nature portfolio

Peer Review File

Multiple mutations in polyketide synthase led to disruption of Psittacofulvin production across diverse parrot species

Corresponding Author: Professor Uri Abdu

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

Version 0:

Reviewer comments:

Reviewer #1

(Remarks to the Author)
In this report, Ghosh Roy et. al., initiate the analysis of a polyketide synthase (PKS) homolog present in diverse members of psittacoformes. This is an interesting subject since animal PKSs are rarely explored, excluding a few recent examples. In this sense, the manuscript published by Cooke T. et. al. 2017 represent a seminal work; where the authors mapped the Mendelian blue locus, which abolishes yellow pigmentation in budgerigar, to a novel polyketide synthase (named MuPKS). The present report aims to extend the studies about such PKSs in other members of Psittaciformes order, carrying the blue phenotype. Although, it does contribute to this area, it needs significant amendment and quality improvements before publication can be considered. Some major points that should be addressed:

1. Perform a rigorous in silico studies and phylogenetic analysis. The content of subtitle: Coding SNPs in conserved residues of PKS completely segregate with the blue phenotype- must be enriched, for example in terms of abundancy, diversity, domain conservation… Lines 135 to 146, please include more precise information. Is the Sanger sequencing genotyping -species/number of specimen- representative?

2. The approach of utilizing WT GgPKS and constructed mutant variants is indirect. It could represent or have been used as control. However, knowing the experimental challenge that represent the heterologous expression of PKSs, the strategy should be better justified and argued. In addition, the construction of GgPKS version carrying simultaneous N428K and T792A substitutions could contribute to clear the confusing statements (in different part of the main text) about segregation of these two variants.

3. Line 209. This subtitle does not include a structural analysis. Is the mapping of relevant residues within the high-resolution structural data of the porcine fatty acid synthase (FAS). Can be move into the discussion section?

Reviewer #2

(Remarks to the Author)
In this manuscript by Roy et al, the authors characterize several missense and nonsense variants at the polyketide synthase (PKS) locus in eclectus parrots, rose-necked parakeets, and galahs. Like budgerigars, these species have been under artificial selection and exhibit a recessive trait called blue caused by loss of psittacofulvin feather pigmentation, which in budgerigars is caused by a missense mutation in the PKS locus. The authors performed targeted sequencing of PKS, and characterized protein-coding mutations that associate with blue by introducing them into the homologous enzyme from chicken (GgPKS), which was previously shown to produce pigments similar to budgerigar PKS when expressed heterologously in yeast. The authors found the missense variants from rose-necked parakeets abolished pigment production by GgPKS in yeast and therefore likely cause the blue phenotype; blue eclectus parrots harbor a nonsense mutation that is likely causative; and the causative mutation from blue galahs was not conclusively found.
This manuscript was fun to read. Artificially-selected recessive traits are fun in general (e.g. Darwin's pigeons etc.), but the blue locus is especially so because the modular nature of PKS enzymes makes it easy to reconstitute psittacofulvin pigment production in yeast and rapidly test many different mutations. There are still many open questions about parrot pigmentation, for instance about what causes variation in red-orange-yellow pigmentation, and whether there is a reductive release mechanism for parrot PKS. But this study is a great follow-up to previous work in budgerigars and lovebirds, and shows how

these sorts of questions might be addressed.
I would be happy to consider a revised version of this manuscript that addresses the following three main points:

Main points:

1) Without seeing more of the mass spectrometry data, it's hard to evaluate the claim that the two major components of the pigment are the same C-16 and C-18 psittacofulvins found in the Cooke et al budgerigar study. And since the authors include no chromatograms from pigments extracted from parrot feathers, their only positive control is wild-type GgPKS. Thus the key conclusions hinge on whether this sample recapitulates the Cooke et al GgPKS results, where it was compared directly to feather pigments. On lines 197-198, it is noted that there were peaks at 243.1385 m/z and 269.1542 m/z, so MS data were evidently collected, but I'd like to see extracted ion chromatograms showing the signal at those two m/z versus elution time in the wild type, negative control, and mutants. These could be shown in Figure 4. On a related note, Figure 4 could be improved by moving panels C-D to the supplement (the absorbance data do not need to be shown in such granular detail in the main figure), and replacing them with an absorbance chromatogram at a single representative wavelength, such as 400 nm, and then stacking the extracted ion chromatograms on top or below.

2) I can't find anything about the SNP frequencies, which makes it hard to evaluate whether any particular variant is causative. I'm assuming each of the 69 specimens listed in Table 1 were genotyped individually. If that's the case, the table ought to be reformatted to show the frequencies, for instance by removing the columns labelled "no. of specimens" and "SNPs found", and replacing them with seven columns: E668*, G112R, K883*, R243K, Q1957*, N428K, T792A, and tallying the number of specimens carrying the particular variant under the appropriate column heading. On a related topic, methodological details are sparse regarding how the variants were called. I'm assuming any site where a heterozygous individual was found among the blue specimens would be excluded from consideration, but to call a site heterozygous one would probably have to manually look at raw chromatograms from Sanger-sequencing the PCR products. The authors should clarify what was done.

3) Artificial selection on a locus like blue will typically sweep the causative nucleotide variant and many other distal variants to high frequency. For the species that exhibit multiple potential causative SNPs, such as P. krameri (E668* and G112R), and P. eupatria (K883* and R243K), do these pairs of alleles always occur on the same haplotype, consistent with a single selective sweep, or do they occur on different haplotypes, consistent with two independent sweeps? It would seem rather unusual and interesting if there were independent sweeps, so even if the data do not allow for haplotype reconstruction, the authors should at least comment on this possibility in the discussion. Again, as mentioned in Main Point #2, Table 1 should be presented in such a way that makes full use of the data that were collected, including haplotype information, if that exists.

Minor points:

Fig 4C-D: Are there chromatograms from T792A? There are UV-A fluorescence data from T792A in panel B, but I don't see it in panels C-D. It could go in the supplement if there is no space.

Fig 5: This figure would be improved by using alphafold to predict the structures of the parrot KS domains instead of showing the corresponding regions of the already-published mammalian PKS structure. It doubt it affects the conclusions much, so it's not a main point for me, but I'm suggesting it because it's pretty easy, and on its way to becoming a routine technique. Don't try running it on the whole protein. Just do the KS domain. Many institutions already have alphafold running on a compute cluster somewhere, but it can also be run on a desktop. I've successfully installed it on Ubuntu Linux myself using the following release: https://github.com/kalininalab/alphafold_non_docker. Once you have the structure predictions as PDB files, you can align the structures using something like Caretta-shape: https://github.com/TurtleTools/caretta. This will give you spatially aligned PDB files that you can open simultaneously in pymol to view as overlaid structures.

