Dr. Frédéric Delsuc
Associate Editor
PCI Evolutionary Biology

Dear Dr. Delsuc,

We appreciate the comments from all reviewers and yours. We feel that after this round of suggestions, our manuscript has been improved. We hope you agree that we have adequately addressed all the issues raised by the reviewers and find that our manuscript is now suitable for recommendation in PCI Evolutionary Biology.

Dear Drs. Juan Opazo and Gonzalo Mardones,

I now have received three reviews from colleagues that have carefuly read your preprint entitled "How many sirtuin genes are out there? Evolution of sirtuin genes in vertebrates with a description of a new family member" submitted for evaluation by Peer Community in Evolutionary Biology.

As you will see in their detailed review reports, all three reviewers found the manuscript interesting and clearly written. I agree with them that your study provides a comprehensive study of this gene family based on detailed evolutionary analyses and I share their appreciation of the efforts you deployed to shed light on the potential function of the newly discovered sirtuin SIRT3-like paralog through functional experiments. All three reviewers also provide valuable suggestions that I think will improve the clarity of the study.

The following points are particularly worth addressing in your revision:

1. More details are needed regarding the construction of the alignment. In particular, the length of the alignment and whether or not it has been processed with a site filtering method before phylogenetic analysis should be indicated. If site filtering has not been performed, as suggested by one of the reviewer, the effect of applying a method such as HMMCleaner on phylogenetic inference might be worth exploring.

We did not clean the alignment before phylogenetic inference. We tried to use HmmCleaner but had several problems installing and executing it. We sent emails to the paper's first and last authors but never received an answer. We tried other options like BMGE, GBlocks, Noisy, and Trimal; however, in all cases, the cleaning procedure resulted in a very short alignment with little phylogenetic information. After that, we think that keeping the phylogenetic tree based on the original alignment seems to be our best option.

2. Two reviewers required additional explanation/clarification regarding the phylogenetic position of the lamprey sequence to the SIRT3 clade and the potential reasons why this sequence is particularly difficult to place in the tree. I agree with them that the potential effect of the compositional bias of the lamprey sequence should be further explored and discussed.

This problem does not exist now, given that the new phylogenetic analyses, including sirtuin sequences from lungfish (as requested by the first reviewer), provide strong support for the monophyly of SIRT3 gene lineage (0.92/95), including the lamprey sequence. This passage was removed from the manuscript

3. Finally, two reviewers had questions concerning the name of the newly identified SIRT3-like paralog and suggest renaming the gene independently of its similarity to SIRT3 by perhaps simply calling it SIRT8. It would be good to hear your informed opinion on this nomenclatural issue.

We agree with this comment. Therefore, we renamed the gene as SIRT8.

I am looking forward to read the revised version of the manuscript.

Best regards,

Frédéric Delsuc on the behalf of PCI Evol Biol

Reviews

Reviewed by Filipe Castro

Opazo and collaborators provide in this manuscript an exhaustive and comprehensive evolutionary analysis of the SIRT gene family in vertebrates species. The study is indeed elegant, clarifying the origin of SIRT3-like genes, and highlighting the importance of taxonomic sampling to infer evolutionary patterns (incidentally, is the gene present in lungfishes?).

We appreciate the positive comments. Regarding the presence of sirtuin genes in lungfish, we searched and found them. Accordingly, we performed all phylogenetic analyses again, following the same protocol described in the manuscript.

The authors further analyze and explore the function of this novel gene utilizing an ample set of approaches – e.g. gene expression, in vivo analysis, modeling and protein function. I have no real comments and congratulate the authors for their work! I understand the authors prevent from speculation but is there a reasonable scenario for both the specific retention and loss of SIRT3-like in these disparate lineages?

Thanks again for the positive comments. It is difficult to anticipate a solid explanation for the loss of this gene in the last common ancestor of mammals, birds, and reptiles mainly because of the little knowledge we have on this newly discovered gene lineage.

A minor doubt relates with nomenclature: I'm not an expert but I wonder whether it would be interesting/possible/advisable (?) to rename the gene independently of its similarity to SIRT3?

