

Mimicry diversification in *Papilio dardanus* via a genomic inversion in the regulatory region of *engrailed–invected*

Martijn J. T. N. Timmermans, Amrita Srivathsan, Steve Collins, Rudolf Meier and Alfried P. Vogler

Article citation details

Proc. R. Soc. B **287**: 20200443.
<http://dx.doi.org/10.1098/rspb.2020.0443>

Review timeline

Original submission: 14 January 2020
1st revised submission: 26 February 2020
2nd revised submission: 31 March 2020
Final acceptance: 31 March 2020

Note: Reports are unedited and appear as submitted by the referee. The review history appears in chronological order.

Review History

RSPB-2020-0079.R0 (Original submission)

Review form: Reviewer 1

Recommendation

Accept with minor revision (please list in comments)

Scientific importance: Is the manuscript an original and important contribution to its field?

Good

General interest: Is the paper of sufficient general interest?

Excellent

Quality of the paper: Is the overall quality of the paper suitable?

Good

Is the length of the paper justified?

Yes

Should the paper be seen by a specialist statistical reviewer?

No

Do you have any concerns about statistical analyses in this paper? If so, please specify them explicitly in your report.

No

It is a condition of publication that authors make their supporting data, code and materials available - either as supplementary material or hosted in an external repository. Please rate, if applicable, the supporting data on the following criteria.

Is it accessible?

Yes

Is it clear?

Yes

Is it adequate?

Yes

Do you have any ethical concerns with this paper?

No

Comments to the Author

In this study, the authors assembled the whole genome sequence of *Papilio dardanus* and conducted genomic comparison based on resequencing of different forms of female wing pattern. As a result, the authors found that the genomic region flanking the previously recognized mimicry genes had elevated *Fst* and LD (linkage disequilibrium) when genomic sequences were compared between two female wing forms. Importantly, there seemed to be an inversion in the region of elevated *Fst*/LD between two wing forms. The authors argued that the region may involve cis-regulatory region, and allelic differences protected by the inversion may control the occurrence of two female wing pattern phenotypes. In *Papilio* butterflies, genomic study of the mimicry supergene locus involving doublesex gene has revealed both the presence and absence of inversion, suggesting that the mechanism maintaining the allelic differences can be variable. The present study reports the possible role of inversion in maintaining female mimicry polymorphism controlled by different genes (engrailed-inverted). This finding suggests that occasional chromosomal inversion in different genomic regions may have caused evolution of different mimicry supergenes in *Papilio* butterflies.

Overall I agree that the contents of the manuscript are sound and important in the study of butterfly mimicry evolution. However, there are some weak points in the genomic analyses. In particular, the Hc allele sequence has not been determined clearly, so the actual sequence differences between Hh and Hc allele have not been depicted. Accordingly, the evidence of inversion appears not to be decisive enough (e.g., lack of Hrr plot showing different orientation between Hc and Hh). Therefore, I recommend that the authors include some cautious discussion about the present resolution of the genomic structure and allelic sequences of the putative supergene locus.

Minor points:

P. 7, second paragraph mentions "The South African f. cenea specimen ... homozygous for the inversion (Hc/Hc)", but in Table 1, its genotype is indicated as Hc/? Please clarify. If that individual was truly Hc/Hc, it could be used to determine Hc sequence.

Please indicate unit of base pairs (bp, kbp, or Mbp?) in figures 2 and 3, and add scale of bp in figure 4.

Review form: Reviewer 2

Recommendation

Major revision is needed (please make suggestions in comments)

Scientific importance: Is the manuscript an original and important contribution to its field?

Good

General interest: Is the paper of sufficient general interest?

Good

Quality of the paper: Is the overall quality of the paper suitable?

Acceptable

Is the length of the paper justified?

Yes

Should the paper be seen by a specialist statistical reviewer?

No

Do you have any concerns about statistical analyses in this paper? If so, please specify them explicitly in your report.

No

It is a condition of publication that authors make their supporting data, code and materials available - either as supplementary material or hosted in an external repository. Please rate, if applicable, the supporting data on the following criteria.

Is it accessible?

Yes

Is it clear?

Yes

Is it adequate?

Yes

Do you have any ethical concerns with this paper?

No

Comments to the Author

By providing a draft genome of *P. dardanus* and by resequencing several individuals displaying different wing colour pattern within this species involved in Batesian mimicry, this manuscript describes a novel case of inversion polymorphism controlling variations in mimetic colour patterns. The uncovered inversion is a putative regulatory region of the genes engrailed and invected. These genes were already identified as functional candidates involved in colour pattern diversification within this species, but the implication of this regulatory region was ignored. By comparing allelic divergence in sympatric and allopatric forms, this manuscript also brings relevant evidences for the role of inversions in the evolution of adaptive polymorphisms. Although I think the results presented here are really interesting, the manuscript sometimes lacks of precision and that the discussion is probably too long and speculative. I also have a number of comments and questions listed below that should be carefully adressed.

Specific comments:

It would facilitate the reviewers' job if the lines of the manuscript were numbered.

Title

I think one of the major novelties of the paper is that the inversion is located in the regulatory region of the engrailed-inverted locus, I would thus recommend changing the title to highlight this original findings. As you mentioned yourself in the introduction, variations in engrailed were already identified as associated with colour pattern variations in *P. dardanus* in a previous study (Thompson et al. 2014).

Abstract

‘multiple forms’, I think it is better to be more precise, if I understood correctly you focused on three forms in this study.

‘We hypothesize that the inversion was instrumental in the origin of the novel phenotype by preventing recombination in the evolutionary process of wing pattern refinement.’ I do not think the results brought by this study can clarify the evolutionary steps leading to the evolution of new mimetic alleles (whether the inversion occurred simultaneously or followed the mutations allowing a new phenotype to emerge for instance). The functional links between the inverted regions and the nucleotide variations it contained with the inverted-engrailed locus, and to a larger extent, to the colour pattern variations are not identified. I think it would be more relevant to contrast the evolution of inversions in the monomorphic and polymorphic populations therefore showing that inversions are indeed favored by balancing selection. This is a better argument drawn from the results presented here.

Introduction

‘survival may be increased if species mimic multiple unpalatable phenotypes’ I think this sentence provides a very bad argument to the evolution of polymorphic mimicry involving group selection. A species is not mimicking anything ! The evolution of new mimetic forms is favored by the negative frequency-dependent selection assumed in Batesian mimics, because rare forms of Batesian mimics have a greater fitness than frequent ones. This should be clearly stated ! Moreover, it is really important to clarify that this drives the persistence of polymorphism within populations (i.e. multiple forms within a given geographic region) rather than at the species level, because as you mentioned in the manuscript, there are also some monomorphic populations of *P. dardanus* as well, where the selection regime and the genetic architecture of colour pattern alleles might differ.

Replace ‘Only the female sex is mimetic’ by ‘Females only are mimetic’

It would be better to avoid using ‘selection for’ that sounds quite finalistic, you may replace ‘negative frequency dependent selection for mimetic protection’ by ‘negative frequency dependent selection generated by predator behavior.’