Lines 54-55: Perhaps add a reference to make it clearer that many PKS genes had already been found in animal genomes, and some of them characterized, before MuPKS was studied.

Lines 102-116: This section can be shortened because a lot of what it describes is obvious from the photographs of Figure 1.

Lines 241-242: "melanin-based blue structural coloration" should be "structural colors", since melanin is necessary but not sufficient to make structural color.

Lines 285-286: The authors state, "we suggest a combinatorial role of N428K and T792A in altering PKS function leading to the blue phenotype", which seems plausible, but it's probably also worth mentioning that the real causative variant might be regulatory, not coding. Just consider chickens—they have a functional PKS too, but they're not red. Rather it's because PKS is not expressed in their feathers.

Typographical and style points:

Multiple places: "blue P. krameri" should be "P. krameri (blue)" and likewise for other taxonomic names, since the proper nomenclature is genus-species-variety.

Multiple places: There should be no space between citation number and the preceding word.

Fig 1 legend: Taxonomic families such as “psittaculidae” and “cacatuidae” should be italicized.

Fig 1,2 legend: Words following “(A)”, “(B)”, “(C)” etc should be capitalized.

Fig 4 legend: “a-HA antibody” should be “α-HA antibody”.

Fig 4 legend: It appears there is a sentence here that is a comment not intended to be included in the final draft: “Because the labels of wavelength are not aligned in all panels in D, it would be helpful to show a vertical dashed line across panels.”

Lines 46-47: “Although the purpose of this unique phenomenon is not clear, the psittacofulvins have likely played a pivotal role in the evolutionary success of parrots”. Re-word, because if they played a pivotal role, then the purpose is clear.

Line 77: “test check” should be either “test” or “check”.

Lines 90-93: “This is the first evidence of variations in the ketoacyl synthase (KS) domain, that shows severe disruptive effect on PKS function as previously shown within the MAT domain in budgerigars and lovebirds, also suggesting combined effect of multi-domanial variations in phenotype determination.” Awkward syntax.

Lines 104-105: “In the following segment, we have characterized variations in blue phenotypes among diverse parrot species analyzed in this study” should be “We have characterized variations in blue phenotypes among diverse parrot species.”

Line 199: “(Cooke et al. 2017)” should be replaced with a citation number.

Line 241: “in combination to” should be “in combination with”.

Line 244: “comparatively less profuse” should be “smaller”.

Line 245: “Coalescence” should be “Comparison”.

Line 264: “hue, share” should be “hue share”

Line 299: “coloration as evolutionarily constrained” should be “coloration is an evolutionarily constrained”

Line 305: “is similarly disruptive as” should be “is as disruptive as”

Line 318: “cause a cease in” should be “halt”

Line 330: “Also facilitates” should be “It also facilitates”.

Reviewer #3

(Remarks to the Author)
Parrots have the ability to produce a unique pigment - psittacofulvin - to color their feathers. In 2017 a PSK gene was identify to be responsible for the production of this pigment in the budgerigars, by studying the blue phenotype. It's not until earlier this year another study on lovebirds (Agapornis spp.) has identified the same loss of function mutation as the one found in budgies. Our knowledge on the mechanism of psittacofulvin production and evolutionary constraint of PKS is still very limited. Therefore, this is a timely study that the authors have compared the PKS genes of wild type and blue individuals of four parrot species to identify the underlying mutations leading to the loss of psittacofulvin pigmentation. Interestingly, they have identified both nonsense and missense mutations, and most of the missense mutations are located in the KS domain. They have performed functional test to determine whether those missense mutations can lead to the loss of psittacofulvin production. This is a well done project with nice results. I only have minor comments (listed below) that hopefully can help to improve the clarify of the manuscript.

Line 32-33: "complete or reduced loss of psittacofulvin production" --> "reduced loss" is not correct. It should be "complete loss of or reduced psittacofulvin production".

Line 54-55: "Since the initial discovery, PKS enzymes have been found in many animal genomes, except for placental mammals" --> based on Cooke et al. 2017? Please add citation(s).

Page 3: Separate the Introduction into multiple paragraphs instead of a single big paragraph.

Line 77: "Therefore, we decided to test check whether PKS homologs play similar roles among variants from diverse psittacoformes members, carrying the blue phenotype" --> revise the sentence.

Line 79: "four popular pet species from the order Psittaciformes" --> Can cite a reference to support this, e.g. Chan et al. 2021. Global Ecology and Conservation, 30, p.e01784.

Line 89: "Interestingly, this variant segregates with a non-causative variant in the MAT domain, leading to blue phenotype" --> "a non-causative variant leading to blue phenotype"? Revise the sentence to make it clear.

Line 119-120: "We started by examining the status of psittacofulvin pigmentation in a typical parrot feather carrying the blue phenotype" --> the author said this, but then they mentioned the result of the wild type green feather first.

Line 143: Fig. 3B should be 3A instead since it's referred to earlier in the text.

Line 147-157: Two nonsense/missence mutations were identified in 3 of the 4 studied species. Did the two mutations in the same species always occur together or occur in different individuals? The authors need to provide this information.

Line 147-157 and Table 1: Provide information of the genotype frequency for each phenotype would be informative. Table 1's title is "Genotype calling" but there's no genotypes provided. Information should be provided to tell whether all those mutations are fixed in the blue mutants (i.e. homozygous for the mutations).

Table 1: Suggestion: Adding the result of functional test and mutation type to the table would help interpretation.

Table 1: Unclear what's mean by "Auxiliary variant". Or are they really "auxiliary"? (Please also see my comment below.)

Line 229: "This asparagine residue is conserved also among fatty acid synthases of lower organisms" --> replace "lower organisms" with other terms.

Line 268: "(Table 1.)" --> remove the ".".

Line 285-286: "we suggest a combinatorial role of N428K and T792A in altering PKS function leading to the blue phenotype" --> as mentioned above, it was not clear whether these two mutations always occur together in the blue individuals. Information should be given in the Results section.
Also, since both mutations in blue E. roseicapilla doesn't lead to a complete loss of psittacofulvin (T792A didn't even reduce psittacofulvin), the authors should also mention alternative explanations such as other loci is involved in psittacofulvin production, or those mutations were lined to the causal mutation(s) in non-coding regions.

Line 319-320: Since both nonsense and missence mutations occurred in the same species for P. krameri and P. eupatria, both mutations or either one mutation can be the causal mutation(s) of the phenotype. In that case, we have no idea which one was the original cause of the blue phenotype. For example, in E. roratus the nonsense mutation was proposed by the author to be the cause of pigment loss. Are the nonsense and missence mutations always occur together? More information and discussion is needed.