It is an interesting question. Initially, we named this gene based on how it appears in the database (SIRT3-like). However, it could be possible to call this newly discovered gene using the following available number; in this case, it will be SIRT8

Reviewed by Nicolas Leurs
I declare that I do not have any conflict of interest with the authors or the content of the article
Evaluation of the various components of the article
Title/abstract/introduction
• Does the title clearly reflect the content of the article? Yes

Thanks

• Does the abstract present the supported findings of the study concerned and no other? Yes

Thanks

• Does the introduction clearly explain the motivation for the study? Yes

Thanks

• Is the research question/hypothesis/prediction clearly presented? Yes

Thanks

• Does the introduction build on relevant recent and past research performed in the field? Yes

Thanks

Materials and Methods
• Are the methods and analysis described in sufficient detail to allow replication by other researchers? Yes although genome assembly version are not stated mainly for synteny analyses

We appreciate this comment; this problem was solved as now we included the genome assembly version in the legend of figure 5.

• Is the experimental plan consistent with the questions? Yes

Thanks

• Are the statistical analyses appropriate? Yes

Thanks

• Have you evaluated the statistical scripts and program codes? NA

We did not implement homemade scripts. All command lines used to run phylogenetic analyses are available in Zenodo.

Results
• Have you checked the raw data and their associated description? Yes

Thanks

• Have you run the data transformations and statistical analyses and checked that you get the same results? No

We did not perform data transformations

• To the best of your ability, can you detect any obvious manipulation of data (e.g. removal)? No

Thanks

• Do the statistical results strongly support the conclusion (p< 10-3 or BF>20)? Yes

Thanks

• In the case of negative results, was a statistical power analysis (or an appropriate Bayesian analysis) performed? NA

It does not apply

• Did the authors conduct many experiments but retain only some of the results? No

Thanks

Discussion
• Do the interpretations of the analysis go too far? No

thanks

• Are the conclusions adequately supported by the results? Yes

Thanks

• Does the discussion take into account relevant recent and past research performed in the field? Yes

Thanks

• Did the authors test many hypotheses but consider only a few in the discussion? No

Thanks

References
• Are all the references appropriate? Yes

Thanks

• Are the necessary references present? Yes

Thanks

• Do the references seem accurate? Yes

Thanks

Tables and figures
• Are the tables and figures clear and comprehensive? Not all of them, font is too small in Fig5a

Thanks for this comment. Accordingly, we have increased the font size.

• Are all the tables/figures useful? Fig5b can be omitted or moved to Supplementary materials

We request to be maintained as this panel provides evidence for the lack of the SIRT8 gene in representative species of amniotes. We think it is a necessary piece of evidence to justify the no presence in public databases.

• Are there too many/too few tables and figures? No

Thanks

• Do the tables and figures have suitable captions such that they can be understood

without having to read the main text? Yes

Thanks

Merits and Strengths
The preprint provides the first detailed and well resolved phylogeny of Sirtuin genes, with an extensive vertebrate species' sampling including representatives of major lineages. An effort was made to prevent biased sampling towards bony fishes or mammals or lineages with additional whole genome duplications. Phylogenetic results allowed the recovery of orthology relationships. Through this extensive and unbiased sampling, the authors show the presence of a previously unidentified Sirtuin gene, an ancestral duplicate of SIRT3, called here SIRT3-like.
The phyletic distribution of this gene is shown to be in every sampled non-amniote gnathostome species. This distribution showcases how sampling bias is the reason the new gene was missed in previous studies. Subsequent analyses were made to localize tissue specific transcription of Sirtuin genes in vertebrates (including SIRT3-like). The elephant Shark SIRT3-like was used for comparative analyses with other vertebrates, through protein structure modeling. The protein was characterized by immunolocalization in transfected cells, enzymatic and functional assays. These results will serve as a good base for further studies on Sirtuin genes and functional evolution of vertebrate paralogs.
The preprint is a very complete study on a newly discovered gene, including analyses ranging from the evolutionary analysis of the duplicative history of the gene family to several demanding functional experiments.

Thanks for all the positive comments.

Flaws, weaknesses and suggestions for improvement
Shouldn't we call this gene SIRT8? As just one additional paralog?

Thanks for this suggestion. Initially, we called this gene based on how it appears in the database (SIRT3-like). In the new version of the manuscript we called SIRT8.

Several instances of unclear evolutionary range/level of comparison in your sentences:
l.79: "ancient group"

We agree with the reviewer, to solve this problem we removed this expression. Now it reads as follows "....is a group of genes that, in mammals, is composed of seven paralogs (SIRT1-7) grouped into four classes (Fig. 1)"

l.81: "variety of biological functions"

This general statement mentions the main biological functions in which sirtuin genes are involved, too many to describe them in detail. Then, at the end of the sentence, we refer

the reader to figure one, providing details and critical references. So we do not see a big problem in keeping it as is.

l.111: “sirtuin genes” => vertebrate or gnathostome genes?

