THE ROYAL SOCIETY PUBLISHING

PROCEEDINGS B

# Convergent evolution of cytochrome P450s underlies independent origins of keto-carotenoid pigmentation in animals

Nicky Wybouw, Andre H. Kurlovs, Robert Greenhalgh, Astrid Bryon, Olivia Kosterlitz, Yuki Manabe, Masahiro Osakabe, John Vontas, Richard M. Clark and Thomas Van Leeuwen

**Article citation details**
*Proc. R. Soc. B* **286**: 20191039.
http://dx.doi.org/10.1098/rspb.2019.1039

**Review timeline**
Original submission: 8 May 2019
Revised submission: 24 June 2019
Final acceptance: 25 June 2019

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

# Review History

## RSPB-2019-1039.R0 (Original submission)

## Review form: Reviewer 1 (Nicola Nadeau)

**Recommendation**
Accept as is

**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?**
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?**
No

**Is it adequate?**
No

**Do you have any ethical concerns with this paper?**
No

**Comments to the Author**
This is a well crafted manuscript that presents clear and compelling evidence for a novel gene involved in carotenoid metabolism.
I have just a few minor corrections to suggest:
Line 125: "has" should be "have"
Line 142: I'm not familiar with the word "teleochrysalids". It is defined in the text as "virgin females", but I wonder if it should actually be "juvenile females".
Lines 181-182: It is not very clear how many RNA pools (2 or 4?) are sequenced or how many individuals are in each (110 or 55?).
Line 190-191, variant calling with GATK's UnifiedGenotyper: I think this assumes diploidy. In a pooled sample the ploidy is effectively much higher. Was this taken into account in the variant calling? How were allele frequencies for the BSA calculated?

## 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?**
Excellent

**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?**
Yes

**Is it clear?**
Yes

**Is it adequate?**
Yes

**Do you have any ethical concerns with this paper?**
No

**Comments to the Author**
Please see attached file. (Appendix A)

## Decision letter (RSPB-2019-1039.R0)

14-Jun-2019

Dear Mr Wybouw

I am pleased to inform you that your Review manuscript RSPB-2019-1039 entitled "Convergent evolution of cytochrome P450s underlies independent origins of keto-carotenoid pigmentation in animals" has been accepted for publication in Proceedings B.

The referee(s) do not recommend any further changes. Therefore, please proof-read your manuscript carefully and upload your final files for publication. 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 me know immediately.

To upload your 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 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, upload a new version through your Author Centre.

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. Please note that PowerPoint files are not accepted.

3) Electronic supplementary material: this should be contained in a separate file from the main text and the file name should contain the author's name and journal name, e.g authorname_procb_ESM_figures.pdf
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. Please see: https://royalsociety.org/journals/authors/author-guidelines/

4) Data-Sharing and data citation
It is a condition of publication that data supporting your paper are made available. Data should be made available either in the electronic supplementary material or through an appropriate repository. Details of how to access data should be included in your paper. Please see https://royalsociety.org/journals/ethics-policies/data-sharing-mining/ for more details.

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=RSPB-2019-1039 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.

5) 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 final version. If you have any questions at all, please do not hesitate to get in touch.

Sincerely,

Dr Sasha Dall
mailto:proceedingsb@royalsociety.org

Associate Editor
Comments to Author:
Dear Authors,

I have now received two careful and constructive reviews for your manuscript "Convergent evolution of cytochrome P450s underlies independent origins of keto-carotenoid pigmentation in animals". Both the reviewers and myself enjoyed reading this manuscript, and found it could be an interesting publication for Proc B. Although the reviews were overall positive there were a few smaller issues raised by both reviewers that should be addressed and will likely further improve the paper. Additionally, in my opinion it would be nice if the broader general significance of the findings were more plainly and directly stated at the end of the paper as opposed to what is vaguely referenced now. I hope that you find the comments of the reviewers helpful in revising your manuscript.