Line 528: "GeneBank" should be "GenBank".

Line 530: "Supplemental material (is) available online."

Figure 3A: Indicate what does the blue colors stand for in the legend.

Figure 1C: There are two polytomy nodes in the tree. The phylogenetic relationship should be resolved in some papers studying parrot phylogeny (e.g. the papers below). Please check and revise.

Smith, B.T., Merwin, J., Provost, K.L., Thom, G., Brumfield, R.T., Ferreira, M., Mauck III, W.M., Moyle, R.G., Wright, T.F. and Joseph, L., 2023. Phylogenomic analysis of the parrots of the world distinguishes artifactual from biological sources of gene tree discordance. Systematic Biology, 72(1), pp.228-241.

Wright, T.F., Schirtzinger, E.E., Matsumoto, T., Eberhard, J.R., Graves, G.R., Sanchez, J.J., Capelli, S., Müller, H., Scharpegge, J., Chambers, G.K. and Fleischer, R.C., 2008. A multilocus molecular phylogeny of the parrots (Psittaciformes): support for a Gondwanan origin during the Cretaceous. Molecular biology and evolution, 25(10), pp.2141-2156.

Figure 1C: "sp." should not be in italic.

Figure 4: In legend it says "Because the labels of wavelength are not aligned in all panels in D, it would be helpful to show a vertical dashed line across panels" --> There's no vertical dashed line in the figure.

Figure 4 legend: "UV-" should be "UV-A"?
Figures should be stand alone. Provide the full term of "WT".

Figure 5: Can also show the position of N428, probably in the supplementary, as the authors proposed that it may also be relevant in line 285-286.

Version 1:

Reviewer comments:

Reviewer #1

(Remarks to the Author)

Reviewer #2

(Remarks to the Author)
Wow! I'm impressed by the level of care the authors took in responding to my points and those of the other reviewers. In particular the yeast overexpression figure has seen a big improvement. The way the genotype data is presented is much better as well. All of my main concerns have been addressed. Congratulations on a very nice study!

Reviewer #3

(Remarks to the Author)
The authors have addressed all my comments properly. I have no further comments.

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

Response to reviewers

Reviewer #1

In this report, Ghosh Roy et. al., initiate the analysis of a polyketide synthase (PKS) homolog present in diverse members of psittacoformes. This is an interesting subject since animal PKSs are rarely explored, excluding a few recent examples. In this sense, the manuscript published by Cooke T. et. al. 2017 represent a seminal work; where the authors mapped the Mendelian blue locus, which abolishes yellow pigmentation in budgerigar, to a novel polyketide synthase (named MuPKS). The present report aims to extend the studies about such PKSs in other members of Psittaciformes order, carrying the blue phenotype. Although, it does contribute to this area, it needs significant amendment and quality improvements before publication can be considered.

We appreciate the reviewer's positive feedback about our work. In this revised version, we incorporate new data to support our conclusions and provide further clarifications below.

1. Perform a rigorous in silico studies and phylogenetic analysis. The content of subtitle: Coding SNPs in conserved residues of PKS completely segregate with the blue phenotype- must be enriched, for example in terms of abundancy, diversity, domain conservation… Lines 135 to 146, please include more precise information. Is the Sanger sequencing genotyping -species/number of specimen- representative?

We agree with the reviewer's comments. To address the criticisms of this part of the result we modified our revised manuscript in several ways. First, as suggested by both reviewer #1 and #3, a comprehensive phylogenetic analysis has been reperformed through Timetree of Life (https://timetree.org/) with the help literature evidence suggested by reviewer #3 (comment on Figure 1C). This analysis has resolved all the polytomy nodes in the tree, providing better assessment of evolutionary relationships among tested species in this study. Moreover, to enrich the result part subtitled "**Coding SNPs in conserved residues of PKS completely segregate with the blue phenotype**", the sequence conservation of the PKS homologs among tested species is analyzed by a multiple sequence alignment of the amino acid sequence of following PKS homologs and added a supplementary table with Percent Identity Matrix created by Clustal2.1 (Table S2)-

GgPKS (PKS homolog of *G. gallus:* LOC420486)
MuPKS (PKS homolog of *M. undulatus*: LOC101880715)
PkPKS (PKS homolog of *P. krameri*: OR452361)
PePKS(PKS homolog of *P. eupatria*: OR452362)
EcrPKS (PKS homolog of *E. roratus*: OR452363)
EorPKS (PKS homolog of *E. roseicapilla*: OR452364)

The multiple sequence alignment showed more than 90% identity match between PKS homologs tested in this study including MuPKS. Additionally, motif prediction by PROSITE webtool (http://prosite.expasy.org/) of the deduced amino acids confirmed the presence of all the functional domain characterized in MuPKS 8, as illustrated in fig. 3.

To answer the query whether **the Sanger sequencing genotyping is representative**, we would like to clarify that we sequenced decent number WT and *blue* specimens per species to ensure representative sampling from domestic parrot population of Israel and confirmed sequence conservation of the affected residues in functional PKS domains, suggesting a conserved mechanism of psittacofulvin biosynthesis in diverse parrots. To simplify the genotyping data, we added specimen numbers in the main text, updated table 1 in the manuscript and provided the exact number of specimens with respective details as supplementary table (Table S2).

2. The approach of utilizing WT GgPKS and constructed mutant variants is indirect. It could represent or have been used as control. However, knowing the experimental challenge that represent the heterologous expression of PKSs, the strategy should be better justified and argued

For functional validation of the candidate variations, we started with expressing PkPKS (OR452361) and PePKS (OR452362) under inducible promoter but organic extracts from yeast expressing these homologs showed no visible yellow pigmentation. Cooke *et al.* already showed that organic extracts from yeast cells expressing chicken PKS (GgPKS; LOC420486) exhibited a yellow pigment, sharing similar pigment components with yeast expressing MuPKS as detected by LC-MS analysis 8. A multiple sequence alignment also showed that GgPKS share around 80% of sequence identity with PKS homologs tested in this study (Table S2). This led us to opting *PKS* sequence from chicken, since yeast strains expressing chicken *PKS* yielded even higher polyketide concentrations compared to yeast strains expressing *PKS* from budgerigar 8. Very recently, in a collaborative project involving red psittacofulvin synthesis, we successfully used the same GgPKS system for confirming of the role of aldehyde dehydrogenase (ALDH3A2) in determining the final product colour of psittacofulvin 17. For better justification, we modified the text accordingly in the result section titled **"Missense variants disrupt feather pigment synthesis in yeast"** as well as the method section titled **"Yeast transformation and GgPKS expression".**

In addition, the construction of GgPKS version carrying simultaneous N428K and T792A substitutions could contribute to clear the confusing statements (in different part of the main text) about segregation of these two variants.