Is the locus H restricted to *P. dardanus* ? If so, you may be more precise by adding : ‘In *P. dardanus*,’ before ‘Wing colours and patterns are controlled by a single Mendelian locus, H, whose various alleles segregate according to a well-defined hierarchy of dominances’. Remove the ‘s’ after ‘dominance in the previous sentence.

‘Segregation analysis in pedigree-broods using AFLP [17] and population genetics [18] have shown that the mimicry switch in *P. dardanus* is genetically linked to the engrailed-inverted locus, a region that codes for two paralogous homeodomain transcription factors involved in anterior-posterior patterning [19]’. As mentioned earlier, the conclusions of these previously published papers seem rather similar to the title of this manuscript, the title should definitely be revised.

‘Genetic variation especially in the first exon of engrailed shows strong association to specific mimetic wing phenotypes, and a high proportion of non-synonymous mutations together with

high allelic divergence suggest balancing selection on this region [20]', does this study focus on the same forms as those studied here? This should be clarified.

'selection to avoid maladaptive intermediates' As already explained, it would be better to replace 'selection to avoid' by 'selection against non-mimetic intermediates'

'A critical aspect of this process is that genetic recombination among functional sites is low, to avoid the formation of intermediates with lower fitness.' Replace 'to avoid' by 'preventing'.

'a mechanism to enhance linkage of co-adapted mutations'. Replace 'a mechanism to enhance' by 'a mechanism increasing'.

'inversions may only arise at a late stage in the evolution of mimicry loci'. It would be useful to provide alternative hypothesis regarding the selection regime and the sympatry with different mimicry rings.

'the evolution of mimicry loci under selection for tight linkage' Replace 'for tight linkage' by 'promoting tight linkage'.

'To understand the genetic architecture underlying polymorphic mimicry in *P. dardanus* we use comparative genomics of multiple female forms'. It would be clearer to state that you focused on 3 forms only.

'This form occurs throughout the entire range of *P. dardanus* and is recessive to all others.' This is misleading, this form is absent in Madagascar, right? It would also be important to highlight that the male-like form 'meriones' is only observed in Madagascar. Overall, because the sympatry/allopatry relationships among forms are important to understand the selection regime acting on colour pattern variations and therefore on the underlying genetic architecture, these relationships should be explained very accurately in the introduction.

Results:

'A draft genome sequence was constructed that allows for the genome wide analysis of structural rearrangements and genetic variation.' Could you estimate the frequency of inversions throughout the genome using pairwise comparisons between the sequenced individuals? This would give a hint on how much inversions occur frequently or not? How likely is it that such inversion occurs close to the gene engrailed just by neutral evolution?

Figure 1: It should be clarified that the male phenotype is the last one on the right of the second row of the Top panel.

'Within the region, sequence divergence (estimated as p-distance) between the hippoconides individuals and the reference genome sequence was slightly higher than for the cenea individuals and the reference sequence' This result seems surprising, I guess the reference sequence stem from the reference genome which is based on a *P. dardanus tibullus*, i.e. an homozygote with two hippoconides alleles. How can you explain that the genetic divergence in the region putatively involved in colour pattern variations is higher between two hippoconides alleles rather than between hippoconides and cenea?

'The inversion was not found in the non-mimetic 'male-like' *P. d. meriones* from Madagascar, which indicates that the bottom-recessive mimetic *f. hippoconides* has the same arrangement as this male-like form'. The hippoconides allele is thus likely to have evolved from an ancestral meriones allele through successive mutations promoted by mimicry. Could you provide an estimation of the level of nucleotide divergence between these two alleles as compared to the hippoconides-cenea pair?

'has the same arrangement as this male-like form, and that this specific arrangement is therefore ancestral.' That makes perfect sense, but did you check that this is indeed the same gene order as in the other Lepidoptera? It is worth to clarify this.

'The four f. cenea specimens represented two distinct subspecies from Kenya and South Africa.' I guess that the 3 specimens from *P. dardanus polytrophus cenea* are from Kenya, and the specimen from South Africa is a *P. dardanus cenea cenea*. This should be clarified in the text of the manuscript. Moreover, since the question of sympatry among morphs is important, it would be worth adding a column in table 1 providing the geographic origin of each sample. It is really important readers without prior knowledge on the geographic distribution of mimetic forms of *P. dardanus* could understand the sympatry/allopatry relationships more easily.

'The sequence data furthermore indicated that the Kenyan specimens carried a second non-inverted allele (Table 1).' This sentence is a bit unclear to me, I guess you mean that the 3(?) sequenced individuals from Kenya displaying the cenea pattern (*P. dardanus polytrophus cenea?*) are all heterozygotes with a cenea allele and a hippoconides allele?

Discussion

'This supports the supergene hypothesis' I am not sure the results obtained here are bringing evidence for the supergene hypothesis, because the data presented here do not pinpoint the causative mutations, and we do not know whether several independent mutations are required for the switch between cenea and hippoconides to happen. Once the inversion has occurred, the lack of recombination promotes the divergence between alleles independently from selection exerted by mimicry, making it highly challenging to discriminate mutations actually involved in the phenotypic divergence. The arguments you provide in the following sections of the discussion are also in favor of the pleiotropic effects of this putatively-regulatory region on engrailed, invected and miR-2768. Such pleiotropic effects are also not demonstrated by the data shown here, preventing to robustly conclude on the supergene hypothesis.

'Invected also contains an intronic microRNA (miR-2768) conserved in Lepidoptera (Fig. 3; Supplementary Figure 12), which has been shown to downregulate cubitus interruptus (ci), a gene that regulates patterning of the wing primordia via the hedgehog signaling pathway in nymphalid butterflies [45].' I think the discussion is generally too long and speculative, regulatory regions can evolve fast and be co-opted, so that comparisons of their role in different species require specific functional studies that are not the aim of the manuscript presented here. I would recommend shortening the speculations of the functional effect of the inversions and focus more on the association between inversion polymorphisms and persistence of several mimetic forms within population cause by nFDS. Your results are really pointing out that inversions can be especially promoted by nFDS, because the lack of recombination allows divergent alleles to be maintained. It would also be interesting to highlight that polymorphism mimicry in *P. polytes* and *P. dardanus* does involve inversions, but that these inversions are different. The location of the inversion uncovered here with respect to the inversions involved in polymorphic mimicry in other Lepidoptera should be mentioned, and the repeated independent evolution of these architectures should be discussed in light of the occurrence of inversions throughout the genomes.

'obviating the need for recombination-suppressing mechanisms' This is unclear: the selection regime involved should be properly explained better. I guess the hippoconides allele benefited from a greater advantage over the non-mimetic allele meriones at some point, so that selection promoting mimicry has led to the evolution of this mimetic form, and completely replaced the ancestral allele in the African populations. Could you provide insights on why this happened on the African continent and not in Madagascar? Does the models of the forms hippoconides and cenea are absent in Madagascar and abundant elsewhere?

