University of East Anglia

Dr Lewis G. Spurgin
BBSRC Research Fellow
School of Biological Sciences
University of East Anglia
Norwich
Norfolk, NR4 7TJ
United Kingdom

L.Spurgin@uea.ac.uk
lewisspurgin.wordpress.com

June 2017

Dear Ben,

Thank you for taking the time to consider our manuscript for publication in *Journal of Animal Ecology*. We were very pleased that the editors and reviewers were all so enthusiastic about the quality of our paper, and were impressed by the detailed and constructive nature of the reviews. We have now been through the Associate Editor and reviewer comments in detail, and made substantial changes to our manuscript in response. Below we outline, point-by-point, how we have responded to each of the reviewer comments. We also reproduce the sections of our manuscript we have changed in response to the reviewer comments. Reviewer comments are in bold, and reproduced manuscript sections are indented and italicised.

Please do not hesitate to let me know if you require any further information or clarifications. I look forward to hearing from you.

Kind Regards,

Lewis Spurgin (on behalf of all authors)

Reviewer 1

This paper reports interesting findings on lifelong telomere dynamics and its environmental correlates in a wild bird species. This is a very interesting study, but a moderate/major revision would be needed before considering this paper for publication in my opinion. With such revision, I am convinced that this paper will be an important contribution in the field of telomeres in ecology and evolution. I have some comments/criticism about the methodology presented by the authors, especially about the poor repeatability found between qPCR runs (see major comment 1) and the low efficiency for the telomere amplification. In addition, one main conclusion reached by the authors (i.e. that lifelong telomere dynamics is different between cohort, see comment 6) is not supported by the data in my opinion, or at least by the stats presented in the paper. Despite such limitations, the research question is novel and merits to be addressed and published in high-quality journal such as Journal of Animal Ecology. I really hope that my comments will be useful to the authors in revising their manuscript.

RESPONSE: We thank Dr Stier for his constructive comments, and are pleased that he believes our work to be novel and worthy of publication. We agree that the main issues he raises were not properly addressed/explained, and respond to these points in detail below.

I am not sure to understand how you calculated the inter-plate repeatability of 0.68 mentioned line 234. Assuming that it is more or less equivalent to the intraclass coefficient of correlation r, it would mean that approximately 50% of the variation in your measurement is linked to the error of measurement. If my understanding is correct, it would mean that your qPCR methodology is quite poorly repeatable. Can you please clarify this point?

RESPONSE: The reviewer is correct that the estimate of repeatability is the intraclass correlation coefficient (this is made clear in the rptR documentation, which we cite). We agree with the reviewer in the sense that repeatability is not as high as we would have hoped, although we note that in the context of the telomere literature, a repeatability estimate of 0.68 is not especially low. However, because of how the samples were run, we are confident that any issues of repeatability have not led to any kind of bias in our results, and in fact make our analyses conservative. We respond to this point in detail, and outline how we clarified our discussion, in our response to the next comment.

You should mention in the methods how the samples were assigned to the different plates (i.e. totally random, balanced between years/age, or all samples of 1 individual on the same plate?). If you allocated more or less randomly your samples to the different plates, can you check (and present as supplementary) that you do not have a "plate effect" on the average TL measurement. If you do observe such plate effect (something not that unusual with qPCR. . .), it might contribute to explain your low repeatability, and maybe part of the telomere lengthening that you describe.

RESPONSE: We are confident that plate effects are not a source of systematic bias in our data. Firstly, we assigned samples completely randomly (i.e. using a random number generator) to plates prior to running the

telomere assays, minimising the chance that plate effects would introduce systematic differences among ages, sexes or cohorts. Secondly, our method of estimating RTL normalises samples across plates, as each RTL value is obtained by subtracting sample Cq values from Cq values obtained from a "calibrator" sample, which is run on every plate (information contained within Bebbington et al. 2016, cited). Thirdly, in response to a request by reviewer 3 we have now also included plate ID as a random effect in the models (see below), and our results are virtually identical.

With regards to the observed telomere lengthening, it is very possible that inter-plate variability accounts some of the observed lengthening in our dataset, but it cannot account for the amount and extent of lengthening we find. Indeed, this is exactly what Figure 2B shows. The lengthening observed "within" samples can be attributed to error across plates, but the additional lengthening observed "among" samples cannot be explained by measurement error.