Reviewer(s)' Comments to Author:

Referee: 1

Comments to the Author(s)
This is a well crafted manuscript that presents clear and compelling evidence for a novel gene involved in carotenoid metabolism.
I have just a few minor corrections to suggest:
Line 125: "has" should be "have"
Line 142: I'm not familiar with the word "teleochrysalids". It is defined in the text as "virgin females", but I wonder if it should actually be "juvenile females".
Lines 181-182: It is not very clear how many RNA pools (2 or 4?) are sequenced or how many individuals are in each (110 or 55?).
Line 190-191, variant calling with GATK's UnifiedGenotyper: I think this assumes diploidy. In a pooled sample the ploidy is effectively much higher. Was this taken into account in the variant calling? How were allele frequencies for the BSA calculated?

Referee: 2

Comments to the Author(s)
Please see attached file

## Author's Response to Decision Letter for (RSPB-2019-1039.R0)

See Appendix B.

# Decision letter (RSPB-2019-1039.R1)

25-Jun-2019

Dear Mr Wybouw

I am pleased to inform you that your manuscript entitled "Convergent evolution of cytochrome P450s underlies independent origins of keto-carotenoid pigmentation in animals" 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 CCBY 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.

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

In this paper the authors study the phenotypic and genetic basis of a red to yellow mutation in a spider mite. They convincingly demonstrate that the mutation affects red ketocarotenoid synthesis and is associated with a frame-shifting deletion in a cytochrome P450 locus. A gene in this family was recently implicated in red ketocarotenoid coloration in birds so this provides a fascinating example of convergent evolution. The conclusions are solid, the paper is well-written and I just have a few suggestions for improvement.

Some further explanation is needed over aspects of the phylogenetic analysis:

- What was the rationale for the particular arthropods chosen?

- The assumption that the 3 variants of CYP384A1 in D. tinctorium represent alleles is questionable - 93.6% similarity seems very low for two alleles, unless this can be justified from other data. In addition, gene conversion among tandem paralogues can lead to high sequence similarity.

Discussion line 360. On the phylogeny, the arthropod clade containing CYP384A1 is present as a single copy in all species sampled, so it is incorrect to claim that this locus is restricted to trombidiforms.

Something about the proposed function of ketocarotenoids in tetranychid mites should be included in the introduction.

Results line 250. β–carotene is described here as plant derived, which is confusing since, as previously stated, these mites can synthesise β–carotene from horizontally transferred genes - presumably some of the β–carotene in the mites could have been derived in this way?

Results line 260. Given that mites in diapause gain darker coloration it is odd that in both wildtype and lemon mites concentrations of carotenoids declines in diapause.

Discussion lines 337 to 351. The suggestion that CYP2J19 hydroxylates as well as ketolates is not supported by the pathways that it is implicated in – in birds, all ketocarotenoids containing hydroxylated terminal rings can be explained by ketolation of hydroxylated dietary precursors. Also, there is no evidence that astaxanthin can be derived from β–carotene in birds.

Minor points

Abstract line 38: “..great number of animal taxa..” appears to be an exaggeration, e.g. within tetrapods and insects, occurrence of a carotenoid ketolase among taxa is patchy.

# Appendix B

Associate Editor
Comments to Author:
Dear Authors,

I have now received two careful and constructive reviews for your manuscript "Convergent evolution of cytochrome P450s underlies independent origins of keto-carotenoid pigmentation in animals". Both the reviewers and myself enjoyed reading this manuscript, and found it could be an interesting publication for Proc B. Although the reviews were overall positive there were a few smaller issues raised by both reviewers that should be addressed and will likely further improve the paper. Additionally, in my opinion it would be nice if the broader general significance of the findings were more plainly and directly stated at the end of the paper as opposed to what is vaguely referenced now. I hope that you find the comments of the reviewers helpful in revising your manuscript.