'Determination of the phenotype likely works in concert with other changes in the engrailed-invected region,' I am not sure I understand this hypothesis given that you did find

recombinations in you broods occurring in the flanking region, making it hard to reconcile the concerted evolution of the inversion and the nucleotide variation in engrailed.

'Inversions are not necessary, but helpful, and generally have the function to reduce recombination,' I strongly disagree with this sentence, the inversion do not have a function per se, it can be promoted if it captures an alleles under positive selection because of mimicry and be maintained in the population because it does not lead to recombinant alleles with poor fitness, this should be explained more clearly.

Decision letter (RSPB-2020-0079.R0)

10-Feb-2020

Dear Dr Timmermans:

I am writing to inform you that your manuscript RSPB-2020-0079 entitled "Mimicry diversification in *Papilio dardanus* via a genomic inversion in the engrailed-inverted locus" has, in its current form, been rejected for publication in Proceedings B.

This action has been taken on the advice of referees, who have recommended that substantial revisions are necessary. With this in mind we would be happy to consider a resubmission, provided the comments of the referees are fully addressed. However please note that this is not a provisional acceptance.

The resubmission will be treated as a new manuscript. However, we will approach the same reviewers if they are available and it is deemed appropriate to do so by the Editor. Please note that resubmissions must be submitted within six months of the date of this email. In exceptional circumstances, extensions may be possible if agreed with the Editorial Office. Manuscripts submitted after this date will be automatically rejected.

Please find below the comments made by the referees, not including confidential reports to the Editor, which I hope you will find useful. If you do choose to resubmit your manuscript, please upload the following:

- 1) A 'response to referees' document including details of how you have responded to the comments, and the adjustments you have made.
- 2) A clean copy of the manuscript and one with 'tracked changes' indicating your 'response to referees' comments document.
- 3) Line numbers in your main document.

To upload a resubmitted manuscript, log into <http://mc.manuscriptcentral.com/prsb> and enter your Author Centre, where you will find your manuscript title listed under "Manuscripts with Decisions." Under "Actions," click on "Create a Resubmission." Please be sure to indicate in your cover letter that it is a resubmission, and supply the previous reference number.

Please note that this decision may (or may not) have taken into account confidential comments.

In your revision process, please take a second look at how open your science is; our policy is that **ALL** (maximally inclusive) data involved with the study should be made openly accessible, fully enabling re-use, replication and transparency-- see: <https://royalsociety.org/journals/ethics-policies/data-sharing-mining/>
Insufficient sharing of data can delay or even cause rejection of a paper.

Sincerely,
Professor John Hutchinson, Editor
mailto: proceedingsb@royalsociety.org

Associate Editor

Comments to Author:

I really enjoyed reading this manuscript, it provides solid evidence for the interface between the evolution of regulatory regions and the animal ecology. I'd like to congratulate the authors. I'd like to encourage them to consider great suggestions from the referees. I'd also like to suggest to merge figure 3 and 4 in a single figure.

Reviewer(s)' Comments to Author:

Referee: 1

Comments to the Author(s)

In this study, the authors assembled the whole genome sequence of *Papilio dardanus* and conducted genomic comparison based on resequencing of different forms of female wing pattern. As a result, the authors found that the genomic region flanking the previously recognized mimicry genes had elevated F_{st} and LD (linkage disequilibrium) when genomic sequences were compared between two female wing forms. Importantly, there seemed to be an inversion in the region of elevated F_{st} /LD between two wing forms. The authors argued that the region may involve cis-regulatory region, and allelic differences protected by the inversion may control the occurrence of two female wing pattern phenotypes. In *Papilio* butterflies, genomic study of the mimicry supergene locus involving doublesex gene has revealed both the presence and absence of inversion, suggesting that the mechanism maintaining the allelic differences can be variable. The present study reports the possible role of inversion in maintaining female mimicry polymorphism controlled by different genes (engrailed-inverted). This finding suggests that occasional chromosomal inversion in different genomic regions may have caused evolution of different mimicry supergenes in *Papilio* butterflies.

Overall I agree that the contents of the manuscript are sound and important in the study of butterfly mimicry evolution. However, there are some weak points in the genomic analyses. In particular, the Hc allele sequence has not been determined clearly, so the actual sequence differences between Hh and Hc allele have not been depicted. Accordingly, the evidence of inversion appears not to be decisive enough (e.g., lack of Hrr plot showing different orientation between Hc and Hh). Therefore, I recommend that the authors include some cautious discussion about the present resolution of the genomic structure and allelic sequences of the putative supergene locus.

Minor points:

P. 7, second paragraph mentions "The South African *f. cenea* specimen ... homozygous for the inversion (Hc/Hc)", but in Table 1, its genotype is indicated as Hc/? Please clarify. If that individual was truly Hc/Hc, it could be used to determine Hc sequence.

Please indicate unit of base pairs (bp, kbp, or Mbp?) in figures 2 and 3, and add scale of bp in figure 4.

Referee: 2

Comments to the Author(s)

By providing a draft genome of *P. dardanus* and by resequencing several individuals displaying different wing colour pattern within this species involved in Batesian mimicry, this manuscript describes a novel case of inversion polymorphism controlling variations in mimetic colour patterns. The uncovered inversion is a putative regulatory region of the genes *engrailed* and *invected*. These genes were already identified as functional candidates involved in colour pattern diversification within this species, but the implication of this regulatory region was ignored. By

comparing allelic divergence in sympatric and allopatric forms, this manuscript also brings relevant evidences for the role of inversions in the evolution of adaptive polymorphisms. Although I think the results presented here are really interesting, the manuscript sometimes lacks of precision and that the discussion is probably too long and speculative. I also have a number of comments and questions listed below that should be carefully adressed.

Specific comments:

It would facilitate the reviewers' job if the lines of the manuscript were numbered.

Title

I think one of the major novelties of the paper is that the inversion is located in the regulatory region of the engrailed-inverted locus, I would thus recommend changing the title to highlight this original findings. As you mentioned yourself in the introduction, variations in engrailed were already identified as associated with colour pattern variations in *P. dardanus* in a previous study (Thompson et al. 2014).

Abstract

'multiple forms', I think it is better to be more precise, if I understood correctly you focused on three forms in this study.

'We hypothesize that the inversion was instrumental in the origin of the novel phenotype by preventing recombination in the evolutionary process of wing pattern refinement.' I do not think the results brought by this study can clarify the evolutionary steps leading to the evolution of new mimetic alleles (whether the inversion occurred simultaneously or followed the mutations allowing a new phenotype to emerge for instance). The functional links between the inverted regions and the nucleotide variations it contained with the inverted-engrailed locus, and to a larger extent, to the colour pattern variations are not identified. I think it would be more relevant to contrast the evolution of inversions in the monomorphic and polymorphic populations therefore showing that inversions are indeed favored by balancing selection. This is a better argument drawn from the results presented here.