We now explain our randomisation approach in the methods, line XX:

> *"Prior to qPCR, we used a random number generator to assign samples to qPCR plates, to ensure that no systematic bias could occur with regards to age, sex, cohort or ecological environment."*

We also include a discussion on error rate on lines XX:

> *"Measurement error is one factor that is an issue in our study, and other studies that use qPCR to measure telomere length (Nussey et al. 2014). Measurement error is unlikely to be a cause of type I error in our study, because we were careful to randomise all samples across qPCR plates, and normalise RTL estimates across plates. However, the noise associated with within and among plate measure may have resulted in a decrease in explanatory power, and possibly in a degree of type II error. Techniques for measuring telomere length with a greater degree of precision may probe helpful in future ecological studies of teloere dynamics, but at present there is still atrade-off between obtaining precise telomere measurements, and utilising the large sample sizes necessary for ecological study."*

1.78 (78%) would not be considered appropriate by many researchers for qPCR methodology (the commonly accepted limit is more or less + or minus 10%, so here between 90% (1.90) and 110% (2.10)). Your efficiency is more than 20% below the expected 100%, which raises some questions about the validity of your methodology. How can you be confident that this low efficiency is not affecting the quality of your data?

RESPONSE: Again, the level of efficiency is lower than we would have liked, although again suts fairly squarely within that found the published literature (see Fairlie et al. 2015, for a recent example), and often it is not even reported. We also note that this problem is not explected to be too drastic, because variaiton in efficiency is taken into account when calculating RTL from the Cq values. And again, we suggest that because of our study design, random error of this kind makes the majority of our analyses conservative, and is accounted for in the analysis of telomere lengthening.

Important information is missing in the methods section regarding the age and life stage at which birds have been sampled. Telomeres might shorten in a few days during the nestling period (e.g. Stier et al. 2015b, 2016), and might be affected by breeding effort (e.g. Reichert

et al. 2014, Sudyka et al. 2014), so it is important in my opinion to know more about the age and status of the birds used in this study.

RESPONSE: We now provide more clarification about how we determined and analyse age, lines XX:

> *"The majority of individuals are ringed (with an individually numbered metal ring and unique combination of colour rings) within the first year of life, and so are of known age. We aged all birds using information on eye colour at first capture (Kondeur et al. 1991) and previous capture history (Richardson et al. 2003). Within the first year of life, birds are classified as nestlings less than one month old (rounded to one month for analyses), fledglings less than six month olds (rounded to six months) or subadults up to one year old (rounded to 10 months). Ages for adult birds were rounded to the nearest year."*

We do not have more detailed data on age for the majority of our individuals, but given that we are studying broad, lifelong patterns of telomere length and dynamics, we feel that the resolution we have is sufficient.

Lines 145-147: Storage and extraction methods are known to potentially affect qPCR results (e.g. Reichert et al. 2017 in press Oecologia). Storing samples in 100% ethanol at RT is not very common among people measuring telomeres as far as I know. Do you have any papers supporting the relevance of such storage method? Can you provide some DNA integrity gel images as supplementary material? Were older samples more likely to be degraded? Such information would be helpful for the reader to assess the quality of your DNA and to have an idea of the appearance of normal vs. degraded DNA samples.

How can you exclude that the cohort effects observed in your study are not linked to differences in storage conditions / DNA quality-integrity rather than true biological variation? Can you try for instance to disentangle/quantify the proportion of variation linked to the year of sampling (I would not expect a strong biological year effect on adults) vs. the proportion of variation linked to the year of birth (cohort effect)?

RESPONSE: Storage in ethanol is standardd procedure for avian blood samples, and allows DNA to be maintained at high quality for many years. We checked DNA integrity for a very large number of samples, and found no evidence of DNA degration in older samples. We now include two gel images to illustrate this in the supplementary material (Fig. S1). We refer to this in the main text, lines XX.

> *We found no evidence of DNA degradation in older samples (Fig. S1).*

We also performed an analysis of RTL and sample storage time, using adult birds, and found no evidence of a relationship - lines XX.