**REPLY: We thank the editor for the support for our current study. We have addressed the comment on describing the general significance more directly, within the strict manuscript length limitations. The manuscript now reads: "Our findings shed light on the evolutionary history of keto-carotenoid production in trombidiform mites and open up new avenues to understand the potential adaptive value of keto-carotenoid-based traits in these invertebrates."**

Reviewer(s)' Comments to Author:

Referee: 1

Comments to the Author(s)

This is a well crafted manuscript that presents clear and compelling evidence for a novel gene involved in carotenoid metabolism.
I have just a few minor corrections to suggest:
Line 125: "has" should be "have"
**REPLY: We have corrected the sentence.**
Line 142: I'm not familiar with the word "teleochrysalids". It is defined in the text as "virgin females", but I wonder if it should actually be "juvenile females".
**REPLY: We thank the referee for pointing this out to us. We have re-defined "female teleochrysalids" as follows: "nymphal females in their final quiescent stage"**.
Lines 181-182: It is not very clear how many RNA pools (2 or 4?) are sequenced or how many individuals are in each (110 or 55?).
**REPLY: We have outlined the RNA collection more clearly.**
Line 190-191, variant calling with GATK's UnifiedGenotyper: I think this assumes diploidy. In a pooled sample the ploidy is effectively much higher. Was this taken into account in the variant calling? How were allele frequencies for the BSA calculated?
**REPLY: GATK variant calling supports different ploidy levels, but the diploid assumption was used and is appropriate here. Across all samples, variants used in the BSA mapping were all at high allele frequencies. As outlined in the manuscript, we selected bi-allelic variants that were fixed within but different between the two parents, and that were at a high allele frequency in the resulting bulk populations. As the bulk populations were derived from crosses of the parents, they are fundamentally different from pools of many wild individuals where there can be rare alleles, and for which the GATK manual suggest altering the ploidy setting from "diploid". For our study it was a very "simple" prediction task for GATK, as opposed to identifying very rare variants in pooled samples of unrelated individuals. The settings we used have been successful in a number of BSA studies for which causal peaks have now been validated by independent methods (e.g., Van Leeuwen *et al.* 2012, PNAS 109:4407-4412; Demaeght *et al.* 2014, Insect Biochem Mol Biol. 51: 52–61; Bryon *et al.* 2017, PNAS 114:E5871-E5880). Allele frequencies (for parents and bulks) were simply parsed from the VCF file that is the output of the GATK pipeline (the "AD" column in the VCF output file). See citation 22 in the**

**original submission, which is referenced at the appropriate point in the BSA methods section, as follows:**

**"The locus responsible for the lemon phenotype was identified by comparing allele frequencies between the three lemon selected samples to the wild-type sample using previously published BSA genetic mapping methods with statistical testing for genotype-phenotype associations by permutation [22];".**

Referee: 2

Comments to the Author(s)

In this paper the authors study the phenotypic and genetic basis of a red to yellow mutation in a spider mite. They convincingly demonstrate that the mutation affects red ketocarotenoid synthesis and is associated with a frame-shifting deletion in a cytochrome P450 locus. A gene in this family was recently implicated in red ketocarotenoid coloration in birds so this provides a fascinating example of convergent evolution. The conclusions are solid, the paper is well-written and I just have a few suggestions for improvement.
Some further explanation is needed over aspects of the phylogenetic analysis:
- What was the rationale for the particular arthropods chosen?
**REPLY: For our phylogenetic analysis, we selected 11 chelicerate species by leveraging our aims of obtaining the widest range of chelicerate species diversity and only using genome assemblies of appropriate quality and with a reliable annotation. The two outgroup species (*D. pulex* and *H. dujardini*) were selected as previous studies provide a reliable annotation of their CYP2 and CYP3 clans (Baldwin *et al.* 2009, BMC Genomics 10, 169, Nelson 2018, Biochim. Biophys. 1866, 141-154).**