Introduction

'survival may be increased if species mimic multiple unpalatable phenotypes' I think this sentence provides a very bad argument to the evolution of polymorphic mimicry involving group selection. A species is not mimicking anything ! The evolution of new mimetic forms is favored by the negative frequency-dependent selection assumed in Batesian mimics, because rare forms of Batesian mimics have a greater fitness than frequent ones. This should be clearly stated ! Moreover, it is really important to clarify that this drives the persistence of polymorphism within populations (i.e. multiple forms within a given geographic region) rather than at the species level, because as you mentioned in the manuscript, there are also some monomorphic populations of *P. dardanus* as well, where the selection regime and the genetic architecture of colour pattern alleles might differ.

Replace 'Only the female sex is mimetic' by 'Females only are mimetic'

It would be better to avoid using 'selection for' that sounds quite finalistic, you may replace 'negative frequency dependent selection for mimetic protection' by 'negative frequency dependent selection generated by predator behavior.'

Is the locus H restricted to *P. dardanus* ? If so, you may be more precise by adding : 'In *P. dardanus*,' before 'Wing colours and patterns are controlled by a single Mendelian locus, H, whose various alleles segregate according to a well-defined hierarchy of dominances'. Remove the 's' after 'dominance in the previous sentence.

'Segregation analysis in pedigree-broods using AFLP [17] and population genetics [18] have shown that the mimicry switch in *P. dardanus* is genetically linked to the engrailed-inverted

locus, a region that codes for two paralogous homeodomain transcription factors involved in anterior-posterior patterning [19]'. As mentioned earlier, the conclusions of these previously published papers seem rather similar to the title of this manuscript, the title should definitely be revised.

'Genetic variation especially in the first exon of engrailed shows strong association to specific mimetic wing phenotypes, and a high proportion of non-synonymous mutations together with high allelic divergence suggest balancing selection on this region [20]', does this study focus on the same forms as those studied here ? This should be clarified.

'selection to avoid maladaptive intermediates' As already explained, it would be better to replace 'selection to avoid' by 'selection against non-mimetic intermediates'

'A critical aspect of this process is that genetic recombination among functional sites is low, to avoid the formation of intermediates with lower fitness.' Replace 'to avoid' by 'preventing'.

'a mechanism to enhance linkage of co-adapted mutations'. Replace 'a mechanism to enhance' by 'a mechanism increasing'.

'inversions may only arise at a late stage in the evolution of mimicry loci'. It would be useful to provide alternative hypothesis regarding the selection regime and the sympatry with different mimicry rings.

'the evolution of mimicry loci under selection for tight linkage' Replace 'for tight linkage' by 'promoting tight linkage'.

'To understand the genetic architecture underlying polymorphic mimicry in *P. dardanus* we use comparative genomics of multiple female forms'. It would be clearer to state that you focused on 3 forms only.

'This form occurs throughout the entire range of *P. dardanus* and is recessive to all others.' This is misleading, this form is absent in Madagascar, right ? It would also be important to highlight that the male-like form 'meriones' is only observed in Madagascar. Overall, because the sympatry/allopatry relationships among forms are important to understand the selection regime acting on colour pattern variations and therefore on the underlying genetic architecture, these relationships should be explained very accurately in the introduction.

Results:

'A draft genome sequence was constructed that allows for the genome wide analysis of structural rearrangements and genetic variation.' Could you estimate the frequency of inversions throughout the genome using pairwise comparisons between the sequenced individuals? This would give a hint on how much inversions occur frequently or not? How likely is it that such inversion occurs close to the gene engrailed just by neutral evolution?

Figure 1: It should be clarified that the male phenotype is the last one on the right of the second row of the Top panel.

'Within the region, sequence divergence (estimated as p-distance) between the hippoconides individuals and the reference genome sequence was slightly higher than for the cenea individuals and the reference sequence' This result seems surprising, I guess the reference sequence stem from the reference genome which is based on a *P. dardanus tibullus*, i.e. an homozygote with two hippoconides alleles. How can you explain that the genetic divergence in the region putatively involved in colour pattern variations is higher between two hippoconides alleles rather than between hippoconides and cenea?

'The inversion was not found in the non-mimetic 'male-like' *P. d. meriones* from Madagascar,

which indicates that the bottom-recessive mimetic f. *hippocoonides* has the same arrangement as this male-like form'. The *hippocoonides* allele is thus likely to have evolved from an ancestral *meriones* allele through successive mutations promoted by mimicry. Could you provide an estimation of the level of nucleotide divergence between these two alleles as compared to the *hippocoonides-cenea* pair?

'has the same arrangement as this male-like form, and that this specific arrangement is therefore ancestral.' That makes perfect sense, but did you check that this is indeed the same gene order as in the other *Lepidoptera*? It is worth to clarify this.

'The four f. *cenea* specimens represented two distinct subspecies from Kenya and South Africa.' I guess that the 3 specimens from *P. dardanus polytrophus cenea* are from Kenya, and the specimen from South Africa is a *P. dardanus cenea cenea*. This should be clarified in the text of the manuscript. Moreover, since the question of sympatry among morphs is important, it would be worth adding a column in table 1 providing the geographic origin of each sample. It is really important readers without prior knowledge on the geographic distribution of mimetic forms of *P. dardanus* could understand the sympatry/allopatriy relationships more easily.

'The sequence data furthermore indicated that the Kenyan specimens carried a second non-inverted allele (Table 1).' This sentence is a bit unclear to me, I guess you mean that the 3(?) sequenced individuals from Kenya displaying the *cenea* pattern (*P. dardanus polytrophus cenea*?) are all heterozygotes with a *cenea* allele and a *hippocoonides* allele?

Discussion

'This supports the supergene hypothesis' I am not sure the results obtained here are bringing evidence for the supergene hypothesis, because the data presented here do not pinpoint the causative mutations, and we do not know whether several independent mutations are required for the switch between *cenea* and *hippocoonides* to happen. Once the inversion has occurred, the lack of recombination promotes the divergence between alleles independently from selection exerted by mimicry, making it highly challenging to discriminate mutations actually involved in the phenotypic divergence. The arguments you provide in the following sections of the discussion are also in favor of the pleiotropic effects of this putatively-regulatory region on *engrailed*, *invected* and *miR-2768*. Such pleiotropic effects are also not demonstrated by the data shown here, preventing to robustly conclude on the supergene hypothesis.

'*Invected* also contains an intronic microRNA (*miR-2768*) conserved in *Lepidoptera* (Fig. 3; Supplementary Figure 12), which has been shown to downregulate *cubitus interruptus (ci)*, a gene that regulates patterning of the wing primordia via the hedgehog signaling pathway in nymphalid butterflies [45].' I think the discussion is generally too long and speculative, regulatory regions can evolve fast and be co-opted, so that comparisons of their role in different species require specific functional studies that are not the aim of the manuscript presented here. I would recommend shortening the speculations of the functional effect of the inversions and focus more on the association between inversion polymorphisms and persistence of several mimetic forms within population caused by nFDS. Your results are really pointing out that inversions can be especially promoted by nFDS, because the lack of recombination allows divergent alleles to be maintained. It would also be interesting to highlight that polymorphism mimicry in *P. polytes* and *P. dardanus* does involve inversions, but that these inversions are different. The location of the inversion uncovered here with respect to the inversions involved in polymorphic mimicry in other *Lepidoptera* should be mentioned, and the repeated independent evolution of these architectures should be discussed in light of the occurrence of inversions throughout the genomes.