First, we would like to clarify that in all *blue* specimens of *E. roseicapilla*, we found two non-synonymous substitutions simultaneously- N428K and T792A (as shown in table 1 and s1). We agree with the reviewer that our finding demands a functional validation of GgPKS expression carrying simultaneous N428K and T792A substitutions.

To resolve this issue, we cloned a fresh GgPKS construct with N428K and T792A substitutions together. We extracted pigment from the yeast expressing this new construct. We repeated the UHPLC-PDA-HRAM MS analysis for all the samples (empty vector, GgPKS WT and existing mutants) adding a control psittacofulvin pigment, extracted from green feathers of *P. krameri.* The results suggested that the combined effect of N428K and T792A has more severe effect than the N428K and T792A independently (figure 4c).

Biochemical validation of GgPKS carrying simultaneous N428K and T792A substitutions has been added along with the independent N428K and T792A data in the results (line 224-230) and updated in figure 4 (added below). To clarify the confusion regarding N428K and T792A substitutions, we made descriptive changes in the manuscript accordingly.

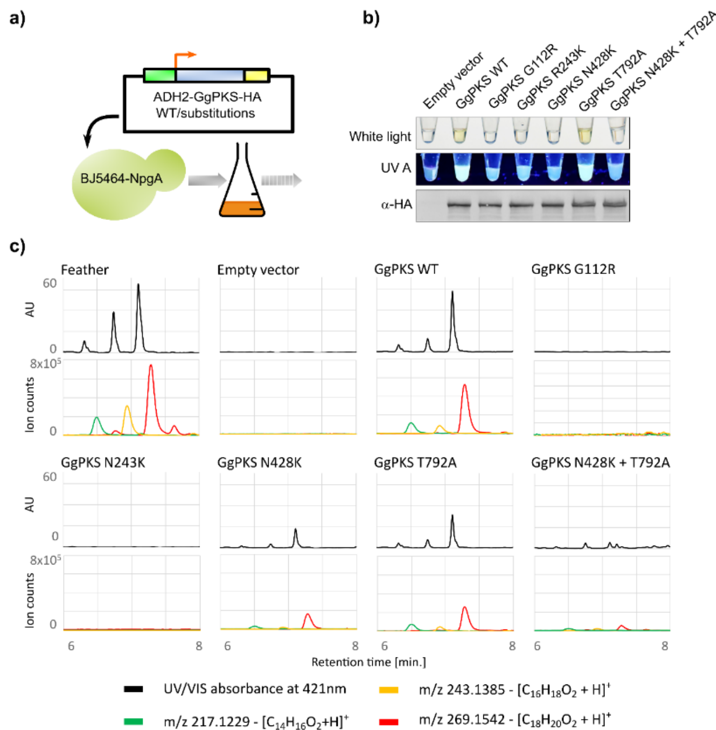

Fig. 4. Biochemical validation of candidate missense variants a) Graphical representation of methanol extraction from yeast strain BJ5464-NpgA expressing HA-tagged GgPKS Wild-type or with the substitutions found in *blue* phenotypes. b) The extracts, illuminated with white light or UV A and corresponding GgPKS expression in total soluble protein extracts from the same yeast cultures by western blot with α-HA antibody. C) the results of the analysis using UHPLC-HRAM QTOF showed that the compounds produced in the feathers of the parrot *P. krameri*, which elute from the Phenyl-Hexyl column at 6.3, 6.7, and 7.1 minutes and have characteristic exact molecular masses of 217.1229 m/z, 243.1385 m/z, and 269.1542 m/z—corresponding to carboxy-psittacofulvins with chain lengths of 14, 16, and 18 carbons, respectively—are also produced in yeast cultures expressing wild-type GgPKS. These compounds are not found in extracts from yeast lacking GgPKS (empty vector) or with GgPKS containing G112R or R243K substitutions**.** A partial production of these compounds is seen in GgPKS with N428K, T792A substitutions and their combinations. Wild-type GgPKS sample shows the highest amount of psittacofulvins, followed by lower peaks in the T792A sample, with the N428K sample yielding the least. When N428K and T792A mutations are combined, the production of psittacofulvins is even lower than when these two mutations are expressed individually.

3. Line 209. This subtitle does not include a structural analysis. Is the mapping of relevant residues within the high-resolution structural data of the porcine fatty acid synthase (FAS). Can be move into the discussion section?

We understand the reviewer's concern thus added the following changes. Instead of showing the corresponding regions of the already-published mammalian PKS structure, we used AlphaFold to predict the structure of the parrot KS, linker, and MAT domains, as suggested by reviewer #2. We mapped onto this predicted model the three mutations in the KS domain and one mutation in the MAT domain, as shown in revised Fig. 5. Thus, we prefer to keep this novel prediction in results section newly named as "**Structural prediction of causal SNPs".**

