Dear Editor and reviewers,

Thank you for your interest.

We have now considered, and hopefully addressed, all the comments from the reviewers in a point-by-point response.

Note that after considering important points from the reviewers and because new studies have been published since our first submission, we decided to reanalyse and rewrite a large part of the manuscript.
The conclusion of the manuscript did not really change, but the angle from which we approach it did.

Because the changes were too major, we are providing a new manuscript without the track changes from the previous version.

I apologize for the time it took to make the revisions.

Best,
David

Response to the review by Dr. Julie Jacquery

In this study, Duneau and colleagues investigate sperm morphology in 15 Daphnia species and map the morphological data on a Daphnia phylogeny. They conclude that increase in sperm length has evolved twice, and that sperm encapsulation has been lost in a clade. Globally, this article is based on few data (measures of sperm length in 15 species), statistical analyses (to support the authors' conclusions) are absent and the discussion is mainly speculative.
Thank you for your thorough comments. We detail below few of the points mentioned here but we would like to say that indeed more species could have been useful (the more the better) but, getting the clonal lines of further species and then getting the sperm from these Daphnia is not an easy task (males are rare to start with). We believe that our choice of species through the phylogeny was suitable to tackle our question.

My main concern is that the authors conclude that sperm length evolved twice in Daphnia, but no statistical analysis supports this conclusion. The authors reach this conclusion through a visual inspection of figure 1. Statistical analyses to support for the conclusion of increased sperm length in the two clades are required.
We have now done the full analysis that supports our claim. This makes our study much stronger. Our dataset allowed to define statistically that there are statistical differences in sperm length and the phylogenetic analysis showed that sperm length has diverged twice.

The authors also hypothesize that the ancestral state in Daphnia is small sperm. I question whether this hypothesis is meaningful for different reasons: 1) Assuming that small sperm is ancestral is almost as parsimonious as the opposite (based on the phylogeny presented in figure 1). 2) Sperm length in Ceriodaphnia (used as an outgroup in the phylogeny) ranges from 2 to 6 um (these values are mentioned in the text, line 106, but are unfortunately not shown on the phylogeny). Therefore, sperm length in Ceriodaphnia is more or less intermediate between the clades with small sperm and a clade with large sperm (the Daphnia sensus lato clade). It is therefore difficult to conclude whether the ancestral state is small or large sperm length. 3) Sperm length in Daphnia is much smaller than in

other taxa (figure 2 from the review of Lupold and Pitnick 2019, https://doi.org/10.1530/REP-17-0536). Hence, it is likely that selection has favored a reduction in sperm length in the Daphnia group.
This is an excellent point and one that makes the statistical analysis recommended very useful. We have now included the mean of the outgroup as described in the literature. However, it does not allow to test statistically the ancestral state. We have been careful to not say whether evolution was toward an increase or a decrease in length, only that it diverged twice.
As we did not have access to *Simocephalus* and *Ceriodaphnia* sperm, hence could not measure any, we can't add them in our graph.

We now position our results on sperm length in the context of SpermTree, the most recent and complete database on sperm size through the animal kingdom. Some *Daphnia* species have among the smallest sperm, but not all species have particularly small sperm considering species with aflagellated sperm. We discuss this in the manuscript.

I also find that most of the discussion is speculative (too far from the results obtained), the authors make many hypotheses to explain the different sperm morphology in Daphnia. The discussion should be more focused on the results obtained.
We reduced the discussion and focused it more on the results.

In addition, the authors do not discuss an important result: the fact that sperm length is extremely small in Daphnia in comparison to all other animals (see data from Lupold and Pitnick, 2019).
This is a good point. Some Daphnia species have among the smallest sperm recorded (~2 µm). We have now positioned Daphnia sperm length more in the context of sperm length in other taxa.

Minor comments:
I am not familiar with the Daphnia taxonomy. I don't know what is the Daphnia sensu lato clade and the Ctenodaphnia clade (mentioned in the abstract and in the text).
We have now added an in-depth explanation of different clades and their relationships.

Ceriodaphnia sperm length should be included in figure 1.
We cannot add it to the graph because as we did not measure sperm from this species, but we did add it in the phylogeny using data from the literature.

Lines 25 and 120: Ejaculate size is imprecise. Do you mean sperm number? Or the volume of the ejaculate?
We meant sperm number and we corrected it.

How many males per species were measured? How many sperm were measured per ejaculate? This should be included in the methods. This is relevant because in another article (https://doi.org/10.1101/2020.02.05.935148, Figure 6B), Duneau et al evidence large inter-individual variation in sperm length in D. magna.
This is now included in the figure. Note however that the variation within ejaculate found in *D. magna* in our other study is much smaller than the variation found here between species. It is extremely rare to find a sperm of 2-4µm in *Daphnia magna* and there is no large sperm in an ejaculate from species with very small sperm.