'obviating the need for recombination-suppressing mechanisms' This is unclear: the selection regime involved should be properly explained better. I guess the *hippocoonides* allele benefited from a greater advantage over the non-mimetic allele *meriones* at some point, so that selection promoting mimicry has led to the evolution of this mimetic form, and completely replaced the

ancestral allele in the African populations. Could you provide insights on why this happened on the African continent and not in Madagascar? Does the models of the forms hippoconides and cenea are absent in Madagascar and abundant elsewhere?

'Determination of the phenotype likely works in concert with other changes in the engrailed-invested region,' I am not sure I understand this hypothesis given that you did find recombinations in you broods occurring in the flanking region, making it hard to reconcile the concerted evolution of the inversion and the nucleotide variation in engrailed.

'Inversions are not necessary, but helpful, and generally have the function to reduce recombination,' I strongly disagree with this sentence, the inversion do not have a function per se, it can be promoted if it captures an alleles under positive selection because of mimicry and be maintained in the population because it does not lead to recombinant alleles with poor fitness, this should be explained more clearly.

Author's Response to Decision Letter for (RSPB-2020-0079.R0)

See Appendix A.

RSPB-2020-0443.R0

Review form: Reviewer 2

Recommendation

Accept with minor revision (please list in comments)

Scientific importance: Is the manuscript an original and important contribution to its field?

Excellent

General interest: Is the paper of sufficient general interest?

Good

Quality of the paper: Is the overall quality of the paper suitable?

Excellent

Is the length of the paper justified?

Yes

Should the paper be seen by a specialist statistical reviewer?

No

Do you have any concerns about statistical analyses in this paper? If so, please specify them explicitly in your report.

No

It is a condition of publication that authors make their supporting data, code and materials available - either as supplementary material or hosted in an external repository. Please rate, if applicable, the supporting data on the following criteria.

Is it accessible?

No

Is it clear?

Yes

Is it adequate?

Yes

Do you have any ethical concerns with this paper?

No

Comments to the Author

This is a very interesting paper bringing new insights on the role of inversions on polymorphic mimicry. I think the authors have correctly addressed my previous comments although I was not able to see the figures and tables in this revised version (I assumed they have not been modified since the previous version ? Except the new column I suggested on table 1 I guess ?).

You may also want to replace 'Being confined to the regulatory region, the inversion may or may not have a functional role in the expression of the novel phenotype' by 'Although the functional role of this inversion, confined to the regulatory region, in the expression of the novel phenotype is currently unknown...' on line 33 of the abstract.

Decision letter (RSPB-2020-0443.R0)

20-Mar-2020

Dear Dr Timmermans

I am pleased to inform you that your manuscript RSPB-2020-0443 entitled "Mimicry diversification in *Papilio dardanus* via a genomic inversion in the regulatory region of engrailed-invested" has been accepted for publication in Proceedings B. Congratulations!!

The referee(s) have recommended publication, but also suggest some minor revisions to your manuscript. Therefore, I invite you to respond to the referee(s)' comments and revise your manuscript. Because the schedule for publication is very tight, it is a condition of publication that you submit the revised version of your manuscript within 7 days. If you do not think you will be able to meet this date please let us know.

To revise your manuscript, log into <https://mc.manuscriptcentral.com/prsb> and enter your Author Centre, where you will find your manuscript title listed under "Manuscripts with Decisions." Under "Actions," click on "Create a Revision." Your manuscript number has been appended to denote a revision. You will be unable to make your revisions on the originally submitted version of the manuscript. Instead, revise your manuscript and upload a new version through your Author Centre.

When submitting your revised manuscript, you will be able to respond to the comments made by the referee(s) and upload a file "Response to Referees". You can use this to document any changes you make to the original manuscript. We require a copy of the manuscript with revisions made since the previous version marked as 'tracked changes' to be included in the 'response to referees' document.

Before uploading your revised files please make sure that you have:

- 1) A text file of the manuscript (doc, txt, rtf or tex), including the references, tables (including captions) and figure captions. Please remove any tracked changes from the text before submission. PDF files are not an accepted format for the "Main Document".
- 2) A separate electronic file of each figure (tiff, EPS or print-quality PDF preferred). The format should be produced directly from original creation package, or original software format. PowerPoint files are not accepted.
- 3) Electronic supplementary material: this should be contained in a separate file and where possible, all ESM should be combined into a single file. All supplementary materials accompanying an accepted article will be treated as in their final form. They will be published alongside the paper on the journal website and posted on the online figshare repository. Files on figshare will be made available approximately one week before the accompanying article so that the supplementary material can be attributed a unique DOI.

Online supplementary material will also carry the title and description provided during submission, so please ensure these are accurate and informative. Note that the Royal Society will not edit or typeset supplementary material and it will be hosted as provided. Please ensure that the supplementary material includes the paper details (authors, title, journal name, article DOI). Your article DOI will be 10.1098/rspb.[paper ID in form xxxx.xxxx e.g. 10.1098/rspb.2016.0049].

- 4) A media summary: a short non-technical summary (up to 100 words) of the key findings/importance of your manuscript.

5) Data accessibility section and data citation

It is a condition of publication that data supporting your paper are made available either in the electronic supplementary material or through an appropriate repository.

In order to ensure effective and robust dissemination and appropriate credit to authors the dataset(s) used should be fully cited. To ensure archived data are available to readers, authors should include a 'data accessibility' section immediately after the acknowledgements section. This should list the database and accession number for all data from the article that has been made publicly available, for instance:

- DNA sequences: Genbank accessions F234391-F234402
- Phylogenetic data: TreeBASE accession number S9123
- Final DNA sequence assembly uploaded as online supplemental material
- Climate data and MaxEnt input files: Dryad doi:10.5521/dryad.12311

NB. From April 1 2013, peer reviewed articles based on research funded wholly or partly by RCUK must include, if applicable, a statement on how the underlying research materials – such as data, samples or models – can be accessed. This statement should be included in the data accessibility section.

If you wish to submit your data to Dryad (<http://datadryad.org/>) and have not already done so you can submit your data via this link

[http://datadryad.org/submit?journalID=RSPB&manu=\(Document not available\)](http://datadryad.org/submit?journalID=RSPB&manu=(Document not available)) which will take you to your unique entry in the Dryad repository. If you have already submitted your data to dryad you can make any necessary revisions to your dataset by following the above link. Please see <https://royalsociety.org/journals/ethics-policies/data-sharing-mining/> for more details.