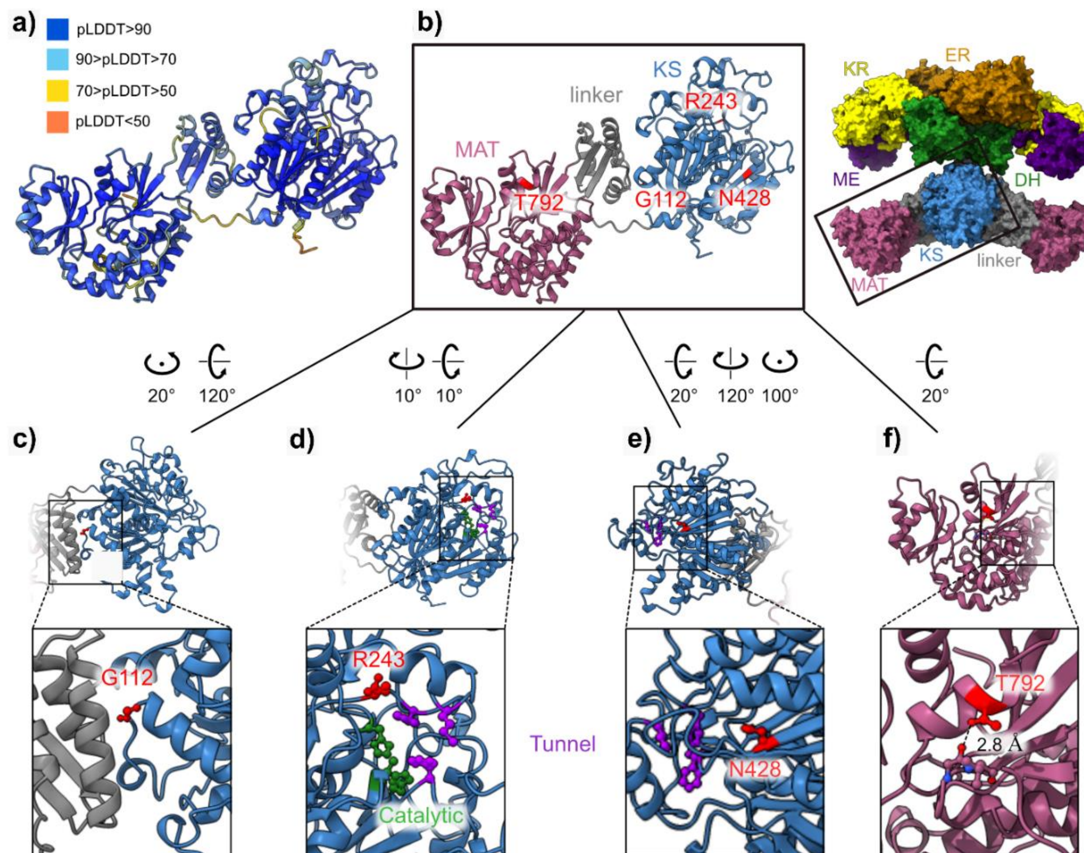

Fig. 5. Structural analysis of causal SNPs. (**A**) Predicted structure of the KS, linker, and MAT domains of the *P. krameri* (*blue*) PKS generated using AlphaFold (Jumper et al., 2021) and coloured according to per-residue confidence (pLDDT) scale (see legend). Note the high confidence in prediction (indicated in blue). (**B**) Left - the prediction shown in (**A**) coloured by domain, showing the positions of the four residues identified as SNPs (indicated in red). Right - the crystal structure of porcine fatty-acid synthase (PDB: 2VZ8) (Maier et al., 2008) with each domain differently coloured is shown for context. Note that the structure is a homodimer containing two polypeptide chains. The boxed region includes the relevant domains for which the *P. krameri* structure was predicted. KS - β-ketoacyl synthase domain, MAT - malonyl/acetyltransferase domain, DH – dehydrase domain, ME - pseudo-methyltransferase domain, KR - β-ketoacyl reductase domain, and ER - enoyl reductase domain. (**C**) Closer view of G112 (in red), showing it is positioned in a loop, at the interface of a linker domain (grey) and the KS domain (blue). (**D**) Closer view of R243 (in red), showing it is positioned near the catalytic residues (coloured green) and the residues that form the entrance to the substrate binding tunnel (coloured purple). (**E**) Closer view of N428, showing its proximity to the residues that form the entrance of the substrate binding tunnel (coloured purple) and that it is buried under an α helix. (**F**) Closer view of T792, showing a predicted hydrogen bond

(indicated with a dashed line) between the sidechain OH group and the main chain of the adjacent loop.

Reviewer #2

In this manuscript by Roy et al, the authors characterize several missense and nonsense variants at the polyketide synthase (PKS) locus in eclectus parrots, rose-necked parakeets, and galahs. Like budgerigars, these species have been under artificial selection and exhibit a recessive trait called blue caused by loss of psittacofulvin feather pigmentation, which in budgerigars is caused by a missense mutation in the PKS locus. The authors performed targeted sequencing of PKS, and characterized protein-coding mutations that associate with blue by introducing them into the homologous enzyme from chicken (GgPKS), which was previously shown to produce pigments similar to budgerigar PKS when expressed heterologously in yeast. The authors found the missense variants from rose-necked parakeets abolished pigment production by GgPKS in yeast and therefore likely cause the blue phenotype; blue eclectus parrots harbor a nonsense mutation that is likely causative; and the causative mutation from blue galahs was not conclusively found.

This manuscript was fun to read. Artificially-selected recessive traits are fun in general (e.g. Darwin's pigeons etc.), but the blue locus is especially so because the modular nature of PKS enzymes makes it easy to reconstitute psittacofulvin pigment production in yeast and rapidly test many different mutations. There are still many open questions about parrot pigmentation, for instance about what causes variation in red-orange-yellow pigmentation, and whether there is a reductive release mechanism for parrot PKS. But this study is a great follow-up to previous work in budgerigars and lovebirds, and shows how these sorts of questions might be addressed.

I would be happy to consider a revised version of this manuscript that addresses the following three main points:

We thank the reviewer for the appreciation of our work and the constructive feedback that helped to substantially improve the manuscript.

Main points:

1. Without seeing more of the mass spectrometry data, it's hard to evaluate the claim that the two major components of the pigment are the same C-16 and C-18 psittacofulvins found in the Cooke et al budgerigar study. And since the authors include no chromatograms from pigments extracted from parrot feathers, their only positive control is wild-type GgPKS. Thus the key conclusions hinge on whether this sample recapitulates the Cooke et al GgPKS results, where it was compared directly to feather pigments.

We understand the reviewer's concern regarding a feather pigment control that evaluates the claim that major components of the pigment are the same C-16 and C-18 psittacofulvins found

in the Cooke *et al*. budgerigar study. To address this issue, we repeated the UHPLC-PDA-HRAM MS analysis adding a control psittacofulvin pigment, extracted from green feathers of *P. krameri.* The results confirmed the presence of the same psittacofulvins as parrot feathers in GgPKS WT and their absence in the empty vector control (Fig. 4C). We have modified the results sections, corresponding figure 4c and method section accordingly. New HRAM-MS data confirmed the identities of the three compounds as C14-carboxyl psittacofulvin (217.1229 m/z), C16-carboxyl psittacofulvin (243.1385 m/z) and C18-carboxyl psittacofulvin (269.1542 m/z) which correspond to compounds found predominantly in green feathers of parrots (Arbore et al. 2024) and extracts of yeast expressing wild-type MuPKS(Cooke et al. 2017). The Raw files of PDA and MS data for feather pigment, pigment extracted from yeast expressing GgPKS WT and mutants are provided as supplementary data.

On lines 197-198, it is noted that there were peaks at 243.1385 m/z and 269.1542 m/z, so MS data were evidently collected, but I'd like to see extracted ion chromatograms showing the signal at those two m/z versus elution time in the wild type, negative control, and mutants .These could be shown in Figure 4. On a related note, Figure 4 could be improved by moving panels C-D to the supplement (the absorbance data do not need to be shown in such granular detail in the main figure), and replacing them with an absorbance chromatogram at a single representative wavelength, such as 400 nm, and then stacking the extracted ion chromatograms on top or below.