We appreciate this comment. The answer is vertebrate sirtuins. We modified the text accordingly.

Figure 1: data concerns only human genes?

We understand this comment. But these genes have been well characterized mostly in humans, having solid evidence regarding their functions. For the sake of simplicity we only mentioned human sirtuins.

l.132: “paralogs” => vertebrate or gnathostome paralogs?

We appreciate this comment. The answer is vertebrate sirtuins. We modified the text accordingly.

l.197 “sirtuin lineages” do you mean gnathostome sirtuin paralogs?

We appreciate this comment. The answer is vertebrate sirtuins. We modified the text accordingly.

(this use of “gene lineage” instead of “paralog” also occur on l. 496

We appreciate this comment. We think that the use of paralog or gene lineage can be considered interchangeable.

and 502;

We appreciate this comment. The answer is vertebrate sirtuins. We modified the text accordingly.

l339-340: “evolutionary conserved” at which taxonomic range was this demonstrated? And do you mean comparing within human sequences or comparing human to other Vertebrates?

The reviewer is correct in pointing this out. However, because there is no systematic analysis of the evolutionary conservation of cellular locations and functions for all members of the sirtuin family, we tuned down our statement. The new text reads as follows “Some SIRT proteins, such as SIRT1, SIRT2 and SIRT3, possess evolutionary conserved, distinct cellular localizations and functions, as evidenced by their analysis in several model species that include Drosophila melanogaster, Mus musculus and Homo

sapiens (McBurney et al. 2003; Blander and Guarente 2004; North and Verdin 2004; Michishita et al. 2005; Haigis et al. 2006 and Fig. 1)"

As the figure 1 is presented, the genomic markers can be confusing specifically with SIRT3 and SIRT3-like gene markers (respectively RIC8A and RIC8B) because one does not know if these genes duplicated through whole genome duplications and therefore the whole loci is implicated or whether genome annotations are incorrect and it is indeed the same gene annotated "A" in one species and "B" in another. Therefore, if the synteny data will not be used subsequently it should be removed.

We understand this concern; however, we think it is helpful to maintain syntenic genes because it represents an alternative way to define the "identity" of a sirtuin gene in the case no phylogenetic analyses are available.

To verify your hypothesis on the belonging of the lamprey sequence to the clade SIRT3 a phylogeny of the marker Bet1 should reveal whether the markers (in human and lamprey) are indeed orthologous. Also give the reference to this Bet1 gene in lamprey.

The reviewer is correct. The Ensembl orthologous prediction algorithm recognizes as 1 to 1 orthologous both human and sea lamprey BET1L genes, providing support to our phylogeny.

http://www.ensembl.org/Homo_sapiens/Gene/Compara_Ortholog?db=core;g=ENSG00000177951;r=11:167784-207399

In addition to the information provided by Ensembl, we performed a phylogenetic analysis including BET1 and BET1L sequences from representative species of all main vertebrate lineages. This result agrees with the orthologous prediction provided by Ensembl.