- 6) For more information on our Licence to Publish, Open Access, Cover images and Media summaries, please visit <https://royalsociety.org/journals/authors/author-guidelines/>.

Once again, thank you for submitting your manuscript to Proceedings B and I look forward to receiving your revision. If you have any questions at all, please do not hesitate to get in touch.

Sincerely,
Dr John Hutchinson, Editor
mailto: proceedingsb@royalsociety.org

Associate Editor
Board Member
Comments to Author:

Apologies for the short comment, the current situation is having a huge impact on my time. I'd like to congratulate the authors for their work, the paper has improved significantly. I share my concern with the referee about data availability, while I trust the authors based on their honesty and past track record, we would like to doublecheck everything will be available at the time of publication.

Reviewer(s)' Comments to Author:

Referee: 2

Comments to the Author(s).

This is a very interesting paper bringing new insights on the role of inversions on polymorphic mimicry. I think the authors have correctly addressed my previous comments although I was not able to see the figures and tables in this revised version (I assumed they have not been modified since the previous version ? Except the new column I suggested on table 1 I guess ?).

You may also want to replace 'Being confined to the regulatory region, the inversion may or may not have a functional role in the expression of the novel phenotype' by 'Although the functional role of this inversion, confined to the regulatory region, in the expression of the novel phenotype is currently unknown...' on line 33 of the abstract.

Decision letter (RSPB-2020-0443.R1)

31-Mar-2020

Dear Dr Timmermans

I am pleased to inform you that your manuscript entitled "Mimicry diversification in *Papilio dardanus* via a genomic inversion in the regulatory region of engrailed-inverted" has been accepted for publication in Proceedings B.

You can expect to receive a proof of your article from our Production office in due course, please check your spam filter if you do not receive it. PLEASE NOTE: you will be given the exact page length of your paper which may be different from the estimation from Editorial and you may be asked to reduce your paper if it goes over the 10 page limit.

If you are likely to be away from e-mail contact please let us know. Due to rapid publication and an extremely tight schedule, if comments are not received, we may publish the paper as it stands.

If you have any queries regarding the production of your final article or the publication date please contact procb_proofs@royalsociety.org

Your article has been estimated as being 10 pages long. Our Production Office will be able to confirm the exact length at proof stage.

Open Access

You are invited to opt for Open Access, making your freely available to all as soon as it is ready for publication under a CC BY licence. Our article processing charge for Open Access is £1700.

Corresponding authors from member institutions

(<http://royalsocietypublishing.org/site/librarians/allmembers.xhtml>) receive a 25% discount to these charges. For more information please visit <http://royalsocietypublishing.org/open-access>.

Paper charges

An e-mail request for payment of any related charges will be sent out shortly. The preferred payment method is by credit card; however, other payment options are available.

Electronic supplementary material:

All supplementary materials accompanying an accepted article will be treated as in their final form. They will be published alongside the paper on the journal website and posted on the online figshare repository. Files on figshare will be made available approximately one week before the accompanying article so that the supplementary material can be attributed a unique DOI.

You are allowed to post any version of your manuscript on a personal website, repository or preprint server. However, the work remains under media embargo and you should not discuss it with the press until the date of publication. Please visit <https://royalsociety.org/journals/ethics-policies/media-embargo> for more information.

Thank you for your fine contribution. On behalf of the Editors of the Proceedings B, we look forward to your continued contributions to the Journal.

Sincerely,

Editor, Proceedings B

mailto:proceedingsb@royalsociety.org

Appendix A

We thank the Associate Editor and the reviewers for their effort and useful comments. We have implemented these comment as much as possible. This required substantial re-writing of the Introduction and particularly the Discussion, which is now shorter and less 'speculative', as requested. We have also merged two of the figures (2 and 4). For your convenience we also include a copy of the paper with all changes marked.

Associate Editor

Comments to Author:

I really enjoyed reading this manuscript, it provides solid evidence for the interface between the evolution of regulatory regions and the animal ecology. I'd like to congratulate the authors. I'd like to encourage them to consider great suggestions from the referees. I'd also like to suggest to merge figure 3 and 4 in a single figure.

- **We have merged figures 4 with figure 2. This seemed logical as these two figures use the same scale.**

Reviewer(s)' Comments to Author:

Referee: 1

Comments to the Author(s)

In this study, the authors assembled the whole genome sequence of *Papilio dardanus* and conducted genomic comparison based on resequencing of different forms of female wing pattern. As a result, the authors found that the genomic region flanking the previously recognized mimicry genes had elevated *Fst* and LD (linkage disequilibrium) when genomic sequences were compared between two female wing forms. Importantly, there seemed to be an inversion in the region of elevated *Fst*/LD between two wing forms. The authors argued that the region may involve cis-regulatory region, and allelic differences protected by the inversion may control the occurrence of two female wing pattern phenotypes. In *Papilio* butterflies, genomic study of the mimicry supergene locus involving doublesex gene has revealed both the presence and absence of inversion, suggesting that the mechanism maintaining the allelic differences can be variable. The present study reports the possible role of inversion in maintaining female mimicry polymorphism controlled by different genes (*engrailed-inverted*). This finding suggests that occasional chromosomal inversion in different genomic regions may have caused evolution of different mimicry supergenes in *Papilio* butterflies. Overall I agree that the contents of the manuscript are sound and important in the study of butterfly mimicry evolution.

However, there are some weak points in the genomic analyses. In particular, the Hc allele sequence has not been determined clearly, so the actual sequence differences between Hh and Hc allele have not been depicted. Accordingly, the evidence of inversion appears not to be decisive enough (e.g., lack of Hrr plot showing different orientation between Hc and Hh). Therefore, I recommend that the authors include some cautious discussion about the present resolution of the genomic structure and allelic sequences of the putative supergene locus.

- **We appreciate the comment, and have added a few caveats, e.g. we now specifically mention that we have not de novo assembled the Hc allele sequence (e.g. on line 212-214). However, we have compare the divergence, have the ends in the sequence reads and prove them with PCR.**

Minor points:

P. 7, second paragraph mentions “The South African f. cenea specimen ... homozygous for the inversion (Hc/Hc)”, but in Table 1, its genotype is indicated as Hc/? Please clarify. If that individual was truly Hc/Hc, it could be used to determine Hc sequence.

- **It is not unlikely to be Hc/Hc as there is high frequency of the Hc allele in S Africa, but it has not been confirmed to be Hc/Hc (No offspring was obtained). We have deleted “(Hc/Hc)” from the main text.**
- **We tried to assemble the Hc allele from this specimen, but unfortunately the sequence reads did not assemble into a full contig.**

Please indicate unit of base pairs (bp, kbp, or Mbp?) in figures 2 and 3, and add scale of bp in figure 4.