Figure 4c is modified as suggested with feather pigment control. Extracted ion chromatograms showing the signal at m/z versus elution time in the feather pigments, negative control, pigments from yeast expressing GgPKS wild type, and mutants are added. As suggested, absorbance chromatograms are shown at a single representative wavelength of 421nm, stacking above the extracted ion chromatograms.

2. I can't find anything about the SNP frequencies, which makes it hard to evaluate whether any particular variant is causative. I'm assuming each of the 69 specimens listed in Table 1 were genotyped individually. If that's the case, the table ought to be reformatted to show the frequencies, for instance by removing the columns labelled "no. of specimens" and "SNPs found", and replacing them with seven columns: E668*, G112R, K883*, R243K, Q1957*, N428K, T792A, and tallying the number of specimens carrying the particular variant under the appropriate column heading.

We understand the reviewer's concern about the SNP frequencies. Thus, Table 1 is updated showing the number of individuals with particular SNPs, which represent the SNP frequencies in each species. Also, SNP frequencies are mentioned in the text (result section: line 149-159) Additionally, phenotypic and genotypic details of each specimen from four different parrot species are listed in Table S5-S8.

On a related topic, methodological details are sparse regarding how the variants were called.

PKS coding sequence (6 exons) of each of the 69 specimens were amplified from genomic DNA using intronic primers (refer to Table S3 and S4.). Primers were designed within intronic regions to produce amplicons approximately 500–700 bp in size. Amplified products were then

sequenced using the forward primers listed in the Table S3 and S4) via Sanger sequencing. Variants were called by manually examining the raw Sanger sequencing chromatograms in Chromas software version 2.6.6. Details of Sanger sequencing protocol is updated in "Genotyping" section of methods (line: 421-433).

I'm assuming any site where a heterozygous individual was found among the blue specimens would be excluded from consideration, but to call a site heterozygous one would probably have to manually look at raw chromatograms from Sanger-sequencing the PCR products. The authors should clarify what was done.

To clarify this point, we would like to mention that all our blue specimens with the causal SNPs were homozygous. (mentioned in Table 1 legend, line 656 and result section, line 149-167)

3. Artificial selection on a locus like blue will typically sweep the causative nucleotide variant and many other distal variants to high frequency. For the species that exhibit multiple potential causative SNPs, such as P. krameri (E668* and G112R), and P. eupatria (K883* and R243K), do these pairs of alleles always occur on the same haplotype, consistent with a single selective sweep, or do they occur on different haplotypes, consistent with two independent sweeps? It would seem rather unusual and interesting if there were independent sweeps, so even if the data do not allow for haplotype reconstruction, the authors should at least comment on this possibility in the discussion. Again, as mentioned in Main Point #2, Table 1 should be presented in such a way that makes full use of the data that were collected, including haplotype information, if that exists.

We thank the reviewer for highlighting this point. We agree that species with multiple potential causative SNPs require haplotype reconstruction. These alleles may either occur on the same haplotype, indicating a single selective sweep, or on different haplotypes, suggesting two independent sweeps. However, since our genotyping relies on Sanger sequencing rather than next-generation sequencing, we cannot draw conclusions about selective sweeps. Nevertheless, as suggested, we have included a discussion on the potential for haplotype reconstruction in the manuscript "Discussion" section (line 346-354). The text added as follows-

"Understanding this selection dynamic is also critical. Two of our sample species, *P. krameri* (E668* and G112R) and *P. eupatria* (K883* and R243K), exhibit multiple potential causative SNPs, each independently capable of producing the *blue* phenotype. When multiple causative SNPs appear in a species, haplotype reconstruction helps determine if they exist on the same haplotype, indicating a single selective sweep, or on different haplotypes, suggesting independent sweeps. However, our use of Sanger sequencing limits our ability to resolve haplotypes with the precision that next-generation sequencing could provide. Future studies with whole-genome sequencing could enable accurate haplotype reconstruction, clarifying whether the SNPs represent one or multiple adaptive events."

Minor points:

Fig 4C-D: Are there chromatograms from T792A? There are UV-A fluorescence data from T792A in panel B, but I don't see it in panels C-D. It could go in the supplement if there is no space.

Ion chromatograms and the absorbance data for mutation T792A are added into figure 4 along with the combination with N428K.

Fig 5: This figure would be improved by using alphafold to predict the structures of the parrot KS domains instead of showing the corresponding regions of the already-published mammalian PKS structure. It doubt it affects the conclusions much, so it's not a main point for me, but I'm suggesting it because it's pretty easy, and on its way to becoming a routine technique. Don't try running it on the whole protein. Just do the KS domain. Many institutions already have alphafold running on a compute cluster somewhere, but it can also be run on a desktop. I've successfully installed it on Ubuntu Linux myself using the following release: https://github.com/kalininalab/alphafold_non_docker. Once you have the structure predictions as PDB files, you can align the structures using something like Caretta-shape: https://github.com/TurtleTools/caretta. This will give you spatially aligned PDB files that you can open simultaneously in pymol to view as overlaid structures.

We used AlphaFold to predict the structure of the parrot KS, linker, and MAT domains, as suggested by the reviewer. We mapped onto this predicted model the three mutations in the KS domain and one mutation in the MAT domain, as shown in revised Fig. 5. Indeed, we came to the same conclusions regarding the effects of mutations on enzyme folding/activity with the porcine structure and the parrot prediction.

Lines 54-55: Perhaps add a reference to make it clearer that many PKS genes had already been found in animal genomes, and some of them characterized before MuPKS was studied.

Two following references are added in this section (Line 56 and 57) to clarify that many PKS genes had already been found in animal genomes, and some of them characterized before MuPKS was studied-

9. Torres, J. P. & Schmidt, E. W. The biosynthetic diversity of the animal world. Journal of Biological Chemistry vol. 294 Preprint at https://doi.org/10.1074/jbc.REV119.006130.(2019)

10. Li, F. et al. Sea Urchin Polyketide Synthase SpPks1 Produces the Naphthalene Precursor to Echinoderm Pigments. J Am Chem Soc 144, (2022).

Lines 102-116: This section can be shortened because a lot of what it describes is obvious from the photographs of Figure 1.

We agree that photographs (Figure 1a and b) are self-explanatory but as the *blue* phenotypes vary among diverse parrot species analyzed in this study (such as *blue* Galah is not "blue" in

color), we prefer to keep the description. However, most of the details are presented in table S1.

