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**
- Does the abstract present the supported findings of the study concerned and no other? **Yes**
- Does the introduction clearly explain the motivation for the study? **Yes**
- Is the research question/hypothesis/prediction clearly presented? **Yes**
- Does the introduction build on relevant recent and past research performed in the field? **Yes**

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**
- Is the experimental plan consistent with the questions? **Yes**
- Are the statistical analyses appropriate? **Yes**
- Have you evaluated the statistical scripts and program codes? **NA**

Results

- Have you checked the raw data and their associated description? **Yes**
- Have you run the data transformations and statistical analyses and checked that you get the same results? **No**
- To the best of your ability, can you detect any obvious manipulation of data (e.g. removal)? **No**
- Do the statistical results strongly support the conclusion (p< 10-3 or BF>20)? **Yes**
- In the case of negative results, was a statistical power analysis (or an appropriate Bayesian analysis) performed? **NA**
- Did the authors conduct many experiments but retain only some of the results? **No**

Discussion

- Do the interpretations of the analysis go too far? **No**
- Are the conclusions adequately supported by the results? **Yes**
- Does the discussion take into account relevant recent and past research performed in the field? **Yes**
- Did the authors test many hypotheses but consider only a few in the discussion? **No**

References

- Are all the references appropriate? **Yes**
- Are the necessary references present? **Yes**
- Do the references seem accurate? **Yes**

Tables and figures

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

- Are all the tables/figures useful? **Fig5b can be omitted or moved to Supplementary materials**
- Are there too many/too few tables and figures? **No**
- Do the tables and figures have suitable captions such that they can be understood without having to read the main text? **Yes**

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.

Flaws, weaknesses and suggestions for improvement

Shouldn't we call this gene SIRT8? As just one additional paralog ?

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

l.81: "variety of biological functions"

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

Figure 1: data concerns only human genes?

l.132: "paralogs" => vertebrate or gnathostome paralogs?

l.197 "sirtuin lineages" do you mean *gnathostome sirtuin paralogs*? (this use of "gene lineage" instead of "paralog" also occur on l. 496 and 502;

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?

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.

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. Similar comment on your argument on lines 217-219.

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

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

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

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.

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

l.110 “*C milii* as a representative species” Please state representative of what clade and justify why you chose this species for subsequent analysis.

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

l.134 please provide the support value

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)

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.

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

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

l.279 What do you mean by interparalog distance? Also, why is SIRT-3 of the elephant shark compared to the spotted gar while SIRT3-like is compared to the coelacanth ?

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

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

Figure 7 Can be moved to supplementary material. Please better specify the “normalization” step. I would consider changing the Figure legend to “Heatmap representation of within-species relative transcriptional levels of sirtuin paralogs between 7 chosen tissues”

l.329 delete “homology”

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

l.335-340 Spacing format issue.

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

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 ?

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

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

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