- The assumption that the 3 variants of CYP384A1 in D. tinctorium represent alleles is questionable - 93.6% similarity seems very low for two alleles, unless this can be justified from other data. In addition, gene conversion among tandem paralogues can lead to high sequence similarity.
**REPLY: We can justify our assumption by additional data; the *D. tinctorium* scaffolds that hold the three variants do not code for additional proteins, making it more likely that the variants represent different alleles and not paralogues. We have addressed the reviewer's suggestion and the manuscript now reads: "Three copies were initially identified in the *D. tinctorium* genome assembly, but were finally considered as putative allelic variants (lowest degree of sequence identity was 93.62% and the scaffolds that hold the copies did not code for additional proteins)."**
Discussion line 360. On the phylogeny, the arthropod clade containing CYP384A1 is present as a single copy in all species sampled, so it is incorrect to claim that this locus is restricted to trombidiforms.
**REPLY: On the phylogeny, the clade that contains CYP384A1 is the Trombidiformes order. We refer the reviewer to supplemental figure 5 to clarify this misunderstanding.**

Something about the proposed function of ketocarotenoids in tetranychid mites should be included in the introduction.
**REPLY: Due to the strict length limitation imposed by Proc Roy B, we were restricted to only outline previous work on the carotenoid metabolic pathway and not the proposed biological functions in spider mites in the introduction. However, we do discuss the proposed functional roles in detail in the discussion (line 376 – 406).**

Results line 250. β–carotene is described here as plant derived, which is confusing since, as previously stated, these mites can synthesise β–carotene from horizontally transferred genes - presumably some of the β–carotene in the mites could have been derived in this way?
**REPLY: The reviewer is correct in that the β-carotene levels detected by this TLC set-up can be both plant derived (likely to be mainly present in the mite gut) and de novo synthesized. Unfortunately, α-carotene (solely derived from plants) and β-carotene cannot be separated by**

**our TLC set-up. We followed the reviewer's remark and removed β-carotene from the list of plant derived pigments to avoid confusion.**

Results line 260. Given that mites in diapause gain darker coloration it is odd that in both wildtype and lemon mites concentrations of carotenoids declines in diapause.

**REPLY: The decline of B-carotene when wild-type mites enter diapause has been previously observed (e.g. Kawaguchi *et al.* 2016, Environmental Entomology, 45, 1568–1573). This is in line with the proposed spider mite carotenoid pathway wherein B-carotene is the precursor to the more red-colored keto-carotenoids. The decrease of astaxanthin when *T. kanzawai* mites enter diapause might indeed appear counterintuitive. Two non-mutually exclusive mechanisms might underlie this pattern. First, the brighter coloration in diapausing mites might originate from a relatively higher astaxanthin deposition in tissues closer to the translucent cuticle. Second, and more likely, astaxanthin esterification might be higher in diapausing mites, hereby lowering the level of free astaxanthin.**

Discussion lines 337 to 351. The suggestion that CYP2J19 hydroxylates as well as ketolates is not supported by the pathways that it is implicated in – in birds, all ketocarotenoids containing hydroxylated terminal rings can be explained by ketolation of hydroxylated dietary precursors. Also, there is no evidence that astaxanthin can be derived from β–carotene in birds.

**REPLY: We thank the referee for the comment. We believe that this section of the text has been misread. We point out here that the ketolase reaction itself might consist of two steps, catalyzed by two different enzymes. In this scenario, the cytochrome P450s hydroxylate the C4/C4' positions and other enzymes oxidize the carotenoid intermediate at these positions. The scenario of a two-step ketolase reaction is possible as disrupting / changing the cytochrome P450 locus would still lead to an absence / shift of keto-carotenoids. However, to avoid the misconception that we are solely focusing on pathways that rely on B-carotene, we now refer to the study on *X. dendrorhous*, by stating: "In addition, in the fungus *Xanthophyllomyces dendrorhous*, a single cytochrome P450 is able to produce keto-carotenoids from a carotenoid precursor [39,40].".**

Minor points

Abstract line 38: "..great number of animal taxa.." appears to be an exaggeration, e.g. within tetrapods and insects, occurrence of a carotenoid ketolase among taxa is patchy

**REPLY: The abstract now reads "a great number of animal species".**