Lines 241-242: "melanin-based blue structural coloration" should be "structural colors", since melanin is necessary but not sufficient to make structural color.

To make the statement clearer, we changed the sentence (line 268) to- "Parrots create green feathers by combining yellow psittacofulvins around the barb cortex with blue structural colors produced by melanin pigments and feather nanostructure."

Lines 285-286: The authors state, "we suggest a combinatorial role of N428K and T792A in altering PKS function leading to the blue phenotype", which seems plausible, but it's probably also worth mentioning that the real causative variant might be regulatory, not coding. Just consider chickens—they have a functional PKS too, but they're not red. Rather it's because PKS is not expressed in their feathers.

We appreciate the positive feedback from the reviewer regarding the interpretation of the role of two variants- N428K and T792A. This part of the discussion is modified though as a result of a new experiment performed with feather pigment control and GgPKS carrying N428K and T792A together. However, the interpretation is more or less consistent. We agree that alteration in psittacofulvin pigmentation could also be result of some regulatory variant, directly affecting PKS expression or some other factor in the pathway defining the final product release. However, our new data suggested that both N428K and T792A are capable of partial disruption of PKS activity independently, the severity of their disruptive nature probably depends on their segregation, as combination of both the variants suggested a lower psittacofulvin production than the independent expression. As we suggested in our text, validating this finding requires an advanced quantitative experiment, which needs to be developed.

Typographical and style points:

Multiple places: "blue P. krameri" should be "P. krameri (blue)" and likewise for other taxonomic names, since the proper nomenclature is genus-species-variety.

Corrected

Multiple places: There should be no space between citation number and the preceding word.

Corrected

Fig 1 legend: Taxonomic families such as "psittaculidae" and "cacatuidae" should be italicized.

Corrected

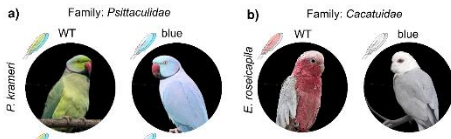

Fig 1,2 legend: Words following "(A)", "(B)", "(C)" etc should be capitalized.

Figure format of the Communications biology suggested mentioning figure segments as small letter (such as, a, b, c). All the figures and legends are modified accordingly.

Fig 4 legend: "a-HA antibody" should be "α-HA antibody."

Corrected

Fig 4 legend: It appears there is a sentence here that is a comment not intended to be included in the final draft: "Because the labels of wavelength are not aligned in all panels in D, it would be helpful to show a vertical dashed line across panels".

Removed

Lines 46-47: "Although the purpose of this unique phenomenon is not clear, the psittacofulvins have likely played a pivotal role in the evolutionary success of parrots". Re-word, because if they played a pivotal role, then the purpose is clear.

Corrected

Line 77: "test check" should be either "test" or "check".

Corrected

Lines 90-93: "This is the first evidence of variations in the ketoacyl synthase (KS) domain, that shows severe disruptive effect on PKS function as previously shown within the MAT domain in budgerigars and lovebirds, also suggesting combined effect of multi-domanial variations in phenotype determination." Awkward syntax.

Rephrased

Lines 104-105: "In the following segment, we have characterized variations in blue phenotypes among diverse parrot species analyzed in this study" should be "We have characterized variations in blue phenotypes among diverse parrot species".

Corrected

Line 199: "(Cooke et al. 2017)" should be replaced with a citation number.

Corrected

Line 241: “in combination to” should be “in combination with.”

Corrected

Line 244: “comparatively less profuse” should be “smaller.”

Corrected

Line 245: “Coalescence” should be “Comparison.”

Corrected

Line 264: “hue, share” should be “hue share”

Corrected

Line 299: “coloration as evolutionarily constrained” should be “coloration is an evolutionarily constrained”

Corrected

Line 305: “is similarly disruptive as” should be “is as disruptive as”

Corrected

Line 318: “cause a cease in” should be “halt”

Corrected

Line 330: “Also facilitates” should be “It also facilitates.”

Corrected

Reviewer #3

Parrots have the ability to produce a unique pigment - psittacofulvin - to color their feathers. In 2017 a PSK gene was identify to be responsible for the production of this pigment in the budgerigars, by studying the blue phenotype. It's not until earlier this year another study on lovebirds (Agapornis spp.) has identified the same loss of function mutation as the one found in budgies. Our knowledge on the mechanism of psittacofulvin production and evolutionary constraint of PKS is still very limited. Therefore, this is a timely study that the authors have compared the PKS genes of wild type and blue individuals of four parrot species to identify the underlying mutations leading to the loss of psittacofulvin pigmentation. Interestingly, they have identified both nonsense and missense mutations, and most of the missense mutations are located in the KS domain. They have performed functional test to determine whether those missense mutations can lead to the loss of psittacofulvin production. This is a well done project with nice results. I only have minor comments (listed below) that hopefully can help to improve the clarify of the manuscript.

We thank the reviewer for the positive feedback. We also appreciate the thought-provoking comments below, to which we try to respond to the best of our ability.

Line 32-33: "complete or reduced loss of psittacofulvin production" --> "reduced loss" is not correct. It should be "complete loss of or reduced psittacofulvin production".

Corrected

Line 54-55: "Since the initial discovery, PKS enzymes have been found in many animal genomes, except for placental mammals" --> based on Cooke et al. 2017? Please add citation(s).

Corrected

Page 3: Separate the Introduction into multiple paragraphs instead of a single big paragraph.

Corrected (split into 3 paragraphs).

Line 77: "Therefore, we decided to test check whether PKS homologs play similar roles among variants from diverse psittacoformes members, carrying the blue phenotype" --> revise the sentence.

Rephrased as "We therefore decided to investigate whether PKS homologs have similar roles in blue variants across different members of Psittaciformes" (line 80)

Line 79: "four popular pet species from the order Psittaciformes" --> Can cite a reference to support this, e.g. Chan et al. 2021. Global Ecology and Conservation, 30, p.e01784.

Suggested reference added

Line 89: "Interestingly, this variant segregates with a non-causative variant in the MAT domain, leading to blue phenotype" --> "a non-causative variant leading to blue phenotype"? Revise the sentence to make it clear.

This part of the introduction is revised (line 87-92). As discussed in comment 2 of reviewer #1, the interpretation of T792A variant is modified. Our new results suggest that while independent expression of either variant (N428K or T792A) can partially disrupt psittacofulvin production, the combined effect of N428K and T792A leads to a more severe disruption than either variant alone (figure 4c). Therefore, we propose that N428K and T792A are capable of partial reduction of psittacofulvin production, while a combinatorial effect of these two variants determines the final *blue* phenotype in *E. roseicapilla*.