- **This information has now been added to figure 2 and 3. Figures 2 and 4 have been merged into a single figure.**

Referee 2

Comments to the Author(s)

By providing a draft genome of *P. dardanus* and by resequencing several individuals displaying different wing colour pattern within this species involved in Batesian mimicry, this manuscript describes a novel case of inversion polymorphism controlling variations in mimetic colour patterns. The uncovered inversion is a putative regulatory region of the genes engrailed and invected. These genes were already identified as functional candidates involved in colour pattern diversification within this species, but the implication of this regulatory region was ignored. By comparing allelic divergence in sympatric and allopatric forms, this manuscript also brings relevant evidences for the role of inversions in the evolution of adaptive polymorphisms. Although I think the results presented here are really interesting, the manuscript sometimes lacks of precision and that the discussion is probably too long and speculative. I also have a number of comments and questions listed below that should be carefully addressed.

Specific comments:

It would facilitate the reviewers' job if the lines of the manuscript were numbered.

- **We have now added line numbers. We apologise for not including these in the original submission**

Title

I think one of the major novelties of the paper is that the inversion is located in the regulatory region of the engrailed-invected locus, I would thus recommend changing the title to highlight this original findings. As you mentioned yourself in the introduction, variations in engrailed were already identified as associated with colour pattern variations in *P. dardanus* in a previous study (Thompson et al. 2014).

- **Following the reviewer's suggestion, we have updated the title to: “Mimicry diversification in *Papilio dardanus* via a genomic inversion in the regulatory region of *engrailed-invected*”**

Abstract

'multiple forms', I think it is better to be more precise, if I understood correctly you focused on three forms in this study.

- **This has been replaced with "three forms"**

'We hypothesize that the inversion was instrumental in the origin of the novel phenotype by preventing recombination in the evolutionary process of wing pattern refinement.' I do not think the results brought by this study can clarify the evolutionary steps leading to the evolution of new mimetic alleles (whether the inversion occurred simultaneously or followed the mutations allowing a new phenotype to emerge for instance). The functional links between the inverted regions and the nucleotide variations it contained with the inverted-engrailed locus, and to a larger extent, to the colour pattern variations are not identified.

- **We have removed statements that relate to the "evolutionary steps to the evolution of new mimetic alleles" and clearly mention that currently no functional links between the inversion and colour pattern variation are known (e.g. line 30-31: "the inversion may or may not have a functional role in the expression of the novel phenotype").**

I think it would be more relevant to contrast the evolution of inversions in the monomorphic and polymorphic populations therefore showing that inversions are indeed favored by balancing selection. This is a better argument drawn from the results presented here.

- **We agree with the reviewer and have clarified that inversion polymorphisms can be maintained by balancing selection (e.g. line 241 onwards).**

Introduction

'survival may be increased if species mimic multiple unpalatable phenotypes' I think this sentence provides a very bad argument to the evolution of polymorphic mimicry involving group selection. A species is not mimicking anything ! The evolution of new mimetic forms is favored by the negative frequency-dependent selection assumed in Batesian mimics, because rare forms of Batesian mimics have a greater fitness than frequent ones. This should be clearly stated !

Moreover, it is really important to clarify that this drives the persistence of polymorphism within populations (i.e. multiple forms within a given geographic region) rather than at the species level, because as you mentioned in the manuscript, there are also some monomorphic populations of *P. dardanus* as well, where the selection regime and the genetic architecture of colour pattern alleles might differ.

- **The section has been thoroughly revised, following the reviewer's suggestions.**

Replace 'Only the female sex is mimetic' by 'Females only are mimetic'

- **Changed**

It would be better to avoid using 'selection for' that sounds quite finalistic, you may replace 'negative frequency dependent selection for mimetic protection' by 'negative frequency dependent selection generated by predator behavior.'

- **Changed**

Is the locus H restricted to *P. dardanus* ? If so, you may be more precise by adding : 'In *P. dardanus*, ' before 'Wing colours and patterns are controlled by a single Mendelian locus, H, whose various alleles segregate according to a well-defined hierarchy of dominances'. Remove the 's' after 'dominance in the previous sentence.

- **Changed**

'Segregation analysis in pedigree-broods using AFLP [17] and population genetics [18] have shown that the mimicry switch in *P. dardanus* is genetically linked to the engrailed-inverted locus, a region

that codes for two paralogous homeodomain transcription factors involved in anterior-posterior patterning [19]'. As mentioned earlier, the conclusions of these previously published papers seem rather similar to the title of this manuscript, the title should definitely be revised.

- **Title has been changed – see above.**

'Genetic variation especially in the first exon of engrailed shows strong association to specific mimetic wing phenotypes, and a high proportion of non-synonymous mutations together with high allelic divergence suggest balancing selection on this region [20]', does this study focus on the same forms as those studied here ? This should be clarified.

- **This specific sentence has been removed from the revised manuscript. The required information is given in the discussion, line 276 onwards.**

'selection to avoid maladaptive intermediates' As already explained, it would be better to replace 'selection to avoid' by 'selection against non-mimetic intermediates'

- **Changed**

'A critical aspect of this process is that genetic recombination among functional sites is low, to avoid the formation of intermediates with lower fitness.' Replace 'to avoid' by 'preventing'.

- **Changed**

'a mechanism to enhance linkage of co-adapted mutations'. Replace 'a mechanism to enhance' by 'a mechanism increasing'.

- **Changed**

'inversions may only arise at a late stage in the evolution of mimicry loci'. It would be useful to provide alternative hypothesis regarding the selection regime and the sympatry with different mimicry rings.

- **This sentence has been removed from the revised manuscript. To limit the amount of speculation in the discussion we decided not to provide alternative hypothesis regarding selection regimes.**

'the evolution of mimicry loci under selection for tight linkage' Replace 'for tight linkage' by 'promoting tight linkage'.

- **Changed**

'To understand the genetic architecture underlying polymorphic mimicry in *P. dardanus* we use comparative genomics of multiple female forms'. It would be clearer to state that you focused on 3 forms only.

- **Changed**

'This form occurs throughout the entire range of *P. dardanus* and is recessive to all others.' This is misleading, this form is absent in Madagascar, right ?

- **Changed this to "is widely distributed on the African mainland"**

It would also be important to highlight that the male-like form 'meriones' is only observed in Madagascar. Overall, because the sympatry/allopatry relationships among forms are important to understand the selection regime acting on colour pattern variations and therefore on the underlying genetic architecture, these relationships should be explained very accurately in the introduction.

- **We have explained the relationships in lines 97-107.**

Results:

'A draft genome sequence was constructed that allows for the genome wide analysis of structural rearrangements and genetic variation.' Could you estimate the frequency of inversions throughout the genome using pairwise comparisons between the sequenced individuals? This would give a hint on how much inversions occur frequently or not? How likely is it that such inversion occurs close to the gene engrailed just by neutral evolution?

- **We agree that this is an interesting question, because we may see a phenomenon that is associated with the engrailed locus just by chance if these types of inversion events are frequent throughout the genome. However, we are not interested in the question about genome evolution in general, but only about the region potentially involved in the mimicry switch. The inversion we find (irrespective of any others that might exist) does show the expected features, i.e. divergence of alleles and association with alternative phenotypes. Finding other inversions doesn't prove or disprove anything about this particular one.**

Figure 1: It should be clarified that the male phenotype is the last one on the right of the second row of the Top panel.