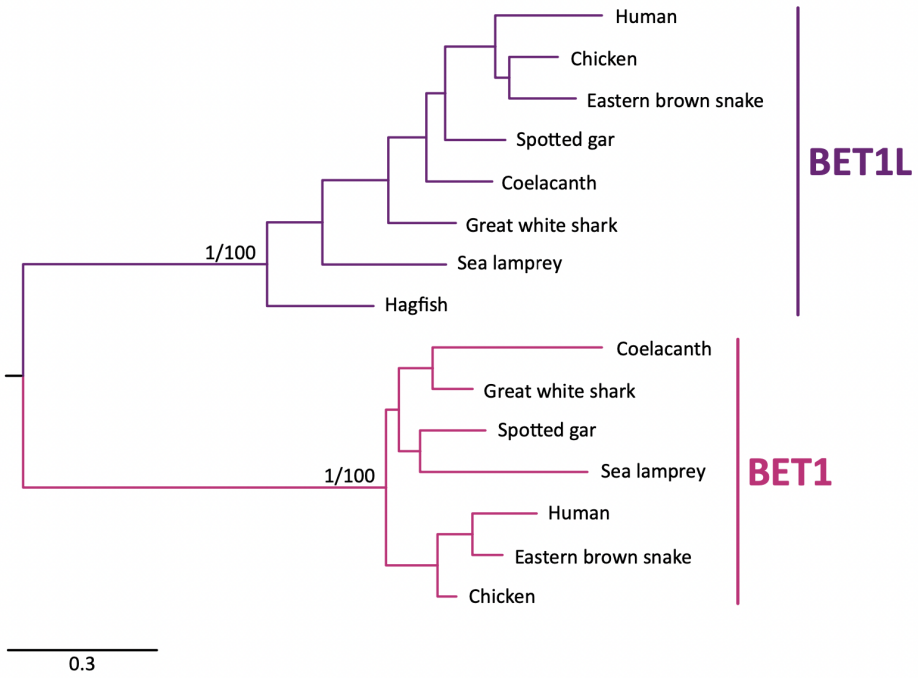

Figure 1. Midrooted maximum likelihood tree showing sister group relationships among Bet1 and Bet1l genes of vertebrates. Numbers on the nodes correspond to support from the aBayes and ultrafast bootstrap values. The scale denotes substitutions per site and colors represent gene lineages.

Finally, all this information will not be necessary for the manuscript, given that the new phylogenetic analyses, including sirtuin sequences from lungfish (as requested by the first reviewer), provide strong support for the monophyly of SIRT3 gene lineage, including the lamprey sequence.

Similar comment on your argument on lines 217-219.

We believe this is a different situation, as in this passage, we are describing the structure of a chromosomal region of the sea lamprey, not invoking orthology.

In the Figure 5, text is too small and figure 5b can be removed as the results do not reveal additional information compared to Fig 5a.

We believe that the dot plots prove that the SIRT8 gene is not present in the amniote species we show in figure 5a. Accordingly, we prefer to keep panel 5b.

Genome assembly version used for Fig5a should be stated in Materials and Methods or in the figure legend.

The reviewer is right. Here is the information that was added to the figure legend.

Human: GRCh38.p13
Opossum: ASM229v1
Chicken: GRCg6a
Gharial: GavGan_comp1
Red-eared slider: CAS_Tse_1.0
Green anole: AnoCar2.0v2
Tropical clawed frog: UCB_Xtro_10.0
Coelacanth: LatCha1
Spotted gar: LepOcu1
Elephant shark: Callorhinchus_milii-6.1.3

Regarding the phylogeny, the outgroup should be visible and a valid justification for the use of NNT should be stated.

We believe that the exclusion of the outgroup in figure 2 is not a big deal. We did not include it because it gives us more space to show the message we want to show to the reader. If readers are interested in seeing the tree with the outgroup, they can check our supplementary material.

We agree with the reviewer that we need to include a valid justification for using NNT as an outgroup. Accordingly, we included a sentence in the methods section.

Additionally, the separation between clade ‘1, 2, 3, 3-like’ on one hand and ‘4, 5, 6, 7’ on the other is not justified since that specific node is poorly supported (0,428/74) and may better be considered as a trifurcation.

We agree with the reviewer that this node is not supported according to the abayes routine and moderately supported according to the ultrafast bootstrap support routine, and it could be considered a trichotomy, although the phylogenetic signal of the alignment is what it is.

According to our new analyses, a deep node also received no/moderate support, implicitly saying that more work is needed to solve the sister group relationships among gene family members.

Was the alignment cleaned (with HmmCleaner for instance) before phylogenetic inference ? This could significantly increase robustness in the analysis by removing low similarity segments from the alignment.

No, we did not clean the alignment before phylogenetic inference. We tried to use HmmCleaner but had several problems installing and executing it. We sent emails to the

paper's first and last authors but never received an answer. We tried other options like BMGE, GBlocks, Noisy, and Trimal; however, in all cases, the cleaning procedure resulted in a very short alignment with little phylogenetic information. After that, we think that keeping the phylogenetic tree based on the original alignment seems to be our best option.

There are no details on the length of the alignment used for the phylogenetic reconstructions (write it on the figure legend ?).

We understand this concern; following the suggestion, we included the information regarding the length of the alignment.

Concerning the second part of the study, including a detailed cellular and biochemical study of the C. milii protein, this is outside my range of expertise. I have no comments on this part of the study.

Minor issues to note:
SIRT should be in italic when talking about the gene, and normal when talking about the Protein.

Done

l.76-77 “fulfill the biological functions with a different combination of paralogs” Could you develop by giving an example?