Male genital papilla drawings from published keys are shown on figure 1, and the reader has to wait until lines 146-149 (discussion) to understand why the authors have included these images. This part could be better integrated in the manuscript.
We address it fully in the new version of the manuscript.

Line 37: It is unclear in this sentence what is triggered by a change in environmental conditions.
We rephrased it to clarify.

Lines 173-174: too many brackets.
We corrected it.

The association between the presence/absence of a sperm capsule and the way males were induced (naturally, or artificially through the addition of methyl farnesoate) is almost perfect. Could it be possible that sperm morphology is affected by the male induction process? Or that male maturity differs between naturally and artificially produced males?
The impact on chemical on male sperm is an important concern indeed. However, it is unlikely that the male chemical induction explains our pattern because the treatment is on the mother and not on the male. We studied sperm morphology when males were more than a week old as they would not ejaculate before they maturated completely (i.e. their chest is opened). Furthermore, male production was also induced for *D. curvirostris* and their sperm had a capsule.

Response to the review by Dr. Renate Matzke-Karasz

Thank you for your thorough review of our manuscript.
We won't address point by point the critics about the quality of the description of the taxonomy as they were in several places. We have taken all the criticism and we do not have any rebuttal. We hope to have addressed it in this new version.

Taxonomy
The authors studied sperm of 15 species of the genus *Daphnia*. However, in the abstract already, they mention a *Daphnia sensu lato* clade and a *Ctenodaphnia* clade without explaining anywhere in the manuscript the concept behind these terms.
As much as I understand (I am not a branchiopod-specialist), the genus *Daphnia* has three accepted and valid subgenera, i.e. *Daphnia* (*Ctenodaphnia*), *Daphnia* (*Australodaphnia*), and of course the nominotypical subgenus *Daphnia* (*Daphnia*), since it contains the type species of the genus *Daphnia*, i.e. *D.* (*D.*) *longispina* as pointed out by Šrámek-Hušek, Strašcraba & Brtek (1962). This seems to be covered by the zoological nomenclatural rules (see ICZN Art. 44.1), which clearly are the basis for any taxonomical, phylogenetic and evolutionary study.
If the above concept of the genus *Daphnia* and its subgenera is correct, I do not understand why the authors use the terms '*Daphnia sensu lato* clade' and '*Ctenodaphnia* clade'. Could it be that they actually mean '*Daphnia sensu* ***strictu*** clade' and '*Ctenodaphnia* clade' and used these terms in the sense of/instead of subgenera? But then again, why wouldn't they use the terms subgenera? This is very confusing, and it's not the only confusing thing: in the course of the manuscript, the authors use the terms clade, group, sub-group, species complex, subgenus, genusX sensu strictu, genusX sensu lato without explaining any of these entities.
The terminology of Daphniidae phylogeny can indeed be confusing. We now added a detailed description of the taxa, clades and their relationships.

I suggest that the authors include a short chapter on the taxonomical status of 'their' species to avoid any misunderstanding and explain how they grouped the species and why.
It is now included in the first paragraph of the methods.

Another problem I see is the treatment of *D. longispina*, *D. zschokkei* and *D. hyalina* in the manuscript (see my comments there). Petrusek et al. 2008 synonymized *D. zschokkei* and *D. hyalina* (and *D. rosea*, which is not part of the present study) with *D. longispina* on the basis of COI and 12S rDNA genes, retaining the name *D. longispina* because of its priority.
In their chapter Materials & Methods, the present authors explain that "*D. hyalina* and *D. zschokkei* are now synonymous of *D. longispina* and should be understood as *D. longispina 'hyalina'* and *D. longispina 'zschokkei'* (Petrusek et al., 2008), hence we merged them on the same branch in the cladogram." However, Petrusek et al. (2008) did not propose such terms, they clearly synonymized the species according to the rules of the ICZN. If the authors do not agree with Petrusek et al. 2008, they should at least discuss it and explain, what they actually intend to express using the terms *D. longispina 'hyalina'* and *D. longispina 'zschokkei'*, since this format is not one covered by the ICZN. Are the authors referring to morphotypes of D. longispina? Or do they intend to introduce subspecies (but then the quote signs don't make sense). Given the synonymization by Petrusek et al. 2008, the authors should explain on the basis of which characters they actually identified their individuals of these 'forms', which they took from lab cultures.
Besides, I think that this topic could have been put much more into the focus of this study by comparing the sperm data of these three formerly separate species. Can sperm morphologies support maintaining them as separate species or support the idea of morphotypes of a single species? Would cross-fertilization tests be feasible?
We now mentioned in the methods that we have data from the different groups that were formerly considered as different species, but we consider them together as *D. longispina* in the analysis.

