocytes. This work is of general interest to the readers however certain sections of this work are unclear. Please find below my comments regarding the same.

Naïve Bayes classifier assume that features are conditionally independent. It is not clear how the authors are utilizing the results obtained from PCA and SHAP analysis for machine learning model building, was the model retrained on the important and uncorrelated features identified from PCA and SHAP analysis? How was correlation between logD and logP (both are related to lipophilicity) handled? It is not clear why authors utilized Naïve bayes classifier instead of PLS-DA or other algorithms. Fig 5b shows that PC1 and PC2 are able to capture 85.6% variability in the data which suggest that the data may be linear hence PLS or other algorithms may work as well.
Response: We thank the reviewer for their insight. We re-evaluated the models in light of above and rejected the Gaussian Naive Bayes multiclassification (GNB) model due to it violating the condition of independent features. Under conditions better suited for the data, random forest (RF) provided better model performance, considering the unbalanced multiclass distribution of the data.

It seems like the authors have utilized a single random split (80:20) to generate the training and test dataset for their machine learning model based on their github code. One of the drawbacks of a single random data split is that the performance of the model dependent on the random state (seed) value. It is highly recommended that the authors follow some kind of cross validation technique to increase confidence in their model prediction and understand the inherent variance within the model as well.
Response: We thank the reviewer for their comment and agree. This was performed and is included in Supplementary Figure 6 and in the methods where a 10-fold cross-validation was performed. Please note that we have substituted GNB with RF for its better performance and suitability to the classification problem above.

Was the dataset used for building the classifier balanced? Did the authors use some kind of oversampling/undersampling technique to ensure a balanced dataset? If not, please specify why the authors did not feel the need to do so.
Response: The dataset was not balanced, and therefore, we did apply alternative techniques to ensure a more balanced dataset. Previous modelling on unbalanced data showed that sampling techniques like oversampling on specific antimalarial drug datasets tend to result in poor models (low recall precision) (van Heerden, 2023). We evaluated different sampling techniques to improve model performance such as oversampling, undersampling and cluster-based undersampling. These techniques did not improve model performance (supplementary file S5) as, e.g. undersampling reduced the chemical diversity and resulted in too many missing features. Oversampling negatively affected

weighted recall and precision. Cluster-based undersampling would not be informative given the small dataset. Therefore, we relied on algorithms less influenced by inherent class bias within the data such as GNB and decision trees, with RF and gradient boosting machines (GBM) performing well.

Additionally, it is recommended to include some information/plots regarding the underlying distribution of the test and training dataset.
Response: Additional plots were included in Supplementary Figure 6.

The authors have utilized a t-test to perform statistical significance analysis but one of the major assumptions regarding t-tests is that the distribution of the population must follow normal distribution, did the authors ensure that the distribution of dataset used in the work follows a normal distribution? If not, it may be better to use non-parametric tests for this case.
Response: We re-evaluated the data for the paper for normal distribution and statistical tests were changed where required. Figure 1 and 2 IC_{50} data is not normally distributed, therefore, we applied a Kruskal-Wallis test with Dunn's multiple comparison test to evaluate significance. Subsequent data were normally distributed and either a t-test or ANOVA was performed and indicated.

Line 477: Authors have "highlighted that the extent of loss of activity of compounds from ABS parasites to mature gametocytes is more pronounced than previously perceived". Did authors perform any statistical significance analysis to compare the results of the different analysis?
Response: We were unable to compare this statistically as the previous extent of loss was based on perceptions from various thresholds from different studies, which makes direct comparison of the data impossible. This highlights and supports our study where we investigated this loss of activity in a comparative analysis between the different stages of the parasite as presented in this work, allowing a concrete evaluation and numeration of the extent of loss of activity.

Minor comments:
750: "compounds belonged to. A total of 2013". Perhaps should be "compounds belonged to a total of 2013"
Response: Changed as requested.

Line 751: Github link doesn't work but the code can be accessed using Github link on line 822
Response: This was checked and corrected, thank you for highlighting this.

Figure S2 a) X Axis labels should be ABS, iGc, mGc instead of ABS, II/III, V. Same applies to Figure S3e.
Response: Changed as requested.