> *Using adult samples greater than one year old, we checked whether RTL was related to sample storage time, and found no evidence of such a relationship (estimate = -0.002, CIs = -0.007, 0.002).*

Would it be more relevant to present Figure 1B only with telomere data from fledglings individuals of the cohort in question. I am assuming from data presentation than it is an average of TL of all the individuals from these cohorts throughout their life, but I might be

wrong. Presenting average TL data over the life of individuals might be somewhat misleading here in my opinion since some cohorts might experience poorer survival than other, etc... In addition, the main cohort effect seems to be on telomere length per se, rather than on telomere dynamics with age (see comment 6 below).

RESPONSE: Figure 1B did not show averages for individuals, but rather all measures for all individuals from each cohort (i.e. including repeats from individuals). The aim here was to show the full distribution of telomere lengths for each cohort. However, on balance we agree that restricting this plot to juveniles is probably less biased. We have now altered the figure accordingly.

Cohort effects (year of birth in analyses of nestling TL) have been reported without being clearly discussed (e.g. Stier et al. 2014, Becker et al. 2015). It might be worth mentioning it in the discussion, and pointing out the difficulties in distinguishing between real cohort effects and potential methodological artefacts linked to sampling, storage, extraction and/or qPCR methods. Were the different cohorts well randomized between qPCR plates?

RESPONSE: See response above re. randomisation. While not an issue for our study, we agree that it is a point worth raising. We now discuss this issue on lines XX:

> *"Although a few studies have shown that temporal variation in telomere dynamic-soccurs in natural populations, these have been limited in the number of seasons they cover (Mizutani et al. 2013; Watson, Bolton and Monaghan 2015; Fairlie et al. 2016). Other studies have found chort effects but not discussed them in an ecological context (Stier et al. 2014; Becker et al.2015). One problem with studying cohort effects is that it can be difficult to tease apart true cohort effects from effects that arise due to sample degradation with aim, and/or batch effects in telomere assays, although neither of these factors were a problem in our study. Indeed, the long-term Seychelles warbler dataset has allowed us to show that temporal variation in telomere dynamics can occur over substantial time periods."*

You mention in several places that telomere dynamics with age is different between cohorts (e.g. lines 241-244; 299-300; 350-54), but in my opinion this is not supported by your data; so such conclusion should be avoided. Indeed, in the first analysis of RTL, including the log age * cohort interaction only marginally improved the fit of the model (delta AIC of 1.40, lines 242-244), which is not sufficient to support your conclusion. In addition, you mention that including cohort ID in the within-individual analyse of TL was even decreasing the fit of the model to your data, again suggesting that telomere dynamics is not significantly impacted by the cohort. Please remove any reference to such effect or try to provide convincing analysis justifying your conclusion.

RESPONSE: We agree that while there is convincing evidence for cohort effects on telomere length, the evidence for cohort effects on telomere dynamics is much weaker. We have therefore removed references to this effect throughout.

While I agree that your results are very interesting and biologically meaningful (age effect, sex effet, link with body size and food abundance), and that they are very similar to those on Soay sheep from Dan Nussey's team, an intra-individual repeatability of 8% is a bit worrying in my opinion. This is especially worrying for using telomere length/dynamics as a "biomarker" as you suggest it in several places in the manuscript. A true biomarker should be at least to some extent consistent within an individual to be useful. You should raise such limitation in your discussion.

RESPONSE: We are not sure that "worrying" is the correct phrase to use, as we anticpiate that other studies in natural population will find similar patterns. Further, we disagree with the sentiment that a biomarker must be consistent over individual lifespans to be useful. We feel that we cover this point in the discussion, lines XX:

> *"Such inconsistent changes in telomere length over lifespans could occur due to changes in the composition of cell types within individual samples, or due to the actual elongation of telomeres (Blackburn et al. 1999(. Determining the mechanism of these changes is essential for how we view telomeres as biomarkers of costs. For example, if telomeres can be lengthened in response to improvements in environmental conditions, this would suggest that they reflect short- to medium-term costs, rather than the cumulative costs that an individual has faced over its lifespan (Bateson et al. 2016). New statistical and technical approaches are therefore now required to determine the mechanisms behind increases in telomere length within individuals, so that biologically informed hypotheses about the ecological causes and consequences of these increases can be generated and tested."*

You mention that in other bird studies such within-individual consistency was much higher because they were done in the lab/nest-boxes. Another major difference with these studies is the shorter time-frame on which they were conducted, in addition to the more controlled conditions to store the samples. Therefore you would expect less noise in theses studies linked to methodological artefact due to sample storage/integrity, which might contribute to explain your low within-individual consistency.