Line 119-120: "We started by examining the status of psittacofulvin pigmentation in a typical parrot feather carrying the blue phenotype" --> the author said this, but then they mentioned the result of the wild type green feather first.

Sentence changed to "We started by examining the status of psittacofulvin pigmentation in a typical parrot feather" as mentioned the result of the wild-type green feather first.

Line 143: Fig. 3B should be 3A instead since it's referred to earlier in the text.

We thank the reviewer for this suggestion. Position of affected residues on the domain structure is named as 3a and conservation status is referred as 3b.

Line 147-157: Two nonsense/missence mutations were identified in 3 of the 4 studied species. Did the two mutations in the same species always occur together or occur in different individuals? The authors need to provide this information.

This is clarified in the table, the result text and in the table S5-S8. Two allelic variations found in the *blue P. krameri* and *P. eupatria* was never found in same individuals as trans-heterozygous. They were always segregated independently in homozygous manner. However, two variants found in *blue E. roseicapilla* (N428K and T792A) were always segregated together in homozygous manner.

Line 147-157 and Table 1: Provide information of the genotype frequency for each phenotype would be informative. Table 1's title is "Genotype calling" but there's no genotypes provided. Information should be provided to tell whether all those mutations are fixed in the blue mutants (i.e. homozygous for the mutations).

All the causal mutations fixed in *blue* specimens. are found as homozygous. To clarify this, we revised the following points-

- result text is modified with SNP frequencies.
- Table1 caption is changed to "Homozygous variants found by Sanger sequencing for PKS across species".

- The column heading of table 1 is also changed *blue* SNPs found in PKS.

Table 1: Suggestion: Adding the result of functional test and mutation type to the table would help interpretation.

We respectfully differ from the reviewer's point of view. Table 1 is referred at the result section subtitled, "Coding SNPs in conserved residues of PKS completely segregate with the *blue* phenotype", which is important to summarize the SNP list. The result of the functional test is coming later in the text. We feel that it is difficult to merge them together. However, we want to draw the reviewer's attention to our modified figure 4c, where ion chromatograms and absorbance data is depicted for each mutant, in much simplified manner that should be easier to interpret the result.

Table 1: Unclear what's mean by "Auxiliary variant". Or are they really "auxiliary"? (Please also see my comment below.)

This phrase is removed from the Table1 legend as per the new interpretation (discussed in comment 2 of reviewer #1 and your comment on line 89 and Line 147-157)

Line 229: "This asparagine residue is conserved also among fatty acid synthases of lower organisms" --> replace "lower organisms" with other terms.

Replaced by "across taxa".

Line 268: "(Table 1.)" --> remove the."."

Corrected

Line 285-286: "we suggest a combinatorial role of N428K and T792A in altering PKS function leading to the blue phenotype" --> as mentioned above, it was not clear whether these two mutations always occur together in the blue individuals. Information should be given in the Results section.

Two variants found in *blue E. roseicapilla* (N428K and T792A) were always segregated together in homozygous manner. (discussed in comment 2 of reviewer #1 and your comment on line 89 and Line 147-157)

Also, since both mutations in blue E. roseicapilla doesn't lead to a complete loss of psittacofulvin (T792A didn't even reduce psittacofulvin), the authors should also mention alternative explanations such as other loci is involved in psittacofulvin production, or those mutations were lined to the causal mutation(s) in non-coding regions.

Our new HPLC result showed that while independent expression of variant T792A can partially disrupt psittacofulvin production, the combined effect of N428K and T792A leads to a more severe disruption than either variant alone (figure 4c). discussed in comment 2 of reviewer #1 and your comment on line 89 and line 147-157. Thus, we eliminate the need for alternative explanation here. However, we agree that the effect of multiple loci or causal mutation(s) in

non-coding regions in psittacofulvin is highly possible and recently shown in dusky lories (Arbore et al. 2024).

Line 319-320: Since both nonsense and missense mutations occurred in the same species for P. krameri and P. eupatria, both mutations or either one mutation can be the causal mutation(s) of the phenotype. In that case, we have no idea which one was the original cause of the blue phenotype. For example, in E. roratus the nonsense mutation was proposed by the author to be the cause of pigment loss. Are the nonsense and missense mutations always occur together? More information and discussion is needed.

As discussed before, two allelic variations (one nonsense and one missense) found in the *blue P. krameri* and *P. eupatria* was never found in same individuals as trans-heterozygous. They were always segregated independently in homozygous manner. The missense substitutions found in *blue P. krameri* and *P. eupatria* are highly conserved across taxa and capable of functional disruption as per domain structure analysis. Moreover, upon heterologous expression, they showed a complete loss of psittacofulvin in the yeast extracts. Thus, we inferred that these missense mutations are as effective as the nonsense variants in causing the *blue* phenotype in the respective species. However, the only variant found in *blue* specimens of *E. roratus* is the nonsense mutation and inferred as the causal mutation. Discussion part is revised with the clarification (line 293-297 and 309)

Line 528: "GeneBank" should be "GenBank."

Corrected

Line 530: "Supplemental material (is) available online".

Corrected

Figure 3A: Indicate what does the blue colors stand for in the legend.

The shades of blue in the alignment represent the percentage identity of the amino acid residues. Mentioned in the legend.

Figure 1C: There are two polytomy nodes in the tree. The phylogenetic relationship should be resolved in some papers studying parrot phylogeny (e.g. the papers below). Please check and revise.

A new phylogenetic analysis has been performed through Timetree of Life (https://timetree.org/) with the help of articles suggested by the reviewer (Smith et al. 2023; Wright et al. 2008). It has resolved all the polytomy nodes in the tree, providing better assessment of evolutionary relationships among tested species in this study

Figure 1C: "sp." should not be in italic.

Corrected

Figure 4: In legend it says "Because the labels of wavelength are not aligned in all

panels in D, it would be helpful to show a vertical dashed line across panels" --> There's no vertical dashed line in the figure.

Figure 4 legend is completely modified. Thus, the line has been removed.

Figure 4 legend: "UV-" should be "UV-A"? Figures should be stand alone. Provide the full term of "WT".

Figure 4 legend is completely modified. UV is corrected as "UV-A”.

Essential context and explanations are added in the legend, which visualizes data independently.

WT is mentioned as Wild-type.

Figure 5: Can also show the position of N428, probably in the supplementary, as the authors proposed that it may also be relevant in line 285-286.

N428 was labeled in the original figure as N399, following the porcine enzyme numbering. In revised Fig. 5, we used a prediction for the parrot enzyme and labeled residues according to the parrot enzyme numbering to avoid confusion.