Returning to taxonomy, I highly recommend that all species named in a manuscript have the (correct) species authorships attached when first mentioned. Species are scientific concepts that must be acknowledged by citing the authorship like in any other scientific concept. Besides, and particularly in groups with a complex history of nomenclature, this will avoid any misconception. In this context, don't miss out on Kotov 2020, Zootaxa 4803 (3): 591–599
We have now added the species authorship.

Image quality
Given that the very purpose of this study is the investigation of sperm morphology (incl. size) it is surprising that the microscopic images are of low quality and that the maximal microscopic magnification used was 40x. With sperm of only a few µm length I would expect a 100x magnification, if not TEM anyway. Consequently, the information on morphological traits is not fully exploited and thus the results and their interpretation must remain relatively superficial.
We do not think that our results are superficial.
Our images may not be as pretty as in other studies from specialised laboratories looking at the different sperm structures, but, the resolution of our images allowed to quantify appropriately the size at a level that allowed to show two very distinct groups. A higher resolution may have allowed to be more precise in the measurements, but it is very unlikely that it would have changed the outcome. Furthermore, we also had to trade image quality with sperm quality. We wanted to have sperm as soon as they were expelled and for this reason they were moving and could not be fixed. For these reason, we did not have the mean to take pictures and measure properly at higher magnification (e.g. x100). TEM could have informed about the structures and maybe the interpretation on their role, but Wuerz et al., 2017, which did such thorough work, was not able to conclude much more on the biology of the sperm of *Daphnia magna*.
The aim of our study was to investigate sperm size, and we think, considering our results, that the approach was appropriate, especially as it even allowed us to discover the presence of filopodia.

The microscopic images of the sperms are presented in very tiny items in an overloaded Fig. 1 and the important, bigger and more meaningful images are pushed into the supplementary material, which I do not recommend.

In our opinion, the message of the Fig. 1 are the phylogeny and the graph. The tiny items allow to understand it better. It is a way to remind the reader what they can see more in details in supplementary. Furthermore, adding all the bigger illustrative, not necessarily pretty, images in a main figure would unnecessarily increase the size of the manuscript.

Other issues

The authors should generally be more precise in phrasing. Why is there no discussion of their results in the light of sperm biology of other crustaceans (rather than mammals, ascidians and nematodes)?
Some conclusions drawn are not well-founded, see comments in the manuscript.
Formatting of figures must be improved.
We hope we have now improved the manuscript and the figures. We did not expand on speculation of sperm biology of other crustaceans, but we now used a different angle to approach our question which we believe make the discussion more general and closer to our results.

Other comments in the manuscripts

There were some extra comments in the manuscript/pdf which were not repeated in the review. We wanted to extract them here and summarize our changes.

- A 'larger ejaculate' could also be achieved by increasing sperm size, not only by more sperm cells.
We have now made it cleared.

- As a reaction against sperm being flushed out, not only the increase of sperm numbers would help (which is implied here). A morphological adaptatation could also be advantageous.
That was indeed what was implied in the following sentences. We fixed this.

- were compacted? have been compacted? It is unclear if you refer to the maturation mentioned in the previous sentence.
We have now made it clear.

- long? What about their width?
We didn't make the distinction, but we are now more specific.

- You did not explain the taxonomical background. Ctenodaphnia is a subgenus, why don't you explain this properly?
Indeed, we meant the group of species within the subgenus and it was unclear. We are now more specific.

- How many sperms did you measure per species? That should go into Mat&Met
We have now added this information in the figure.

-These are publications protected by copyright restrictions that must be followed. You will need licenses for re-use of their figures.
Thank you for pointing this out. We have now redrawn the picture and cited it as such.

- wouldn't it be the total volume of sperm material that is restricted by the size of the spermiduct?
This is a good point. We changed "number of sperm" to "amount of sperm material".

-These figure items are too small, and there is no indication what/where the genital papilla actually is in each item.

We address this by redrawing the structure and zooming on the important part. We indicated the region of the genital papilla with one arrow in the species where they are prominent.

- If more 'tickets' are needed (i.e. higher sperm numbers) for an increased fertilization success (as suggested in the previous sentence), why would they invest in larger sperm at all?
We have rewritten this part to be clearer.

- Wouldn't these filaments (or 'filapodia' in Wuerz et al. 2017) need at least microtubules, if they are supposed to play a role in 'crawling'? However, Wuerz et al. 2017 did not find any cytoskeletal elements in them, suggesting that the filaments "may function in contact and possibly fusion with the oocyte"
We have rewritten this part and removed the speculation about sperm mobility, we instead focused on their role in attachment.

- conserved evolution doesn't sound correct to me .. evolution is not conserved, but the trait is... maybe restricted evolution? limited evolution?
We rewrote entirely this paragraph.