The study of Gitelman (2007)(PMID: 17567594), represents an example. In this study, twist genes, which are essential for embryonic development, were studied in different species of animals. The author shows that in species of the genus Xenopus, twist2 was lost, but the twist1 paralog acquired, and therefore preserved, twist2 function. Thus, this example shows that species with different gene repertoires are able to fulfill the biological functions associated with the twist gene family. We included this reference at the end of the sentence.

l.110 “C milii as a representative species” Please state representative of what clade

We appreciate this comment. We think this information is already contained in the sentence in which this expression is included. The sentence already written on the manuscript reads as follows: “We also aim to characterize the subcellular localization, enzymatic activity, and mitochondrial activity of the newly discovered SIRT8 gene lineage in the elephant shark (Callorhinchus milii) as a representative species”

And justify why you chose this species for subsequent analysis.

We used this species because it was the one in which we have more confidence in the curated annotation. This was important because subsequent in vitro analysis relied on commercial synthesis of the curated cDNA instead of obtaining it from tissue of the species.

Figure 1 : Please specify that protein size is in amino-acids in the figure.

Done

l.134 please provide the support value

This passage was removed from the manuscript, given the results of the new phylogenetic analyses.

l.136-137 “could be explained by the sequence itself, instead of the well-known compositional bias of cyclostome genomes” Please specify what is the problem with the sequence. Please also consider that you may not be expecting a one-to-one orthology relationship between the lamprey and gnathostome sequences (doi: 10.1101/gr.184135.114)

This passage was removed from the text, given that the new analyses, including lungfish sequences, provided strong support for the monophyly of the SIRT3 clade, including the lamprey sequence (0.92/95).

l.143-144 If syntenic markers are conserved (and orthologs cf above) between human and lamprey, it should be enough to state that this sequence is indeed a SIRT3 gene.

This passage was removed from the text, given that the new analyses, including lungfish sequences, provided strong support for the monophyly of the SIRT3 clade, including the lamprey sequence (0.92/95).

Figure 3: the sentence “the scale denotes substitutions …lineages.” Appears twice in the Legend.

Thank for noting this mistake, it was fixed

l.237-248 This paragraph should be reformulated, as such the message is not clear.

We are not sure what is confusing in this paragraph. We are mainly trying to interpret the meaning of gene loss.

l.279 What do you mean by interparalog distance?

This is when we are comparing distances among paralogs i.e. SIRT3 and SIRT8

Also, why is SIRT-3 of the elephant shark compared to the spotted gar while SIRT3-like is compared to the coelacanth?

Here we are just showing the extreme values, all comparisons were performed.

l.282-289 here, either you say what are the criteria to reach these descriptors (eg Class I) and test them, or you simply don't need these lines

We think that the rationale for these inferences are present in the mentioned statement. We say "based on how sirtuin genes are evolutionarily related and the information already known for the other family members……..."

l.318 "seems not to be unexpected" Please avoid double negations.

The reviewer is correct. We reworded this sentence and now it reads as follows: "This observation agrees with the literature as for SIRT3 there are many….."

Figure 7 Can be moved to supplementary material.

We prefer to keep the figure in the main text; we think it reveals important information, especially for the newly discovered gene family member.

Please better specify the "normalization" step.

According to the reviewer's suggestion, we added more details regarding the normalization process in the method section.

I would consider changing the Figure legend to "Heatmap representation of within-species relative transcriptional levels of sirtuin paralogs between 7 chosen tissues"

We changed the figure legend according to the reviewer's suggestion.

l.329 delete "homology"

Done

l.335 specify the "central" position, except if it is the positions cited in the next sentence (not quite clear in the current version)

We appreciate this comment. To address this issue, in this sentence we added text that specify the amino acids comprising the central portion of each SIRT8 analyzed.

l.335-340 Spacing format issue.

Revised

l.361-2: please specify that P63 is a cytoskeleton-associated protein 4 and TGN46 a Trans-Golgi Network Glycoprotein 46.

We appreciate this comment and have added the recommended UniProt name of P63, which in fact is cytoskeleton-associated protein 4. However, for TGN46 UniProt recommends the use of Trans-Golgi network integral membrane protein 2 and as such we have modified the text accordingly.

Figure 8c: TGN46 appears to colocalise with SIRT3-like as the blue fluorescence is different from that of non-transfected cells. Could you explain this ?