RESPONSE: It is true that the shorter timescales of these studies may be the reason for the higher within-individual consistency, and we have now added a note to this effect, lines XX:

> *"Consistent with a pattern of sporadic changes in telomere length with age, we found that within-individual telomere measurements were only weakly correlated. These findings are in contrast to other avian studies in which within-individual telomere length measurements were highly consistent, and individual-level telomere shortening occurred throughout the juvenile period and into adulthood (Heidinger et al. 2012; Boonekamp et al. 2014). However, the lifelong telomere dynamics found in Seychelles warblers are strikingly similar to those found in Soay sheep (Fairlie et al. 2016). This discrepancy in results may be because in our study, and that of Fairlie et al. (2016), individuals were born and reared in the wild, as opposed to in nestbox or laboratory conditions. Alternatively it may be because our longitudinal telomere measurements have been taken over longer time periods."*

However, as we argue above, we do not think it is likely that the pattern is a result of sample storage/DNA integrity. Past discussions with the Nussey group suggests that they are also confident their telomere assay is not negatively affected by long-term sample storage.

Body mass is known to be less affected than tarsus length by bias linked to different experimenters in long-term studies like yours. Do you have any body mass data, and if yes why not including it in this paper?

RESPONSE: We do have body mass data, but decided to exclude it for a number of reasons. Firstly, with tarsus length we wanted to control for variation in structural size - even with measurer effects, tarsus length is a better indicator of structural size than body mass. Body mass is more of a measure of (short-term) condition, and varies according to the time of day, season, etc. In some respects body mass is likely to respond in a similar way to the environment as telomere length (especially if telomeres can get longer). There is thus a circularity issue with including bodymass as a predictor of telomere length. Finally, from apractical perspective, including body mass introduces collinearity issues, particularly with tarsus length, but also with other ecological variables.

Line 16: Telomeres shorten in response to oxidative stress and cellular division. It is important to mention that here since you find higher telomere shortening very early in life that might be linked to cellular division rather than oxidative stress.

RESPONSE: Yes, this is a good point. We now clarify in the discussion, lines XX

> *"Our study adds to the substantial body of literature from humans and wild animals showing that telomere length decreases with age, and that this decrease is most rapid in early life - most likely as a consequence of cellular division..."*

Line 30: "Telomere length and the slope of its relationship with age" is not a straightforward formulation in my opinion. You could use "Telomere length and the rate of telomere change with age", it has more biological meaning in my opinion.

RESPONSE: We have now removed this sentence, as suggested in one of the previous comments.

Line 44: "environmental costs" is not clearly defined, so please rephrase or define clearly what you mean by environmental costs here.

RESPONSE: We have changed the phrase to "individual responses to the environment".

Lines 55-56: While I acknowledged that Von Zglinicki 2002 is the common citation used by people to support the role of oxidative stress in telomere shortening, this is in vitro work with cultured cells using supra-physiological doses of ROS. I would suggest to mention that this relationship is observed "in vitro" or provide a reference supporting such assumption in vivo (if you find one, because I am still looking for solid evidence of this relationship..).

RESPONSE: We have clarified that this is in vitro.

Lines 64-65: I am not sure that Horn 2010 and Simons 2015 are appropriate here, however you could cite Stier et al. 2015 Exp. Gerontol. reviewing relationships between age and TL as well as TL and survival in non-mammalian vertebrates species.

RESPONSE: We have changed the citations as suggested.

Line 82: To help a bit the naïve reader, I would give here some recent examples of "environmental conditions" that are known to affect telomere length; for instance anthropogenic noise (Meillère et al. 2016), altitude (Stier et al. 2016), heat stress (Simide et al. 2016) or infection (Asghar et al. 2016).

RESPONSE: We have added in the suggested references.

Lines 231-234: information on sample size, efficiencies and repeatability would be better placed in the methods section in my opinion, since it is not results per se.

RESPONSE: We disagree here - it is the outcome of the molecular methods. We tried moving it but this disrupted the flow of the manuscript se have left it as it was.