- **Added this information onto the figure directly and to the figure caption.**

'Within the region, sequence divergence (estimated as p-distance) between the hippoconides individuals and the reference genome sequence was slightly higher than for the cenea individuals and the reference sequence' This result seems surprising, I guess the reference sequence stem from the reference genome which is based on a *P. dardanus tibullus*, i.e. an homozygote with two hippoconides alleles. How can you explain that the genetic divergence in the region putatively involved in colour pattern variations is higher between two hippoconides alleles rather than between hippoconides and cenea?

- **This was a mistake in the text and we are glad that the reviewer spotted it. In agreement with expectations, it is actually lower between *hippoconides* alleles than between *cenea* and *hippoconides* (as can be seen in the newly added panel in figure 2).**

'The inversion was not found in the non-mimetic 'male-like' *P. d. meriones* from Madagascar, which indicates that the bottom-recessive mimetic *f. hippoconides* has the same arrangement as this male-like form'. The hippoconides allele is thus likely to have evolved from an ancestral *meriones* allele through successive mutations promoted by mimicry. Could you provide an estimation of the level of nucleotide divergence between these two alleles as compared to the hippoconides-cenea pair?

- **We have calculated p-distance between the meriones sample and the reference genome (hippoconides). This is indeed higher than for hippoconides and cenea samples. This information is now included in figure 2 (bottom panel).**

'has the same arrangement as this male-like form, and that this specific arrangement is therefore ancestral.' That makes perfect sense, but did you check that this is indeed the same gene order as in the other Lepidoptera? It is worth to clarify this.

- **We looked at this in Timmermans et al., 2014. We did not find evidence for changes in gene order near engrailed-inverted and Hh alleles compared to other Lepidoptera (*Heliconius* and *Bombyx*).**
- **This sentence is about the orientation of the regulatory region. We currently don't have information on the orientation in related species to which we could align this region.**

'The four f. *cenea* specimens represented two distinct subspecies from Kenya and South Africa.' I guess that the 3 specimens from *P. dardanus polytrophus cenea* are from Kenya, and the specimen from South Africa is a *P. dardanus cenea cenea*. This should be clarified in the text of the manuscript.

- **This had been clarified in the revised version of the manuscript.**

Moreover, since the question of sympatry among morphs is important, it would be worth adding a column in table 1 providing the geographic origin of each sample. It is really important readers without prior knowledge on the geographic distribution of mimetic forms of *P. dardanus* could understand the sympatry/allopatry relationships more easily.

- **We have now added this information to table 1.**

'The sequence data furthermore indicated that the Kenyan specimens carried a second non-inverted allele (Table 1).' This sentence is a bit unclear to me, I guess you mean that the 3(?) sequenced individuals from Kenya displaying the *cenea* pattern (*P. dardanus polytrophus cenea*?) are all heterozygotes with a *cenea* allele and a *hippocoonides* allele?

- **This has now been clarified in the text.**

Discussion

'This supports the supergene hypothesis' I am not sure the results obtained here are bringing evidence for the supergene hypothesis, because the data presented here do not pinpoint the causative mutations, and we do not know whether several independent mutations are required for the switch between *cenea* and *hippocoonides* to happen. Once the inversion has occurred, the lack of recombination promotes the divergence between alleles independently from selection exerted by mimicry, making it highly challenging to discriminate mutations actually involved in the phenotypic divergence. The arguments you provide in the following sections of the discussion are also in favor of the pleiotropic effects of this putatively-regulatory region on engrailed, invected and miR-2768. Such pleiotropic effects are also not demonstrated by the data shown here, preventing to robustly conclude on the supergene hypothesis.

- **We agree with the reviewer and have extensively modified the discussion section, it now for example contains the paragraph:**

"We have not determined the sequence of the *cenea* (H_c) allele and do not know whether several independent mutations are required for the switch between f. *cenea* and f. *hippocoonides* to happen, but the fact that a recombination suppressing inversion exists suggests a genomic architecture consistent with the supergene hypothesis (although due to the linkage of mutations within the inversion, it will not be possible to uncover the functional sites without functional studies). " [lines 212-218]

'Invected also contains an intronic microRNA (miR-2768) conserved in Lepidoptera (Fig. 3; Supplementary Figure 12), which has been shown to downregulate *cubitus interruptus* (*ci*), a gene that regulates patterning of the wing primordia via the hedgehog signaling pathway in nymphalid

butterflies [45].’ I think the discussion is generally too long and speculative, regulatory regions can evolve fast and be co-opted, so that comparisons of their role in different species require specific functional studies that are not the aim of the manuscript presented here.

I would recommend shortening the speculations of the functional effect of the inversions and focus more on the association between inversion polymorphisms and persistence of several mimetic forms within population cause by nFDS. Your results are really pointing out that inversions can be especially promoted by nFDS, because the lack of recombination allows divergent alleles to be maintained. It would also be interesting to highlight that polymorphism mimicry in *P. polytes* and *P. dardanus* does involved inversions, but that these inversions are different. The location of the inversion uncovered here with respect to the inversions involved in polymorphic mimicry in other Lepidoptera should be mentioned, and the repeated independent evolution of these architectures should be discussed in light of the occurrence of inversions throughout the genomes.

- **We have shortened the discussion and removed and/or toned down speculative sections on the functioning of the switch.**
- **A study on the repeated, independent evolution of these inversion architectures is indeed interesting, but in our opinion outside the scope of the current manuscript. Such a study, if thoroughly conducted, would need to be a comparative analysis and include several mimetic species.**

‘obviating the need for recombination-suppressing mechanisms’ This is unclear: the selection regime involved should be properly explained better. I guess the hippoconides allele benefited from a greater advantage over the non-mimetic allele meriones at some point, so that selection promoting mimicry has led to the evolution of this mimetic form, and completely replaced the ancestral allele in the African populations. Could you provide insights on why this happened on the African continent and not in Madagascar? Does the models of the forms hippoconides and cenea are absent in Madagascar and abundant elsewhere?

- **We agree that the section was confusing. The whole paragraph has been carefully rewritten (lines 234-252).**

‘Determination of the phenotype likely works in concert with other changes in the engrailed-inverted region,’ I am not sure I understand this hypothesis given that you did find recombinations in you broods occurring in the flanking region, making it hard to reconcile the concerted evolution of the inversion and the nucleotide variation in engrailed.

- **Our data suggests that recombination is reduced outside the inversion too (spanning a region ~75kb).**

‘Inversions are not necessary, but helpful, and generally have the function to reduce recombination,’ I strongly disagree with this sentence, the inversion do not have a function per se, it can be promoted if it captures an alleles under positive selection because of mimicry and be maintained in the population because it does not lead to recombinant alleles with poor fitness, this should be explained more clearly.

- **Paragraph has been rewritten in accordance with the reviewer’s suggestions (e.g. see lines 292-294)**