We are not sure what the reviewer is pointing out with the statement "blue fluorescence is different" in transfected and non-transfected cells, and the relationship of that statement with the statement "TGN46 appears to colocalize with SIRT8". Nevertheless, we think that the reviewer's concern relates with both a possible effect on the Golgi apparatus of transfected cells and a possible colocalization of SIRT8 and TGN46. Although we did not formally compare the pattern of fluorescence of the cellular compartments that we analyzed in transfected and non-transfected cells (mitochondria, Golgi apparatus, nucleus and endoplasmic reticulum), our implicit conclusion is that the overall pattern of fluorescent signal detection for TGN46 in transfected and non-transfected cells is the same, i.e., a concentrated perinuclear (semi-spherical) array of punctate structures. The Golgi apparatus is expected to show an overall similar structure between different cells, meaning that among different cells it is expected to observe some level of diversity in Golgi apparatus arrangement that include lack of spherical shape. Moreover, among different cell lines, the arrangement of the Golgi apparatus, the so-called morphology of the Golgi apparatus, varies considerably including the appearance of what is known as "Golgi ribbon". On the other hand, the SIRT8 fluorescent signal displays a convoluted array of tubules and punctate structures that extends beyond the perinuclear area. To conclude that SIRT8 colocalizes with TGN46 extensive superposition of fluorescent signals should be observed. In contrast, the fluorescent signal of MitoTracker Orange displays a convoluted array of tubules and punctate structures that extensively superpose to that of SIRT8. Thus, we stand our conclusion that SIRT8 only colocalizes with MitoTracker Orange. However, to provide a quantitative assessment of colocalization, we now include an unbiased analysis that confirms our initial conclusion. We included the analysis method and the results of the analysis in the manuscript, and prepared a supplementary figure with the results. Regarding a possible effect of SIRT8 expression on the structure of the Golgi apparatus or other compartments, an evaluation of its fine structure by electron microscopy is necessary, which we think is beyond the scope of our present manuscript. However, we performed a quantitative analysis of fluorescence microscopy images of the shape of the Golgi apparatus in transfected and non-transfected cells, as indicated by TGN46 staining

of images similar to those shown in Figure 8c and 8d. We estimated the circularity index of the Golgi apparatus, as implemented in the ImageJ software, and found non-significant differences when using a two-tailed, paired t-test. We do not include these data in our revised manuscript because, although informative, it is not sufficient to conclude that the expression of SIRT8 does not affect the structure of the Golgi apparatus.

l.432 Please change “much potent” to “much more potent”.

Done

l.519 & 523 Please avoid back to back parentheses.

Done

l.544 Please state how many ultra-fast bootstrap replicates were performed.

Done

Reviewed by anonymous reviewer
This study rexamines and clarifies the evolutionary history and repertoire of the multifunctional vertebrate Sirtuin gene family. The authors find evidence for a new family member that has been lost in amniotes. Discovery of such previously unknown genes is not especially unusual in vertebrate gene family studies that broadly sample genome-data from across vertebrate phylogeny, however the extensive effort made in this study to shine a light on the function of this new gene is both rare and highly commendable.

I have no major concerns about the study, although I note that I lack sufficient expertise to critique the functional experiments, and only a few minor points that the authors may wish to consider:

1. Regarding the placement of SIRT3 the authors state "This situation could be explained by the sequence itself, instead of the well-known compositional bias of the cyclostome genomes" (line 136-137). Is it not possible that the opposite to this is occuring here, with compositional bias affecting the placement of the lamprey sequence but not the hagfish?

We appreciate this comment. Although we believe that this discussion will not be necessary for the new version of the manuscript. The new phylogenetic analyses, including sirtuin sequences from lungfish (as requested by the first reviewer), provide strong support for the monophyly of SIRT3 gene lineage, including the lamprey sequence.

Similarly (although it seems unlikely) from the results shown it seems that paralogy of the lamprey sequence to SIRT3 cannot be ruled out?

Based on our new phylogenetic analyses, it seems that we can say confidently that this lamprey sequence is SIRT3. Additionally, the phylogenetic analyses of BET1 and BET1l (see a response to the second reviewer) also point in this direction.

2. It might be worth noting that the 'high transcription levels' in ovary for the newly discovered SIRT3-like, although still interesting, lack biological replicates at least within species (especially given the Zebrafish findings).

We agree with this comment. To notice this issue we included the following sentence "These results should be taken with caution given the lack of biological replicates, at least within species."