Lines 267-271: You mention that "longitudinal telomere length increased with log age...", and cite fig 2C, 2D to support your statement. As far as I understand fig 2C, it seems that telomere length decrease with age before more or less 2 years of age, and then is not changing overall with age later in life. Yet, such figure does not fit with the estimate you give in the text, so the presentation of your results is quite confusing here.. I do not see how your results are supporting this particular statement. In my opinion, you should probably state that the probability of detecting a within-individual telomere shortening is decreasing with age // the probability of detecting telomere elongation is increasing with age.

RESPONSE: We agree that this was confusing, and have simplified along the suggested lines - now lines XX:

"We found that the likelihood of telomere lengthening increased with log age... Increases in telomere length were most likely to be observed shortly after the juvenile period, at around four years of age, and later in life (although sample sizes for older birds are much smaller; Fig. 2C,D)"

Line 302: the sentence is too vague, especially "particular life stage".

RESPONSE: We have now simplified the sentence to make it clearer, lines XX:

"Despite an overall pattern of telomere shortening with age in the Seychelles warbler, we found evidence of within-individual increases in telomere length, and that the extent of these increases could not be explained solely by qPCR measurement error."

Line 321: same comment as above "specific periods".

RESPONSE: We are not completely sure what is unclear about this, but have tried rephrasing. Now lines XX:

"Increases in telomere length were not consistent over individual lifespans, but occurred in bouts, against a backdrop of overall lifelong telomere shortening. "

Lines 337-338: Such change in blood cell types would be quite convincing for mammals since "mammalian people" are measuring telomeres from different kind of WBCs and also immature nucleated RBCs.. However, in birds more than 99% of the cells in the blood are RBCs, therefore a change in blood cell types is quite unlikely. However, cellular turnover of RBCs might change with age/life stage, and you can expect that RBCs that have been in the blood stream for a longer time have shorter telomeres due to oxidative stress, though this remain to be tested properly.. Therefore you should probably mention the "cellular composition of the blood (e.g. age of RBCs, proportion of WBCs)" instead of "composition of cell types".

RESPONSE: This is a good point and we have made the suggested edit.

Lines 348-349: you might cite Becker 2015 and Stier 2014 here too.

RESPONSE: These bave been added in response to an earlier comment.

Lines 350-54: Please reformulate according to major comment 6 above.

RESPONSE: We have now removed this section.

Line 356: Atema 2015 is not in the wild

RESPONSE: We have removed this citation as part of a previous comment.

Lines 367-369: In my opinion, Heidinger et al. 2012 does not fit here (they do not look at differences in early-life conditions), but you might cite Reichert 2015, Stier 2014, 2015a, 2015b, Meillère 2016.

RESPONSE: We have removed the Heidinger reference, leaving two citations in - we felt this was sufficient.

Lines 378-379: It is unclear here that you speak about a comparative study of the sex effect on TL.

RESPONSE: We rephrased accordingly, lines XX:

> *"Also worth noting is that the effect of telomere length on survival is strongest in male Seychelles warblers (Barrett et al. 2013), although comparative research suggests that the nature of the relationship between sex, telomeres and survival is not yet clear (Barrett and Richardson 2011)."*

Reviewer 2

I congratulate the authors on this fantastic paper. It's very rare for me to read a manuscript and not be able to come up with a single meaningful criticism, but I really can't fault this effort. The authors tackle important emerging questions about telomere dynamics and the importance of telomere length as a biomarker in ecology using a fantastic long term data set. Importantly, and unlike many recent studies of telomere length in wild animals, this study is based on a high quality longitudinal data set spanning a long time series. I particularly like the great care the authors have taken to address the question of within-individual lengthening of telomere length, applying a variety of published methods to test whether observed lengthening is greater than that expected due to measurement error. These results, combined with the thorough analyses of ageing patterns and cohort effects, represent an important addition to our current understanding of the complexity of telomere dynamics in the wild. They also show interesting associations between telomere length, environmental quality, sex and size which offer important insights into the ecological and physiological factors responsible for telomere dynamics. All in all, I really enjoyed reading this paper and have no meaningful suggestions for improvements. I expect this excellent manuscript will make a very strong and widely cited addition to the emerging literature on telomere dynamics in wild populations.

RESPONSE: We thank Dr Nussey for his extremely kind words. It is not often that one receives such a positive review!