Supplementary: Data analysis section "Statistical significance was determined either with unpaired Student's t-tests using GraphPad Prism (version 9) program" is there a typo here or do authors utilize another approach to determine statistical significance
Response: Corrected

Fig S6a perhaps a violin plot may be better to represent this data.
Response: Changed as requested

Reviewer #3 (Remarks to the Author):

In this manuscript, the authors evaluate the anti-plasmodial activity of several antimalarial compounds on Plasmodium falciparum asexual and sexual stages (gametocytes). By performing a systematic meta-analysis of published IC50 data from 118 compounds currently used and in clinical development, they conclude that most compounds exhibit a loss of activity between asexual and sexual parasites. To circumvent the variations between different assay platforms and conditions used in different laboratories to calculate the IC50s, they confirm these findings on 10 compounds using the same assay and the same luciferase reporter line. They further investigate the confounding factors contributing to the reduced drug activity against gametocytes. Their data show that for most drugs, the activity profiles correlate with the drug target abundance between the stages, but this is not the case for compounds losing >10-fold of their activity in gametocytes. Using cheminformatics and machine learning, they

identified physicochemical properties of the drug (topological polar surface area, solubility, and lipophilicity) as the most critical chemical features that contribute to compound efficacy. Accordingly, they provide evidence that PSAC and aquaglyceroporin, which are absent in mature gametocytes, contribute to compound uptake in asexual stages. Finally, they show that the different lipid composition of the infected erythrocyte in asexual parasites compared to gametocytes results in reduced diffusion of compounds into gametocytes. They conclude that the loss in activity of compounds against gametocytes correlates to specific requirements in the physicochemical properties to allow for effective diffusion across membranes. The authors address an important question that will likely attract interest in the scientific community and that must be considered for development of new antimalarials. Although several conclusions made by the authors only confirm what was already known in the field, there are several novel aspects to these findings. The experiments performed and reported were convincing. In most cases, the data provide the evidence needed for the authors to make their conclusions.

Major comments
- The results reported lines 399-401 do not correspond to the data shown in figure 6D: there is not significant changes for MMV048 at all stages, nor for DHA and Blasticidine at immature gametocyte stages.
Response: Thank you for highlighting this, we corrected the graphs to show the significance in Figure 6D (which was in shown in the supplementary figures, but somehow lost in the final draft of the figures, we apologise).

- This is surprising that inhibition of PSAC with furosemide decreases the activity for MMV048 and DHA in asexual stages, since PSAC is not active in ring stages, which represent 95% of the asexual blood stage parasites used in this study.
Response: No, the experiments were performed on trophozoites where PSAC is active. This was stated in the methods and iterated in the figure legends and text.

- All the statistical analyses are done with unpaired Student's t-tests. However, ANOVA should be more appropriate in several figures (for example in Figure 1, 2 and 6 where several conditions are compared).
Response: The statistical tests were re-evaluated and checked for normal distribution and changed as suggested.

- The authors address how the abundance of parasite protein targets correlates with the activity profiles of the drugs; however, they do not address or at least discuss how the different concentration of hemoglobin in the red blood cell between asexual (ring stages) and gametocytes contributes to this phenomenon. It is an important point since the mechanism of action of several antimalarials, including chloroquine or artemisinin derivatives, relies on the degradation of hemoglobin.
Response: This is correct. We added a sentence to this discussion as the influence of inhibition of hemoglobin degradation cannot be discarded to explain the loss of activity. However, aside from the quinolines and artemisinins classes, the analysis we performed was specific to drugs that has specific protein target associations, allowing us to perform the correlations focussed on specific target protein abundance levels. Indeed, Figure 4 and Supplementary file S4 indeed only show that the analysis did not include compounds such as chloroquine or artemisinin where these haemoglobin targeting (or pleiotropic effects) is ascribed as mode-of-action.

Minor comments
- the legend of Figure 1 is not clear
Response: This was corrected.

- Figure 2B: Primer location should be shown on the schema
Response: Added as requested.

- l 240: The transgenic reporter line expressing tdTomato red fluorescence under the Pfef1α promotor is not described in the reference 36 (Portugalizia et al, 2019).
Response: Apologies, the correct reference was included.

- Supp Fig S4C: All gametocytes shown in this figure look more like stage IV gametocytes than mature stages V. Also, the morphology of Cabamaquine-treated parasites does not appear different than the other ones (at least on the shown pictures).

Response: This was corrected to more clearly indicate stage V gametocytes.

- l 452: the sentence "compounds that prefer ABS parasites are also found outside this preferred absorption model" should be rephrase, since only 6 out of 31 compounds are outside of the ellipse.
Response: Corrected as suggested

- l 486: replace gametogenesis by gametocytogenesis
Response: Corrected.

- in the discussion, l 514 – 517, the authors discuss the caveat that compounds would need to be able to access erythroblastic islands where gametocyte sequester. They should also discuss that the activity profiles of several drugs could be modified if gametocytes develop within erythroblasts, as it has been suggested in PMID: 25009232 and PMID: 32589714.
Response: This is correct, and has been included in the discussion.

A point-by-point response to the last minor requested changes from reviewer 1:

This revised manuscript is substantially improved both in clarity and data presentation. Most of my previous issues have been satisfactorily addressed. Below are some outstanding issues, most of which are relatively minor and could be addressed with simple clarification of sentence structure.

Line 112: "As expected" – why was this expected? Also low dose primaquine is an effective anti-gametocyte treatment. However it needs liver metabolism to the active form which is missing from in vitro assays. Probably not the best antimalarial to be discussing in this context.
Response: The 'as expected' comment refers to the previously published activities of known antimalarials as referenced. We agree that primaquine indeed has to be activated by liver metabolism, and we removed reference to this in this statement.

Line 117: Confusing sentence structure – as read, it appears that CHX is a 4AQ or and 8AQ.
Response: Thank you for picking this up. We corrected this to read: …..aminoquinolines and also the diamine pentamidine and antifungal cycloheximide…..

Line 152L/Fig 2A: Please add ML10 to the list of compounds in cluster 6 seeing as the dose response curve is included in Fig 2B.
Response: Thank you, we included this in the figure's cluster 6 list.

Lines 225-230: "Under these conditions, gametocytogenesis still occurred, similar to the sexual commitment rates for most of the compounds evaluated…." What does "similar" mean? Does this statement mean that there was no statistically significant difference? If so, it would be good to state this as it would strengthen argument.
Response: We corrected the statement and included statistical data to show that indeed gametocytes could still be produced from the majority of compounds to strengthen the argument.

Lines 233-235: Are the authors demonstrating that a 130-fold loss in activity results in only a doubling of immature gametocytes? Whilst any increase is clearly undesirable, how concerning actually is this? If it takes >100-fold loss of activity to increase the population by 100% this does not seem that much to me.
Response: We thank the reviewer for picking this up, and yes, we recognize that it is unclear. We apologise since this sentence was a carryover from the previous version, where we specifically referred to one compound. However, with the re-evaluation under the 'therapeutically relevant concentrations' that we provided in the revised version, this statement does not hold anymore and was deleted. We apologise for this omission, and checked all other statements for accuracy.

Lines 239-253/Fig3: Did not reproduce well in the pdf and some of the data is cut off from Fig2C and E so it is not possible to comment on this section.
Response: We apologise for this, the figure seemed fine on our end, even in the final pdf submitted. We checked again for its accuracy.

Line 271: "Taken together, only compounds that cannot effectively target immature gametocytes will not necessarily prevent onward transmission of the parasites." This is a very confusing sentence structure – please clarify. Does it mean, "Taken together, compounds that are ineffective at targeting immature gametocytes may also be ineffective at preventing onward parasites transmission to the mosquito"?
Response: We corrected the sentence structure to more clearly reflect the meaning, which is exactly as this reviewer perceived it to be.

Line 378: Are you sure PfAQP is absent from mature gametocytes? This needs a citation. In Pb, genetically tagged AQP is present in gametocytes (but higher expression in asexuals): https://doi.org/10.1016/j.ijpara.2012.10.006
Response: We changed this statement in the text to reflect the fact that PfAQP abundance is vastly reduced in mature gametocytes compared to ABS, as evident from transcriptome (van Biljon, 2019) and proteome (Silvestrini 2010) data. This correlates to the statement made by the reviewer pertaining to the situation in Pb, which we also now referred to in text.