Reviewer 3

Dear authors, You have an impressive dataset of longitudinal data and your work on telomere dynamics can be an important contribution to the current knowledge. However, I have some major concerns regarding your manuscript.

Your introduction could be improved, first of all, by citing correctly. Also, there is some more recent literature on the topic, which you could cite. See more in comments on the introduction below.

RESPONSE:

Further, I have some methodological questions. I would like to see more information on how samples were distributed over qPCR plates. Were samples of the same individual on the same plate? If not, how did you take plate ID into account in your statistical analyses? How exactly did you relate territory quality to RTL and delta RTL change? Only for birth year? Reading your abstract and aims in the introduction, I expected more.

RESPONSE:

In the discussion, you speculate on effects of environmental conditions on telomere dynamics. Is there a relationship with telomere lengthening between years and insect abundance/territory quality in the year of lengthening? I guess you could test this with the data you have? This could add to your discussion on potential telomere lengthening, which remained rather methodological. In my opinion, the ecological context in the discussion could be improved.

RESPONSE:

L 53: In Monaghan & Haussmann 2006, data for 4 species are presented whereof two show telomere shortening. I therefore don't consider this the best citation for your statement. Also, given that by now there are plenty of others that show this relationship - including other Haussmann or Monaghan papers.

RESPONSE:

L 54: The statement is correct but reference Haussmann, Vleck & Nisbet 2003 is wrong! In this paper the result of a cross-sectional analysis is presented and discussed that one can classify birds into age classes according to their telomere length. Please cite appropriate papers.

RESPONSE:

L 56: If you use the reference von Zglinicki 2002, please state that this is based on cell culture research. Or site a convincing reference where it has been shown in vivo. As far as I know the relation between oxidative stress and telomere shortening in vivo is less clear.

RESPONSE:

L 65: Please either cite a meta-analysis or some more like Barrett et al. There are quite some studies with original data showing a relationship between telomere length and survival. Given all the literature available, I do not consider to cite Simons 2015 a good choice here.

RESPONSE:

L 65-66: Here you state there is little evidence of a direct relationship. Only a few lines above, in sentence line 59-61, you write about a potential causal relationship. Please consider to rewrite this, so that it is in one paragraph.

RESPONSE:

L 80: Njajour 2007: Where is this reference?

RESPONSE:

L136: You describe how you assess territory quality etc. But how do you relate it to RTL or delta RTL in your statistical analyses? Do you only use territory quality in the hatching year? You could relate territory quality between RTL measurements to telomere dynamics. Is the variation of territory quality within-individually over time smaller than among-individually? Is the difference between years much larger than between territories?

RESPONSE:

L 165: "time point t + 1" Is the time span between repeated samples always the same? If not, please add a sentence to make it clear. Are samples taken in the same month every year?

RESPONSE:

Some more information on the sampled individuals your study is based on would be helpful and could be added to the supplement. E.g. cohort sizes.

RESPONSE:

L 168: How do you explain a non-normal distribution of telomere lengths? Data selection or methodological effect? How large is the variation of samples among cohorts?

RESPONSE:

L 180: Van de Pol & Wright 2009

RESPONSE:

L 205: first and least?

RESPONSE:

L 234: Please provide intra-plate repeatabilities as well. Do you have repeats that are not directly next to each other on the plate?

RESPONSE:

Given an inter-plate repeatability of 0.68, how do you take plate ID into account in your statistical analyses? Random effect plate ID?

RESPONSE:

For the samples analysed repeatedly, which measurement did you use? Or average?

RESPONSE:

L 238-239: If you would add a translation into base pairs lost per year, it would become easier to interpret your results. As the relationship is not linear give examples for a few ages.

RESPONSE:

L 247: Please compare cross-sectional and longitudinal telomere shortening and discuss it.

RESPONSE:

L 252: It is always the same time span between sampling, right? Could the low repeatability/correlation be partly due to inter-plate differences? Did you run all samples of the same individuals on the same plates or partly on different plates?

RESPONSE:

L 307: Again the citation Haussmann et al. is not correct here.

RESPONSE:

L 340: Does telomere lengthening or shortening relate to environmental conditions in your study?

RESPONSE:

Generally, the discussion of this part of potential causes of telomere dynamics remains very methodological. The ecological context could be improved by adding own results as mentioned above or more citations.

RESPONSE:

L 389: Sentence incomplete

RESPONSE:
