Peer Community In Evolutionary Biology

A genomic perspective is needed for the re-evaluation of species boundaries, evolutionary trajectories and conservation strategies for the Galápagos giant tortoises

Michael C. Fontaine *based on peer reviews by 4 anonymous reviewers*

Etienne Loire, Nicolas Galtier (2017) Lacking conservation genomics in the giant Galápagos tortoise. Missing preprint_server, ver. Missing article_version, peer-reviewed and recommended by Peer Community in Evolutionary Biology. 10.1101/101980

Submitted: 21 January 2017, Recommended: 26 September 2017

Cite this recommendation as:
Fontaine, M. (2017) A genomic perspective is needed for the re-evaluation of species boundaries, evolutionary trajectories and conservation strategies for the Galápagos giant tortoises. *Peer Community in Evolutionary Biology*, 100031. 10.24072/pci.evolbiol.100031

Published: 26 September 2017

Genome-wide data obtained from even a small number of individuals can provide unprecedented levels of detail about the evolutionary history of populations and species [1], determinants of genetic diversity [2], species boundaries and the process of speciation itself [3]. Loire and Galtier [4] present a clear example, using the emblematic Galápagos giant tortoise (*Chelonoidis nigra*), of how multi-species comparative population genomic approaches can provide valuable insights about population structure and species delimitation even when sample sizes are limited but the number of loci is large and distributed across the genome. Galápagos giant tortoises are endemic to the Galápagos Islands and are currently recognized as an endangered, multi-species complex including both extant and extinct taxa. Taxonomic definitions are based on morphology, geographic isolation and population genetic evidence based on short DNA sequences of the mitochondrial genome (mtDNA) and/or a dozen or so nuclear microsatellite loci [5-8]. The species complex enjoys maximal protection. Population recoveries have been quite successful and spectacular conservation programs based on mitochondrial genes and microsatellites are ongoing. This includes for example individual translocations, breeding program, “hybrid” sterilization or removal, and resurrection of extinct lineages). In 2013, Loire et al. [9] provided the first population genomic analyses based on genome scale data (~1000 coding loci derived from blood-transcriptomes) from five individuals, encompassing three putative “species”: *Chelonnoidis

becki*, *C. porteri* and *C. vandenburghi*. Their results raised doubts about the validity/accuracy of the currently accepted designations of "genetic distinctiveness". However, the implications for conservation and management have remained unnoticed. In 2017, Loire and Galtier [4] have re-appraised this issue using an original multi-species comparative population genomic analysis of their previous data set [9]. Based on a comparison of 53 animal species, they show that both the level of genome-wide neutral diversity (πS) and level of population structure estimated using the inbreeding coefficient (F) are much lower than would be expected from a sample covering multiple species. The observed values are more comparable to those typically reported at the "among population" level within a single species such as human (*Homo sapiens*). The authors go to great length to assess the sensitivity of their method to detect population structure (or lack thereof) and show that their results are robust to potential issues, such as contamination and sequencing errors that can occur with Next Generation Sequencing techniques; and biases related to the small sample size and sub-sampling of individuals. They conclude that published mtDNA and microsatellite-based assessment of population structure and species designations may be biased towards over-splitting. This manuscript is a very good read as it shows the potential of the now relatively affordable genome-wide data for helping to both resolve and clarify population and species boundaries, illuminate demographic trends, evolutionary trajectories of isolated groups, patterns of connectivity among them, and test for evidence of local adaptation and even reproductive isolation. The comprehensive information provided by genome-wide data can critically inform and assist the development of the best strategies to preserve endangered populations and species. Loire and Galtier [4] make a strong case for applying genomic data to the Galápagos giant tortoises, which is likely to redirect conservation efforts more effectively and at lower cost. The case of the Galápagos giant tortoises is certainly a very emblematic example, which will find an echo in many other endangered species conservation programs.

References:

[1] Li H and Durbin R. 2011. Inference of human population history from individual whole-genome sequences. Nature, 475: 493–496. doi: [10.1038/nature10231](https://doi.org/10.1038/nature10231)

[2] Romiguier J, Gayral P, Ballenghien M, Bernard A, Cahais V, Chenuil A, Chiari Y, Dernat R, Duret L, Faivre N, Loire E, Lourenco JM, Nabholz B, Roux C, Tsagkogeorga G, Weber AA-T, Weinert LA, Belkhir K, Bierne N, Glémin S and Galtier N. 2014. Comparative population genomics in animals uncovers the determinants of genetic diversity. Nature, 515: 261–263. doi: [10.1038/nature13685](https://doi.org/10.1038/nature13685)

[3] Roux C, Fraïsse C, Romiguier J, Anciaux Y, Galtier N and Bierne N. 2016. Shedding light on the grey zone of speciation along a continuum of genomic divergence. PLoS Biology, 14: e2000234. doi: [10.1371/journal.pbio.2000234](https://doi.org/10.1371/journal.pbio.2000234)

[4] Loire E and Galtier N. 2017. Lacking conservation genomics in the giant Galápagos tortoise. bioRxiv 101980, ver. 4 of September 26, 2017. doi: [10.1101/101980](https://doi.org/10.1101/101980)

[5] Beheregaray LB, Ciofi C, Caccone A, Gibbs JP and Powell JR. 2003. Genetic divergence, phylogeography and conservation units of giant tortoises from Santa Cruz and Pinzón, Galápagos Islands. Conservation Genetics, 4: 31–46. doi: [10.1023/A:1021864214375](https://doi.org/10.1023/A:1021864214375)

[6] Ciofi C, Milinkovitch MC, Gibbs JP, Caccone A and Powell JR. 2002. Microsatellite analysis of genetic divergence among populations of giant Galápagos tortoises. Molecular Ecology, 11: 2265–2283. doi: [10.1046/j.1365-294X.2002.01617.x](https://doi.org/10.1046/j.1365-294X.2002.01617.x)

[7] Garrick RC, Kajdacsi B, Russello MA, Benavides E, Hyseni C, Gibbs JP, Tapia W and Caccone A. 2015. Naturally rare versus newly rare: demographic inferences on two timescales inform conservation of

Galápagos giant tortoises. Ecology and Evolution, 5: 676–694. doi: [10.1002/ece3.1388](https://doi.org/10.1002/ece3.1388)

[8] Poulakakis N, Edwards DL, Chiari Y, Garrick RC, Russello MA, Benavides E, Watkins-Colwell GJ, Glaberman S, Tapia W, Gibbs JP, Cayot LJ and Caccone A. 2015. Description of a new Galápagos giant tortoise species (Chelonoidis; Testudines: Testudinidae) from Cerro Fatal on Santa Cruz island. PLoS ONE, 10: e0138779. doi: [10.1371/journal.pone.0138779](https://doi.org/10.1371/journal.pone.0138779)

[9] Loire E, Chiari Y, Bernard A, Cahais V, Romiguier J, Nabholz B, Lourenço JM and Galtier N. 2013. Population genomics of the endangered giant Galápagos tortoise. Genome Biology, 14: R136. doi: [10.1186/gb-2013-14-12-r136](https://doi.org/10.1186/gb-2013-14-12-r136)

Reviews

Evaluation round #2

DOI or URL of the preprint: **10.1101/101980**
Version of the preprint: 2

Authors' reply, 04 September 2017

Download author's reply

Decision by Michael C. Fontaine, posted 22 July 2017

Modifications required

Dear Dr. Loire and Galtier,

Thank you for your patience. Please accept again our sincerest apologies for the unusual long delay in delivering the reviews of your manuscript. We have received three additional thoughtful reviews of your preprint manuscript entitled "Preserving microsatellites? Conservation genetics of the giant Galápagos tortoise." All three reviewers agree that this new version of the manuscript has improved compared to the original version and hold the potential to make a valuable contribution to the field and practices in conservation genetics. I do agree with them. However, the three reviewers still raise important issues that need to be consider carefully before I can make a recommendation of this manuscript. Reviewer #1 especially has provided a thorough and very critic assessment of the analyses of the transcriptomic data provided in this MS and raise serious issues about the limitation of the sampling of individual tortoises as a demonstration that genomic data contradict previous results obtained with more classical markers. I do agree with the reviewer assessment, and I suggest that the authors tune down their comparison of the transcriptomic data with previous studies. As pointed by the three reviewers a more measured discussion would have merit and could make a significant contribution if it would cover the topics related to taxonomic and conservation units (e.g. population, (sub-)species, management units) since this is at the heart of the argumentation, why genomic data are now required to define those units and assess their evolutionary trajectories in a more insightful and reliable way. Reviewer #1 implies that some genomic work is currently ongoing on the Galapagos giant tortoise. I suggest the authors to look into that, since the management plan of the Galapagos tortoises states that NGS data are underway. I recommend that the authors consider carefully the comments and suggestions of the three reviewers, and address the issues raised in a satisfactory way before I can make a recommendation. All the best

Michael C Fontaine

Reviewed by anonymous reviewer 1, 16 July 2017

I thank the authors for removing Fit as a proxy measure of species divergence and acknowledgement that the negative Fit in the previous version was "troublesome". The replacement with F does not improve the figure. The attempt to refocus the manuscript on the lack of loci currently used in management plans, rather than insist that microsatellites are not useful markers is a major improvement. However, I still have reservations about the usefulness of this letter.

1) Samples of mixed ancestry:

The reviewer considers an interesting possibility, which is that there would actually be substantial population structure in C. nigra, but that our sample would not reflect it because it would by chance consist in individuals of "mixed ancestry". We, of course, cannot formally reject this hypothesis.

The data on microsatellites and MtDNA does not show "substantial" population structure, but some. At least 2 of the 5 samples are of mixed ancestry see Russell et al 2010 & Russell et al 2007 and tables provided below. The samples are not a random selection of populations, but a selection of zoo animals who's origin was initially unknown. In Russell et al 2007, only 12.2% of captive individuals were determined not to have MtDNA and micostat data that were consistent- which are 3 of the 4 transcriptome samples. A larger dataset in 2010 reduced the q values of the same samples but now 2/4 were consistent, and a larger analysis of multiple samples Garrick et al 2014 estimated that q values below 0.8 were considered to be F1 or backcrosses.

ZUZ1= Santa Cruz, with a q below 0.768 but now consistently assigned across datasets

ZUZ10= PBL with a q of 0.330, indicating hybrid origin but consistently assigned across studies.

ZUZ20=PBL or AGO with a q of 0.6 also indicating it is of hybrid origin, has been assigned to two different populations.

ZUZ30= assigned to three different populations, VA, VR, and VD with a q of 0.470 indicating it is of hybrid origin.

Further to that, ZUZ20 and ZUZ30 potentially from different populations were mixed up in your lab (Loire et al 2013). The fifth sample is from C. becki, a population elegantly shown to be a lineage undergoing "natural" fusion (Garrick et al 2014), but this sample is without microsatellite data to compare it to the published datasets. You then go on to discuss specifically the GAE05F and the mixed up sample GA05H, I think this overstates the quality and certainty of your data.

The potential for these samples to be hybrids is not mentioned in Loire et al 2013. This is important as it limits the usefulness of your already small sample size with regard to discussing species delimitation, it may also be why the other groups, with greater access to samples, have not cited your paper. With the addition of Ballenghien et al 2017 the authors now report 80% of their own samples have been contaminated by the sequencing centers, I doubt this will add confidence to the tortoise community to utilize these data.

Many other studies have demonstrated low genetic diversity is in the Galapagos tortoises, your data, with more (>1000) transcriptome loci - show the same low diversity. These data are already published (Loire et al 2013, Ballenghien et al 2017), not novel, nor show conclusively more than can be gained Garrick et al 2014.

It is well established that the tortoises are a recently diverging species group, and from Garrick et al (2014) view under pd or biological species concepts would be considered full species. The key here, is lack of loci and management measures based on these loci. From Russell's et al's own work demonstrating a change of q with larger samples indicates there is not enough loci at present, supporting the authors view. I fully concur with the authors, that sterilization, seems extreme, as does the desire to bring back lineages from the dead, a movement to NGS is certainly required for management. I do not think this letter is timely nor expresses a new nuance that can be gained specifically from the author's limited transcriptome data.

2) The utility of summary statistics to compare diversity across species

Still, please note that the total amount of nuclear genetic diversity we detect in this sample is extremely low, compared to typical multi-species samples.

Using a relative summary statistics to compare genetic diversity across species with different historical processes is flawed (Charlesworth 1998). Fst -like statistics are strongly influenced by within population

divergence, markers that sample different parts of the genome will therefore provide different levels of differentiation. Relative measures of differentiation are not appropriate if the goal is to compare species with different population history (Charlesworth 1997; 1998). Elevated levels of F can be due to selection or inbreeding rather than reduced gene-flow (Charlesworth 1998; Cruickshank & Hahn 2014). Genetic diversity can be low due to substantial population decline (e.g. bottlenecks, as known to be the case in the tortoises), small founder effects, and or catastrophic events (e.g. volcanic eruptions). The negative Fst values obtained for the Galapagos tortoises even after "contamination correction" (Ballenghien et al 2017) remains a concern for these data. Using an absolute measure of diversity would be better, as the authors themselves are aware.

The reviewer's hypothesis would imply that (i) the distinct C. nigra entities have diverged by no more than ~0.1% genome-wide, and (ii) hybridization is so common that hybrids dominate in our random sample.

Previous studies confirm at least 2/5 samples are hybrids. Your lab then mixed 2 of the samples, so the actual number of samples that you have that can be traced back to an individual is n=3. The samples are not random but animals that were zoo residents from unknown origin. You have not sampled genome wide but transcriptomes in at least 2 hybrids. In 2013 you state ZUZ01 was from Rotterdam zoo, but this animal was transferred to London Zoo in 2010. It is important there is clear understanding on what the samples represent. If the authors do not concur with the concept of "purebred" and "mixed ancestors" at all in this system, their points would be strengthened by data from samples that the tortoise population geneticist would consider purebred taxonomic units.

Again, we perfectly agree that this sample of size 5 is not sufficient, far from it. It is just a demonstration that more data are needed before making drastic management decisions, since the first genome-wide data set gathered in this species contradicts the hypotheses on which management policy gas been based so far. We modified the text in several places to account for this comment, explicitly stating that our sample is insufficient to provide a definitive answer on population structure in C. nigra (l 138-139).

Yes management plans should be based on more genetic loci, but no, you have not demonstrated that your approach from transcriptome data is an improvement. You have not demonstrated that the management policy is in contradiction with the transcriptome data. You show low genetic diversity (as seen before) and the management policy is designed to maintain the diversity that exists.

Management plan:

There is certainly improvement by actually citing the management plan and papers on which the authors are criticizing.

The paper would be improved with more written about species, population and conservation units. However, the authors state that this is out-with the scope of their letter, yet I disagree, as it is at the heart of the problem. I question whether this is timely (data published 2013) and I am not convinced that there is enough new data or analyses to warrant such a letter at present.

So much of this manuscript is about what constitutes a suitable set of markers to delimit species or management units, but the authors have not addressed the taxonomic history of this iconic group or whether that matters (surely all populations are a priority for conservation). I concur with the authors that prudent use of data would be a useful contribution to Galapagos work, and in that, I include prudent use of the author's own limited data.

I do not see there is merit in "putting pressure on groups" that have large datasets already, after they have already stated that they are doing NGS data on these samples. I can see the authors feel others should have cited their dataset, but mixing samples in the laboratory, having at least 2 that are have mixed ancestry (which they seem to be unaware of), and the problem they have experienced of contamination at sequencing centers, limits the usefulness of the data. The authors clearly state they are not in a position to do more, I suggest they wait until more data is available or they get more samples.

Minor points:

Lines 66-70 should be modified to make this clearer. The species were described morphologically (Van Denburgh1914) and later confirmed by genetics.

Line 76- who died, not “which died”.

Garrick et al 2012 is not in the references.

Line 80: Some problems, however, still hamper the task of defining C. nigra conservation units. For instance, mitochondrial DNA and microsatellite data do not fully agree (Poulakakis et al. 2012).

I would replace “conservation units” with taxonomic units to strengthen your argument.

Line 124 ” from 4 to 11 individuals”, but says 5 or more individuals in figure legend.

Line 165-174: Should reference Garrick et al 2014

Line 175: several papers exist already on ancient vs modern gene flow. You should rewrite this last paragraph, it is not as well written as some of the comments to the referees.

I reiterate from my previous review, a more measured discussion would have merit, highlighting the issues raised, however even if they were to do so, I cannot recommend this letter based on relative summary statistics of limited data, the transcriptome data has been published before, the management plan of the Galapagos tortoises states they are doing NGS, so I don’t see what purpose this letter would serve.

Samples used in Loire et al 2013

MtDNA and microsatellites are not consistent on population/species assignment for at least 2/5 samples. ZUZ10 and 20 mixed by authors own laboratory. If GA05H is assigned to PBL or AOG in Russell et al 2010 this may explain why it was more different to GA05A than other individuals assigned to PBL (Line104-105). Lack of microsat loci, is a problem, but is so is the lack of transcriptomes of individuals with unequivocal population assignment (i.e. with a higher q value than 0.8 for a population assignment the amount Garrick et al consider to be ‘purebreds’).

See attached PDF file to see the tables

Download the review

Reviewed by anonymous reviewer 3, 21 July 2017

Review of Loire & Galtier: Preserving Microsatellites? Conservation genetics of the giant Galapagos tortoise

In this letter Loire & Galtier discuss an interesting and highly relevant topic of using genetic/genomics to understand species’ evolutionary biology, particularly in conservation biology contexts. The authors present a re-analysis of recently published transcriptome data, the results of which are discordant with previous studies using microsatellites (upon which the current conservation management strategy is based) and call for additional genome-wide studies to inform conservation management of this species.

Having read the revised manuscript and the previous thorough reviews and responses from the authors (as well as revisiting Loire et al. 2013 to fully understand the context of the new results), overall I feel that the authors provided thoughtful responses and concordant revisions to the manuscript. The majority of the initial questions and points of clarification that arose in my mind whilst reading the revised version were discussed in the previous reviews. Thus, I will limit my comments to a handful of suggestions-most of which were brought up in the initial reviews but I feel deserve further clarifications in the manuscript in order for readers to understand the arguments, context, and key take home messages. With these changes, I can support the recommendation of the letter by PCI Evol Biol and hope that this can encourage further discourse of the application of genomics in conservation biology.

- Despite the authors’ thoughtful responses to comments stating that their key message is that further examination/valuation of giant Galapagos tortoise genomics is needed, as well as their argument is not against microsatellites as markers but rather the limited # of markers used in previous studies, both of these points are absent from the abstract.

-Additionally, I agree with the previous reviewer’s comment that the title (which seems to have remained unchanged) is misleading about the issue being with the # of markers vs the type. I wouldn’t necessarily withhold recommending the letter if the authors feel strongly about the title, but personally I think the authors could come up with a better fit to accurately encapsulate the points and fuel a healthy discourse.

-I thoroughly agree with the previous reviewers' suggestions regarding adding more background information on the conservation context/history, and think the authors did a very nice job of providing that in the revised version. I also think the revisions to the comparative analyses make the findings and relevance clearer.

-I agree that species concepts and taxonomy is a large topic that is beyond this scope of this letter, however I do think it is important to be clear about what species concepts have been used/are suggested in this context. Marine mammal science just published a special issue devoted to this topic of delineating species and subspecies for conservation http://onlinelibrary.wiley.com/doi/10.1111/mms.2017.33.issue-S1/issuetoc The authors may find some content and visuals helpful, in particular : Taylor et al. Guidelines and quantitative standards to improve consistency in cetacean subspecies and species delimitation relying on molecular genetic data

-There are several places where I think the authors may want to provide further in text clarification for their methods/arguments to communicate with non-geneticists (i.e., particularly conservation biologists/managers). In particular, the issue of using transcriptome/coding data, which is more constrained, as well as tissue type, is raised and responded to in the reviews, and the authors include citations to other papers where this is discussed to support their methods. I understand the authors don't want to spend a large amount of background when it is better explained/validated elsewhere (and thus this approach is common and acceptable in most situations). However, here the primary goal of this letter is to get the (likely skeptical) reader to consider their argument and not dismiss the results as in err when they are discordant with the previous studies. In particular, to get those involved in conservation programs to pay attention and consider valuing future genome-wide studies, etc. Many conservation biologists, and many conservation geneticists for that matter, are comfortable with markers like microsats, but are not well versed in different NGS approaches (for context, I work as a conservation geneticist using NGS approaches in a government wildlife management agency). Most have learned about using these assumed neutral markers for understanding evol processes, etc., and may immediately dismiss using coding regions like transcriptomes if they don't understand it. Similarly, the low sample # vs. high loci justification...despite multiple studies demonstrating the utility and validity of this (which the authors explain very well in response to the reviews and include citations in text), this remains not well understood and could lead to dismissal of their results from readers who think we cannot learn anything from 5 samples. In sum, I don't suggest that the authors spend a lot of time belaboring the points that this approach is scientifically sound (and I think they do a good job of being transparent with the limitations of the data in the revised version), but I do think a bit more in text explanation would ensure that the audience they seek to reach understands-if the authors really want to have their work impact the conservation of this species, conservation managers need to understand it.

-Finally, the point discussed regarding the disagreement/uncertainty about whether hybridization/mixing occurred naturally prior to human intervention seems key to the conservation context, and the application of genome-wide analyses to better reconstruct the history (and that we really cannot do this any other way), I think is under-emphasized and would help support the authors' argument for future genome studies.

Reviewed by anonymous reviewer 4, 13 July 2017

The authors offer a critique on current Galapagos tortoise conservation practices based upon a reported discrepancy between transcriptome data and mtDNA/microsatellite data. The transcriptome data, generated previously by the authors, suggests the samples used in that study were drawn from a panmictic population (Loire et al. 2013), whereas studies mtDNA (Caccone et al. 2002) and microsatellite data (Ciofi et al. 2002) conclude there is population structure in C. nigra. In the present study, the authors do a comparative analysis of the inbreeding coefficient F versus the log-transformed non-synonymous diversity s. The authors suggest the low value of F estimated for C. nigra indicates low genetic diversity and population structuring compared to other species, questioning the presence of genetically differentiated clusters in C. nigra and, subtly, the need to define subspecies in C. nigra.

I think the previous reviewers have already provided a lot of valuable comments that I will not repeat here,

so I will focus on sharing my thoughts on the structure and set-up of the paper. In my opinion, the paper has the potential to be an interesting opinion, or even a review, on current conservation practices in Galapagos tortoises and how genomic data should be integrated into it to change it for the better. I therefore really encourage the authors to follow up on this. However, I'm struggling with the present structure of the story. Is the problem the conservation units, the traditional markers, or the "extreme" translocation experiments done? I think there is a choice to be made here by the authors.

To me the issue here does not seem to be the traditional markers themselves or the number of loci. Generally, the problem with traditional markers/low numbers of loci is a lack of power in detecting population structure when gene flow is high or populations are not in migration-drift equilibrium. Marine species seem to suffer from this in particular. Here increasing the number of markers (e.g. SNPs) may enhance more resolution. Of course, the biological relevance of "low but significant" estimates of genetic differentiation comes to mind following this. Judging from the work of Caccone et al. (2002) and Ciofi et al. (2002) it seems that they do provide convincing evidence for population structure. Even when there is only weak evidence for population structure, it can be better to employ precautionary approach and subdivide a population more than seems necessary to avoid local extinction of potentially undetected population segments. In addition, I agree with the previous reviewers that microsatellites, and mtDNA as well, have valuable traits themselves that make them useful for conservation, e.g. the fast-evolving nature of microsatellites or the use of mtDNA for demonstrating male vs. female-mediated dispersal.

From the above section, it also follows that the conservation units don't seem to be the real issue. Delineating conservation units is a challenging thing, and requires careful evaluating of all available scientific data for a species. Ultimately, it is a management decision.

Judging from the answers of the authors to the previous reviewers, I carefully conclude that the problem seems to be that we can now do genomics at an affordable cost, and it's not being done for the Galapagos tortoise despite its conservation status and current costly management practices heavily based upon genetics. I feel the story can be focussed and streamlined more to fit with this aim. For example, the focus can be narrowed by emphasizing possible genomic consequences of the current management practices (e.g. really focus on inbreeding, potential losses of adaptive potential?) in a "what if we're wrong?" manner using the precautionary principle. I do believe the authors have the potential to write a really important opinion here, but the scope needs to be narrow enough to make the point of the authors clear.

As the authors mention in their reply to the first reviewers, they are aware that a full discussion of defining species, sub-species, conservation units and how genetic data fits in here is beyond the scope of the paper. I fully agree with this. However, in the current format that discussion (different subspecies of C. nigra) seems to be at the core of the paper. The authors provide one analysis, the comparative analysis demonstrating a low inbreeding coefficient in Galapagos tortoises compared to other species, and conclude it is unlikely that multiple differentiated gene pools are unlikely and question the need to define sub-species. This makes the story convoluted, because this ties into the classic discussion: what is a species, sub-species, population?

Given the conservation issue of Galapagos tortoises and the fact that genomic resources are available at an affordable cost, I fully agree with the authors and encourage them to follow up on this paper. However, I would recommend to carefully look at the story, evaluate which problem is most important to the authors and restructure the story accordingly. As mentioned before, an opinion paper would be one possibility. Or if the authors are up to it, a review of Galapagos conservation genetics and the need for genomic data.

Evaluation round #1

DOI or URL of the preprint: https://doi.org/10.1101/101980
Version of the preprint: 1

Authors' reply, 06 May 2017

Dear Dr. Loire and Galtier,

We have received two thoughtful reviews of your pre-print manuscript entitled "Preserving microsatellites? Conservation genetics of the giant Galápagos tortoise." Both reviewers agree that this is a valuable and interesting letter which raise important points and issues regarding the ongoing conservation practices of the Galapagos Giant tortoises based in insufficient genetic data. However, both reviewers also raised important issues and reserves regarding the current version. Both reviewers felt that the authors should do a better job at explaining the context, the actual issues, and limitations of previous studies. Furthermore the results provided in this current version as an attempt to illustrate the poor representativeness of a dozen of loci to delineate the genetic structure and taxonomic delimitation are based on another data set which also raise a lot of questions. The two reviewers have made extensive recommendations which could greatly improve the current version and I encourage the authors to address them carefully in revised version in order to be consider for a recommendation by PCI Evol Biol.

Thanks for this. Please find below our detailed, point-by-point response to the reviewers' comments.

Very briefly:

- we agree with many of the suggestions regarding the presentation of our work, which we have substantially amended; in particular, we clarify that we are not criticizing the usage of microsatellite loci, but rather of an insufficient number of loci;
- we give more details about ancient and recent conservation initiatives in *C. nigra*, the influence of published genetic data analyses, and our suggestions for improvement;
- we discuss in more detail the controversy between various sources of data, existing biases in published analyses, and the potential and limitations of our approach regarding substructure detection and species delineation; we now refer to the Roux et al (2016 PLoS Biol) study, which was yet unpublished when we first submitted this manuscript;
- we have updated estimates of piS and F thanks to a recent improvement of our SNP-calling algorithm (Ballenghien et al 2017 BMC Biol) accounting for cross-contamination and similar sources of noise in read counts.

Reviewed by anonymous reviewer, 2017-04-13

This paper serves to highlight the previous transcriptome work of the authors (Lore et al 2013) demonstrating that there is low genetic diversity in Galapagos tortoises and make comment on the current conservation management based primarily microsatellite data. The authors have re-used transcriptome data from 5 wild caught individual zoo animals (Lore et al 2013) who's collection data was not known but genetically assigned to 3 species/subspecies (Russello et al. 2010). They reconfirm they have found very low level of genetic diversity and the taxonomy of the group needs to be addressed. They compare the diversity within Galapagos tortoises to a range of other species diversity and finally compare the difference to a pair-wise comparison of transcriptome data from across the tree of life. They question the usefulness of the current management plan for Galapagos turtles based on microsatellite diversity.

I found this article thought provoking based on an interesting group, however found the discussion of the taxonomy, current management plan and evidence of natural hybridization in the system requires more work, the limited data and the conclusions drawn overreach the evidence as currently presented.

Thanks for this comment. Please note that our "conclusion" is mainly a message of prudence and a call for more data to be generated, given the conflicts that exist between currently available data sets.

Comments: There is a general lack of background regarding the taxonomy of the Galapagos turtles. The recognition of separate species or interbreeding subspecies is at the heart of the authors argument.

More therefore could be written regarding the taxonomic status as currently recognised Whether the taxa are species or subspecies does not affect their need for conservation. However, their ability to hybridize and whether or not that should or shouldn't be prevented when populations are so small is what is the issue is here. The management program I am assuming from the authors summary removes "hybrids" from the breeding pool; with the assumption, that hybridization is due to anthropogenic interference rather than natural causes. (This needs more references to demonstrate this is what is happening).

However, lineage fusion is also being argued for the diversity found in C. becki (Garrick et al 2014). C. becki one of the 3 species chosen in their sample, seems to be well established as a taxon that has mixed ancestry (e.g. (Poulakakis et al. 2012)). 50% of the sampled individuals have mixed ancestry (Garrick et al. 2012). Only 29% of the PBL population were found to be pure-bred individuals (Garrick et al. 2014), but crucially in support of the authors argument lineage fusion is thought to predate human mediated translocations. More should be written about this species to support the authors' view point.

Good point. We added a sentence highlighting the commonness of hybrids in the field (l 86).

The heart of this paper is based on the provenance of the 5 samples and the robustness of their assignment to different species/subspecies tortoise. I recommend these data be presented in a table in the supplementaries. There seems to be conflicting results for species assignment using microsatellite data for 3 of the 5 samples between studies (Russello et al. 2007; Russello et al 2010). Russello et al. (2007) states the differences between MtDNA and microsat assignment for these samples is due to them having mixed ancestry, the different methods also assign them differently. After adding data from a new population, he reports different assignment again for the same samples, with lower q (Russello et al. 2010). Could the authors add a sentence or two, to assure provenance. ZUZ10/ZUZ20 are reported as mixed up when the authors sequenced their transcriptome and assigned them GA05H and GA05G in the original transcriptome paper (Lore et al 2013). It is therefore not clear which microsatellite data belongs to which transcriptome. The same samples that were assigned to PBL in 2007 but were assigned to different populations in 2010, both were thought to mixed ancestry. All 4 samples had q-values below 0.79, there seems to be no data on the 5th sample.

Can the authors comment on these samples population assignment. Is it unclear? 1. Can you provide evidence that the samples are not from mixed ancestry? The micosats suggest all 4 samples are. And if they are, comment on how that would affect your conclusions about panmixa? 2. The fifth sample added to the data set in 2013, has a MtDNA assignment but I could not find a microsatellite profile for comparison. 3. Does mixed ancestry explain the genetic differences found for GAO5H mentioned in the text? 4. Do the authors have evidence that the microsatellites are not robust at identifying populations with confidence? It would strengthen the authors point.

The reviewer considers an interesting possibility, which is that there would actually be substantial population structure in *C. nigra*, but that our sample would not reflect it because it would by chance consist in individuals of "mixed ancestry". We, of course, cannot formally reject this hypothesis. Still, please note that the total amount of nuclear genetic diversity we detect in this sample is extremely low, compared to typical multi-species samples. The reviewer's hypothesis would imply that (i) the distinct *C. nigra* entities have diverged by no more than ~0.1% genome-wide, and (ii) hybridization is so common that hybrids dominate in our random sample.

Again, we perfectly agree that this sample of size 5 is not sufficient, far from it. It is just a demonstration that more data are needed before making drastic management decisions, since the first genome-wide data set gathered in this species contradicts the hypotheses on which management policy gas been based so far. We modified the text in several places to account for this comment, explicitly stating that our sample is insufficient to provide a definitive answer on population structure in *C. nigra* (l 138-139).

Coding data is usually constrained and this will not necessarily reflect the historical process. The authors have a clear expectation that the coding proteins would be divergent and demonstrate measures of selection, and write as if surprised by these low Fit levels, despite the many other papers

demonstrating low nuclear divergence in the turtles (e.g. (Burns et al. 2003)).

Speciation can be driven by small areas in the genomic landscape and not observed by a statistical average across all variation (Fit). The genetic differences between the pied and collared flycatcher for example closely related species but still able to hybridize, were found to be small, restricted to a small part of the overall genome that control the production of gender cells, chromosome structures, rather than gene adaptation (Ellegren et al. 2012). It may not be until speciation is further along in the evolutionary process that divergence will be evidenced in the proteome. The lack of diversity may therefore be because it is transcriptome data rather than genomes-scans.

We fully agree: the level of population differentiation is expected to vary between regions of the genome, as observed in many published data sets. This is exactly why we argue that a dozen of loci is not a sufficient sample. This rationale is at the heart of our recent Roux et al (2016 PLoS Biol), in which we show that accounting for among-loci heterogeneity in gene flow is required for reliable species delineation and inference of population history. The Roux et al (2016 PLoS Biol) paper is mainly based on transcriptomes, similar to the data set discussed in the current manuscript. We agree that genome-wide data would be even more informative, and indeed urge *C. nigra* population geneticists to move to the genome scale. We now cite Roux et al (2016 PLoS Biol) as an additional demonstration of the need for genome-scale data in population/conservation genetics.

The authors argue that tortoises have not accumulated diversity of well-established species, and compare it to a range of pair-wise species comparisons, some of which are highly divergent, which has the effect of inflating this point. The pair wise comparison of (Microtus arvalis + Microtus glareolus); these actually represent two different genera; "glareolus" is Myodes glareolus (the European common bank vole).

I find the gap in the graph between the single species and pair wise data needs a little bit further explanation. Could sampling pairs that are incipient species fill this gap?

Transcriptomes of other relatively young radiating species, maybe a more relevant comparison to include; e.g. cichlids (Brawand et al. 2014), rather than different subgenera of hares. If Fit is greater in young divergent taxa (which it is for cichlids) this would add more support to your conclusion. Or comparisons with Darwin's Finches?

In this manuscript we do not aim at sampling the continuum of divergence. Our goal is rather to illustrate that (1) *C. nigra* harbours less genetic diversity and population structure than many species of animals that nobody would think of splitting, and (2) the typical level of population structure in multi-species samples is much higher than the one we observe in *C. nigra*. To make this rationale, we need a random sample of species and species pairs. This is why we used the Romiguier et al (2014 Nature) and Galtier (2016 PLoS Genet) data sets, which have been gathered independently of population structure and species delineation issues.

Adding taxa that are intermediate between the "one species" and "several species" status, such as cichlids and Darwin's finches, would be very interesting in principle – this is the idea of the Roux et al (2016 PLoS Biol) paper – but would not add much to the argument that C. nigra clearly falls in the "one species" side. One strength of this data set, furthermore, is homogeneity of data type and data analysis pipeline (Gayral et al 2013 PLoS Genet, Ballenghien et al 2017 BMC Biol). Adding data from other studies would make the comparison less clear-cut. The *M. arvalis/M. glareolus* twospecies sample was removed from the analysis.

Management implications: If a view is taken that the turtles represent 1 species and multiple subspecies, should their present management plan be changed? The authors, I am sure would not want to retain unnatural crosses brought about by translocation in a system. Apart from stopping the current inbreeding and sterilization program, the authors have not offered an alternative management plan, and their opinion is not based on data that is expansive (taxon sampling) or conclusive (genomic wide sampling). Putting this in context with the C. becki "natural" fusion would help strengthen the arguments to leave populations alone (if that is their view).

Does it automatically follow that we should assume the populations are not in a process of separation and therefore manage them as a single population based on a single Fit measure ?

We are not saying that, or have strong recommendation at this stage. Rather, we suggest that part of the money devoted to conservation in this taxon should be used to generate genome-wide data in a large sample of Galapagos tortoises and access to the relevant information, rather than pay for transportation of animals by helicopter in order to recreate genotypes that have no demonstrated value, or species that have no demonstrated existence. We make this suggestion even more explicit in the last paragraph of the revised version, which was fully rewritten.

The group clearly appears to be a species undergoing radiation, with mtDNA and microsatellites showing structuring that is significant between populations ((Russello et al. 2010, Garrick et al. 2012). The tortoises are famed for demonstrating speciation in progress as Darwin himself observed morphological differences related to the environmental selection.

Arguments could be made about what constitutes a good species, what species concept is being followed, but non were made. Apart from morphological, MtDNA and microsat evidence is there any evidence of incompatibility between subspecies; e.g. lower fitness, breeding avoidance? If not, this would add support to your argument (see Garrick et al 2014).

Such evidence does not exist, as far as we know. Very little is known on the relationship between genotypes, phenotypes and fitness in *C. nigra*, as we recall in the revised version. For instance, whether shell shape is heritable in *C. nigra* is unknown (Fritts 1984 Biol J Lin Soc, Chiari et al 2009 PLoS ONE), in absence of common garden experiments or crosses specifically addressing this question. Current efforts to avoid "hybridization" are in no way based on arguments suggesting that hybridization could be harmful to fitness. Rather, inbreeding depression was explicitly identified by, e.g., Milinkovitch et al (2013 Evol Appl) and Edwards et al (2013 Biol Conserv) as a potential problem in *C. nigra* repatriation programs.

I am sure the authors could add more discussion regarding the founder effect, bottle-neck, life history and demographic effects on Fit; compare the data to other island endemics that have experienced recent bottleneck e.g. lemurs are in the dataset.

Lemurs have been on Madagascar for dozens of millions of years; they do not appear to be particularly informative on island-induced founder events. Our data set includes a large number of species, each with its own history of fluctuation in Ne, isolation and secondary contact. We did not modify the manuscript based on this comment.

The authors write: "…current practice whereby so called "purebred" individual are crossed is pointless". Could you reference this please, I could not find reference to this in the management plan nor, any of the papers quoted in the article.

This sentence was replaced by a longer description of past and current management plans. We now refer to Cayot (2008 Galapagos Res), Milinkovitch et al (2007 BMC Ecol) and Edwards et al (2013 Biol Conserv). Milinkovitch et al (2007 BMC Ecol) indicate that an individual carrying the "wrong" genotype was removed from an island to make sure it would not participate to reproduction – despite the acknowledged risk of inbreeding depression (Milinkovitch et al 2013 Evol Appl). Edwards et al (2013) make recommendations, based on genotypes at microsatellite loci, on which individuals should be crossed with which, and which should be released in the field.

Apart from referring to a news and views article, there is limited evidence presented of what the current management plan is for the Galapagos turtles. For those less familiar with this system it would help if the authors could provide references to the management plan itself, especially that demonstrates that it is being conducted primarily on microsatellite data a central theme to the paper. The microsatellites confirm morphological described subspecies/species also delimited by MtDNA (Ciofi et al. 2006). I understand they are not using microsatellites to delineate taxa, they found microsatellites to delineate the already morphologically described taxa and using these markers to identify hybrids (Russello et al. 2007; 2010; Garrick et al 2014) and make breeding decisions in captivity (Milinkovitch et al 2004).

We now describe in more detail the management strategy in *C. nigra* (l 60-80), based on the report by Cayot

(2008 Galapagos Res). The management plan started in the 60's long before genetic data were available, and relied on the idea of restoring islands independently (Cayot 2008 Galapagos Res), which was probably a wise decision given the available information at that time. MtDNA and microsatellite data revealed some genetic structure between islands, and between populations within large islands. Genetic data lead to a taxonomic revision of the taxon that split *C. nigra* into 13 distinct species, partly in agreement with the pre-existing subspecies. Genetic data have significantly influenced management plans in, e.g., promoting the re-creation of so-called "extinct lineages" (Edwards et al 2013 Biol Conserv, Nichols 2015 Nature) or removing so-called "hybrids" from the field (Milinkovitch et al. 2007 BMC Ecol). More generally, publications based on genetic data have promoted the view that there exist "purebred" Galapagos tortoises, each belonging to one particular species, and individuals of "mixed-ancestry", most of which are supposed to originate from unnatural causes.

We disagree with this interpretation of mtDNA and microsatellite data *C. nigra*, which we think is biased towards over-splitting and based on a simplistic model of population history, in which gene flow is a priori assumed to be of human origins. We predict that by adding more loci, a different picture might appear, in which assigning individuals to "species" will be more and more difficult. At any rate, our main point here is not to criticize previous work, but rather to call for prudent decision making, knowing the limited power of existing genetic markers, and the unexpected results reported by Loire et al (2013 Genome Biol) and the current analysis.

The Galapagos Islands conservation program states they are "continuing using advances in genetic to refine tortoise conservation". Is NGS therefore already in progress as part of the management plan?

It would seem there are large numbers of blood samples of tortoises (>1700 samples from Volcano Wolf alone, Garrick et al. 2014), more data is therefore feasible to get and is required to make such strong statements on management plans.

This is what we are asking for. One of our goals with this manuscript is to put pressure on *C. nigra* population geneticists so that they generate and publish genome-wide data in this taxon. Our group is not in position to do that – we do not have access to the samples, besides the five we collected in European zoos.

Minor recommendations: There are 54 animal species reported but, 52 in graph and with 4< individuals and in the supplementary table 1. There are more than 54 animals if the single and pair data are compared.

We now clarify that our data set includes 53 species, i.e., the 54 species of Romiguier et al (2014 Nature) with at least four individuals minus harvest ant *Messor barbarus*, in which F is strongly negative due to a highly specific mating system (Romiguier et al 2017 Mol Ecol).

116 species listed in the supplementary table S1, but there are 44 species pairs in Galtier et al 2016 (on Git hub). Where is the data from? Access to the data, could be improved with accession numbers, and a list of what individuals were pooled as the populations' esp. in the pair wise comparison in Table S1. This will also hopefully include some sampling localities, as it is difficult to understand what sampling regime has been used e.g. which 4 samples were chosen to represent humans, or where the ants were sampled, from a single colony, are they actually clones? Are all transcriptomes of a similar tissue type? Or mixed? I would like further clarification on the data used; i.e. how many loci were in each set, what proportion loci analysed were used, what is the portion that shows no variation, all this could be added to Table S1? A table of the genetic variation between the 5 samples would be useful. I am sure this won't change the overall conclusion reached, but a little more transparency would make reading the paper easier.

This information is available from the main text or Supplementary Information of the source articles Perry et al (2012 Genome Res), Romiguier et al (2014 Nature), and Galtier (2016 PLoS Genet).

This transcritpome data is also taken from only 1 tissue type (blood) while this is understandable on live animals, there is no discussion regarding the limitations of expression and therefore data from 1 tissue type. Neither do we get to know what type of tissue the other transcriptomes are from.

This information is provided and discussed in Gayral et al (2013 PLoS Genet), Romiguier et al (2014 Nature)

and Galtier (2016 PLoS Genet).

The authors write; "C. nigra harbours lower genetic polymorphism and population structuring than most vertebrates or other animal species…" However most of the vertebrate samples are mammals (16/22), (more comparative reptiles (n=1) would be appropriate in this dataset) and 32% have lower Fit values than the turtles.

A low Fit, originally proposed as a measure of inbreeding, does not per se mean there is no population structuring. Ideally the two components of within and between population diversity need to be assessed but with only 5 individuals, makes what can be done inconclusive. Fit is below 0, which means there are slightly more heterozygotes than expected than a panamictic scenario; the cause of this could be discussed- mixed ancestry samples, or deleterious mutations?

We agree that our negative estimates of F were reflecting a bias in our approach. This is why we decided to take a comparative approach rather than commenting on the estimated values. In the revised version we used an improved SNP-calling method (Ballenghien et al. 2017 BMC Biol), which accounts for cross-contamination or other sources of read count overdispersion (see below response to reviewer 2).

Figure 1: the square on the human in the graph needs to be shifted so it does not overlap with the turtle.

Main conclusion: Conserving microsatellites? I think the inference about the conservation program is premature. I fully support the authors' desires to encourage NextGen data to be completed across the Galapagos tortoise group. However, they have demonstrated that transcriptomes from blood may not be the most useful dataset, but there may be other genomic islands of diversification for example in the sex chromosomes or non-coding regions as found in other groups.

The tortoises are clearly a recent incipient speciating group, undergoing a radiation and suffered from multiple anthropogenic threats. Determining definitively whether these populations are species or subspecies, or populations, requires more data. No one will argue with the need to preserve all subspecies and populations, but the question remains on how to treat mixed ancestry individuals as natural or non-natural outcomes of habitat and species interference.

The authors here, suggest a halt to the practise of preventing these "hybirds" to contribute to the mating population. I have empathy with their view that sterilization seems a very extreme measure, as does crossing "purebreds", but I am not persuaded yet, by their argument as presented that hybrids from the islands should not be removed. Writing further on the significance of lineage fusion (Garrick et al. 2014), the ability to identify recent (<200 years vs older hybridization events) would strengthen their argument.

I have empathy with the authors, but unfortunately these data, and current literature review of other evidence are not enough to say these turtles are panmictic. A more measured approach to writing an article about these issues I do think however has merit.

Thanks for your time in reading and commenting on our ms. We hope having clarified our view and improved the manuscript based on your comments. Regarding your very last comment: we do not mean to argue that *C. nigra* is a panmictic entity, and have modified the manuscript to avoid suggesting it. We rather state that our approach revealed no obvious departure from panmixy in *C. nigra*, whereas it did so in many other species of animals and vertebrates.

Reviewed by anonymous reviewer, 2017-04-13

The letter by Loire and Galtier provides an interesting provocative reflection about the "current" strategies/practices in conservation biology relying on classical (some may say outdated) approaches in conservation genetics. They present this opinion in the framework of the conservation of the giant Galapagos tortoises, where multiple species have been suggested based on genetic data such as short sequences of the mitochondrial genome and a dozen of microsatellite loci, but also backed up by morphological and geographical differences. Here, the author underline how poorly reliable such species definition based on:

1- the low number of genetic loci used to identify and characterize the distinct species (mainly mitochondrial and ~10 microsatellite loci); 2- the low mitochondrial sequence divergence between the recognized species, which is typical of the amount of divergence seen within species in reptiles; 3- that hybrid genotypes or unexpected mito-types were considered as originating from non-natural causes and were removed from previous analyses, thus artificially increasing genetic distinctiveness; 4- and finally, that links between morphological and genetic variation was anything but clear, with the "saddleback" morphotype being observed in unrelated species.

Given the limited and conflicting amount of evidence used to delimit those species, the authors question the drastic/spectacular conservation initiatives that are currently being taken to preserve the threatened tortoise species, which involve (among other) translocating animals, conduct breeding program to reconstruct original gene pools (some being extinct), based on the limited amount of genetic information currently available. The authors call for a much more extensive genetic survey based on genome-scale data in order to reconsider the taxonomic status of the species currently recognized.

Globally, I fully agree with this assessment. In such circumstances, conservation assessment should make use of the now affordable genome-scale dataset, especially when conflicting evidence occurs. This would ensure that population and species identification is reliable enough to justify and lead the conservation initiatives, that estimation of connectivity (gene flow) between distinct groups and admixture proportion are accurately estimated and accounted for in management programs, and that characterization of the gene pools are actually representative to the entire genome, not just a few microsatellite and mitochondrial loci. However, I have a few reserves regarding the way this letter is currently written, and how the argumentation is build, before I can recommend this manuscript.

Thanks for your excellent assessment and suggestions.

1- This letter looks like a late reaction to the comment in Nature of Nichols (2015) and a critic about how species identification has been identified based on the papers initiated in 2002 (by Caccone et al. and Ciofi et al. 2002) and updated in 2015 by Poulakakis et al. using the same methodology (low number of loci, morphology and geographic data). I am thus wondering why the authors decided to write this letter now, almost 2 years after the publications of those previous papers which presented the conservation initiatives undertaken at that time, and 4 years after the publication of the paper of Loire et al 2013. I understand that this paper was already discussing (I suppose) the poor representativeness of the genetic data collected to establish the conservation units, and was ignored by the scientific community leading these conservation initiatives. To me, this is not clear why in the present letter is coming so late. I think the author should clarify why they are reacting now. The issues raised in the current letter are indeed timely needed and should be heard by the conservation politics and managers. May be the authors could also provide an update of the status of the ongoing conservation initiatives, instead of just referring to the 2015 comments (of this information are available, of course).

Our research group has no specific focus on the Giant Galapagos tortoise. This species was one out of the ~100 we selected for a long-term project on comparative population genomics in animals (e.g. Gayral et al 2013 PLoS Genet, Romiguier et al 2014 Nature, Galtier 2016 PLoS Genet, Roux et al 2016 PLoS Biol). Unlike most of the species sampled in this project, we decided to specifically analyse the *C. nigra* data and write a paper on it (Loire et al 2013 Genome Biol) because: (1) we have an interest in reptile and turtle molecular evolution (e.g. Lourenco et al 2013 J Evol Biol, Figuet et al 2014 Genome Biol Evol), (2) we had detected interesting patterns of sequence variation, potentially reflecting adaptation to local environmental conditions, (3) *C. nigra* is an iconic species and a conservation target, so publishing the first genome-scale data in this taxon was, we thought, useful, and (4) our data did not appear consistent with the prevailing view regarding the taxonomy of this group, i.e., the existence of several species – even though we clearly stated that our sample was insufficient to draw firm conclusions.

Having published this, we were expecting *C. nigra* population geneticists to seize the problem, address the conflict by generating genome-wide data in a sample of sufficient size, and draw the appropriate conclusions

regarding conservation. To our surprise, the following publications on *C. nigra* population genetics (Garrick et al 2014 Mol Ecol, Garrick et al 2015 Ecol Evol, Poulakakis et al 2015 PLoS ONE, Jensen et al 2016 Peer J) did not mention our 2013 paper or consider its message, but rather continued to rely on the same interpretation of the same mtDNA and microsatellites, as if our data were non existent. Our surprise further increased when we read the comment by Nichols (Nature, december 2015) and realized that this same 12-locus dataset was still taken as a justification for spectacular, costly interventions that may well be just irrelevant or even harmful to the animals.

Bewildered, we questioned the pertinence of our 2013 analysis and decided to check our results by adopting a comparative perspective across the many species of the project. The results were (we think) clear-cut: our approach apparently has the power to identify strong population structure and obvious species boundaries when they exist, and these apparently do not exist in *C. nigra*. We contacted Prof. Caccone in order to share our result, ask her opinion and hopefully make progress regarding the discrepancy – without success. Hence our decision to publicize this analysis, to at least let people know that the controversy exists, that more data are critically needed, and that important management decisions are made in spite of the conflicting pieces of evidence.

2- The authors state that mitochondrial and microsatellite data do not fully agree in the definition of the conservation units without providing any details on the actual issues and discrepancies. However, this is a central argument justifying the doubt raised by the authors about the validity of the currently recognized conservation units, populations, and species identification. Yet the authors do not really enter in any details and just quote Poulakakis et al (2012) as a reference, and stop short after that. The authors must make a better job at describing the actual conflict between loci, the potential or real issued in the interpretations and definitions of the conservation units, etc. If the species are actually not reproductively isolated pools, but are still very distinct populations, would that change the conservation initiatives?

This is a difficult question, which goes beyond the scope of this very manuscript. There are well known pros and cons of protecting local populations. Promoting gene flow between populations can act against local adaptation, but forced inbreeding can have deleterious consequences as well – only to speak about genetic aspects. Another related, difficult question is which level of gene flow is ”natural”, and whether we should intervene in order to decrease, increase, or equate the ”natural” level of gene flow. All these questions are, in our opinion, regrettably not addressed in current literature on *C. nigra* conservation genetics.

Conservation of giant Galapagos tortoises clearly is a success in that population density has dramatically increased over the past five decades, thanks to a breeding/repatriation program that was decided and implemented before genetic data were generated, as we now clarify. We are here criticizing the most recent decisions – sterilization/removal of so-called ”hybrids”, relocation of animals, re-creation of so-called ”extinct lineages” – which are said to be justified by genetic data. We disagree with the interpretation of existing genetic data, and ask for more data to be generated before such extreme decisions are taken.

We do not have strong recommendations yet, in absence of sufficient information. But if genomescale data and a large sample did concur with our preliminary data that population structure in *C. nigra* is ”naturally” weak, then for sure we would recommend that these initiatives be discontinued. We added a section explaining our view in more detail (last paragraph of the revised manuscript). We suggest that, based on the existing data, *C. nigra* might well be best described as a perturbed metapopulation, not a collection of endangered species, which would make a big difference in terms of management.

3- The authors also present the microsatellite loci as poorly representative of the nuclear diversity. However, I have the feeling that the authors are mixing different aspects leading to a lot of confusion.

a. The heart of the author’s arguments is the low number of loci used to identify distinct gene pools, population and species rather than the type of markers. If I am mistaken, then the authors provide much more details about why microsatellites are not appropriate. Both mitochondrial and microsatellite data were used to recognized distinct populations and species in this system, not just

microsatellites. So why stigmatizing the usage of microsatellites (see below)?

A very good point – see our response below.

b. The title of the paper even question whether we should keep using microsatellite loci at all in conservation genetics. Later in the text the authors question how representative the diversity at microsatellite level are compared to the genetic diversity of the genome. These points make little sense. Indeed, this is the low number of loci which is the real problem, not the type of loci. Hundreds and even thousands of microsatellite loci occur in genomes, but conservation geneticists usually use a dozen of them in their studies for practical reasons. Hopefully this is about to changes with the new NGS approaches (Gymrek 2017). Yet microsatellite variation is part of the diversity of genomes, among other source of variation such as SNPs, inversions, CNVs, etc. Furthermore, microsatellites display some appealing features making them very useful and informative even today. Their very fast evolutionary rate can be useful to track recent divergence, admixture proportion, very recent gene flow between population, and very recent changes in demographic histories provided that enough loci are used. Studies on model species showed how fast evolving these markers are (Gymrek 2017), how performant they can be to address key questions related to population structure and evolution within species, e.g. population structure, admixture and gene flow, such as in human (Rosenberg et al. 2002), chicken (Rosenberg et al. 2001). Furthermore these markers also show a remarkable clockwise evolution in human and potentially also in other species, making them very suitable for retracing the evolutionary history over recent time scale (Sun et al. 2009), and can also be linked with gene expression, making them more and more interesting as well (Gymrek et al. 2016). The main problem with microsatellite data is related to the difficulty and labor intensive work they require to optimize and score them accurately with classic approaches, which limited the number that can be handled. However, new NGSbased technics can solve this kind of limitations. Thus, these types of makers are still potentially very useful for tracking very recent population split, interruption of gene flow between groups, even if the level of divergence between these groups is low.

c. Although the very fast evolutionary rate of microsatellite loci can be very useful to address some questions happening at very recent time scale, they show a lot of limits related to homoplasy at larger evolutionary timescale, making them less suitable for addressing population and species evolution at larger timescale, of course. Thus, these markers might indeed not be so adequate, but it does not invalidate their usage when divergence is low. However, the authors do not really discuss any of these points, and just question their relevance without providing any details.

Now, I do not think the issues the authors want to raise is really about which type of markers should be used, but whether enough coverage of the genome has been achieved to resolve the taxonomic status of the different populations of tortoises. In the present letter, this is really that a dozen loci which is poorly representative of the genetic variation across the genome of the tortoises. This argument would have been equally true with only a dozen of SNPs scattered across the genome. I would thus suggest to change the title, removing the question "Preserving microsatellite?" and adjust the argumentation around the number of loci instead of stigmatizing the use of microsatellite loci.

We totally agree. The problem is the small number of loci, not the nature of these loci, and the manuscript was perhaps in part misleading in this respect. Our point is, because they only rely on these 12 specific loci, what people actually do in the first place by enforcing crosses between "genetically similar" individuals and removing "hybrids" is to favour specific allelic associations across these very loci – which might or might not correspond to true gene pools.

In the new version we try and make it clear that what matters is the number of loci, not the fact that these are microsatellites. We modified the last section of the text to clarify our point. We recall that there is no documented genetic incompatibility between gene pools in this species, no documented relationship between genotypes and fitness, no evidence for local adaptation, no demonstration that the major polymorphic phenotype, carapace shape, is heritable, and doubts regarding the very existence of population substructure.

This is why we suggest current practice might not preserve anything biologically relevant besides particular combinations of alleles at ~12 specific loci.

4- There seems to be some confusion around the definition and identification of conservation units, populations and species in the previous papers and the authors in the present MS does not provide much clarification. The authors could shed some clarity on the biological units (populations and species) and what this could mean in term of conservation units for the present case of the groups giant Galapagos tortoises.

a. Indeed, based on previous mitochondrial data, the authors state that the level of divergence between named species observed are not higher than typical amounts of within species variation. They also show this with their own transcriptome data previously published in Loire et al (2013). However, the authors could certainly be more explicit and phrase it in term of level of divergence along the speciation continuum, and that this level of divergence is typical of what we see between populations within a given species. I suppose that the 4 to 11 individuals in each species considered in Romiguier et al (2014) included individuals from distinct populations. If correct, the authors should insist on that fact. Otherwise, the argument is falling short if the individuals considered in these studies only come from a single population or if this information is not known.

True: samples in Romiguier et al (2014 Nature) consist in individuals from distinct populations, as we now indicate (l 111).

b. In addition, there is a common confusion among conservation geneticists (and even more among conservation biologists in general) about the difference between genetic divergence between species and genetic differentiation between populations. The authors should make sure that this terminology is clear and accessible in their MS. Indeed, the divergence between the named species seems quite low and comparable to the divergence we see between populations within species. Thus, I agree that calling these groups as distinct species may seems exaggerated. However, this is a semantic discussion of what do we call distinct populations and distinct species, and how should be defined the conservation units, based on what variables and where do we put the threshold on genetic divergence? Answering these questions are certainly quite hard, subjective and at even philosophical. However, the authors limit their argumentation mainly to genetic divergence and make an attempt on genetic differentiation (although see my comment below about that), but does not enter into any details and consideration of how should we define conservation units, based on what criteria. If the authors want to have a real impact on the decisions and be heard by the manager, they have to provide more details and be more specific.

See above our response to comment 3. Again, we think that this complex issue goes beyond the scope of this piece. Not only scientists have their word to say here. The idea of reintroducing and protecting giant Galapagos tortoises in every island and every volcano of the archipelago is perfectly fine with us irrespective of population genetics.

Our only concern here is that decisions are currently made about which individual should be crossed with which, and at which place they should be located, based on questionable genetic data and analyses. We oppose the suggestion that some *C. nigra* individuals would have less value than others (e.g. Edwards et al 2013 Biol Conserv), and should be sterilized, kept in captivity or translocated only based on their 12-loci genotypes in a context of uncertain population genetic structure.

c. Since viable hybrids seems to exist, the distinct named species are apparently still interfertile. Thus calling them distinct species seems exaggerated (following the classic evolutionary definition of a species). That being said, this does not mean that these groups (populations) are not very differentiated populations. Actually, Poulakakakis et al 2015 showed, based on a descent sampling for the two groups (20 to 30 individuals for each) they have studied with a dozen of microsatellite loci and mitochondrial data, that the 2 populations were very differentiated from each other with remarkably high FST and RST values for microsatellite data (respectively 0.20 and 0.30 in average, Fig.5 of Poulakakakis et al

2015). This already quite high for such type of loci (Jakobsson et al. 2012). In contrast, Loire and Galtier contest these results using FIT values estimated for thousands of SNP loci from only 5 individuals (one or 2 in each named species). They show in their MS that FIT values are very low, and actually negative, and conclude that population structure is actually almost inexistent. Although I agree that the Loire et al 2013 data set offers a representative genomic sampling with ~1000 of loci, 1 or 2 individuals per named species (or population) and 5 individual total does not offer at all a representative sampling of the allele frequency spectrum in each group, which is required to estimated such FIT values (FST, or FIS). Tracking departure from Hardy Weindberg Equilibrium (HWE) based on 1 or 2 individuals per population is completely inappropriate, even with ~1000 loci, because there is no way the data can be informative on the distribution of the homozygotes and heterozygotes genotypes at given locus and test for departure from HWE expectations. FST, FIS and FIT are based on allele frequency calculation, which cannot be estimated with 1 or 2 individuals per group and a total of 5.

There is indeed a conflict between the microsatellite analysis of Poulakakis et al (2015 PLoS ONE, and previous papers by the same group) and our transcriptome-based SNP analysis regarding population differentiation. Before responding to this comment, we would like to recall that we do not claim having the final say on this issue, but rather call for additional data to be generated.

On Poulakakis et al (2015):

First, a rectification: the values for Fst and Rst between the two Santa Cruz populations reported in Figure 5 of Poulakakis et al (2015 PLoS ONE) are [0.07; 0.14] and [0.14;0.21], respectively, not 0.20 and 0.30 as suggested by your comment (see dark grey bars in their Figure 5; 0.20 and 0.30 correspond to the mean/median across many pairs of *C. nigra* populations). So these newly defined "species" differ from each other by Fst of the order of 0.1, which is similar to estimates reported, based on microsatellites, between pairs of human populations, somewhat reinforcing the parallel we are making. Secondly, and most importantly, their analysis (and previous analyses by the same group) only includes so-called "purebred" individuals, meaning that individuals carrying recombinant/mixed/intermediate genotypes have been removed. Obviously, such a pre-filtering of individuals inflates the observed level of differentiation between populations. So this Fst~0.1 is an overestimate – yet the two populations have been called distinct species by Poulakakis et al (2015 PLoS ONE). We describe this analysis in more details in the revised text and explain that comparing the published microsatellite-based Fst in *C. nigra* to estimates obtained in other species is tricky given their methodology.

On our measure of Fit:

We understand the concern about sample size, but still argue that our approach has the potential to detect departure from the Hardy-Weinberg equilibrium. Our analysis considers a single population and calculates the expected number of homozygous genotypes (summed across loci) under HWE. This number is compared to the observed number of homozygous genotypes (summed across loci), and inbreeding coefficient Fit, a measure of homozygote excess, is calculated based on these two numbers. We use the estimator of Weir and Cockerham (1984 Evolution), which is unbiased even in case of small samples. The sampling variance of one-locus Fit scales inversely with sample size (i.e., number of individuals, Curie-Cohen 1982 Genetics). Given the way data was here combined across loci, the sampling variance of our multilocus Fit also scales inversely with the number of loci (SNPs), assuming they are independent. So the two dimensions (individuals, loci) are essentially equivalent in terms of sampling variance at equilibrium. Empirically, our comparative approach demonstrates that this index has the power to detect strong population structure when it exists – just compare one-species vs. two-species samples in our figure 1. Roux et al. (2016 PLoS Biol) is another illustration of the potential of this kind of data to inform on population history and species boundaries.

Let us be clear: we do not mean to argue that sampling 5 individuals is an optimal way to measure population differentiation. Rather, we simply note that if population structure was strong, we would expect to detect an excess of homozygous genotypes in this sample, which according to mtDNA and microsatellites, consists in individuals from several distinct "species". One source of confusion might be that our analysis considers a single population per species, whereas the Fst/Fis/Fit framework normally involves two levels. In single

population analyses, Fit equals Fis and Fst is not defined. To try and avoid confusion we now call the statistics F, introduce it as the inbreeding coefficient, and describe our calculation more thoroughly in the text and the legend to figure 1.

We added these additional pieces of information in the revised version, in which we highlight the need for new data and unbiased assessment of population differentiation in order to resolve the conflict between existing datasets.

d. Furthermore, the authors try to provide an ad-hock comparison of the FIT value between tortoise and human, in order to provide a comparative picture of the amount of genetic structure seen in the two species. They show that the negative FIT values found in human are comparable with those found in tortoises, arguing how absurd the is currently the species delimitation. However, the authors should note that FST and FIT values between human populations estimated in the literature are far from being negative, questioning how their estimated value is representative of differentiation among human population. Sticking with microsatellite loci for comparative purposes with the Galapagos tortoise case, Rosenberg et al (Rosenberg et al. 2002; 2005) showed that genome-wide estimates of FST values based on hundreds microsatellite loci across the genome can reach up to 0.12 between the most differentiated human populations, and is not at all negative as the authors seem to suggest. The authors should probably inspect the variance of their own FIT estimator before making any definitive conclusion. I suspect the variance would be very large given the very low (population) sample size. This also suggests that the averaged FST of 0.30 previously reported in Poulakakis et al (2015) between some turtle groups is extremely high and suggest very strong population differentiation. This is especially true since the maximum theoretical expectation for such highly polymorphic type of markers is usually bounded to a maximum of ~0.35 (Jakobsson et al. 2012). So even if the divergence between the groups is low, suggesting that population split very recently and did not accumulated much fixed differences, genetic differentiation between them is substantial. Now whether or not these should be called distinct populations, (sub)-species is a never-ending taxonomic debate in the divergence continuum of speciation, but they do certainly qualify as distinct conservation units. Now again, a dozen of loci might not be representative enough to estimates precisely population genetic parameters across the genome. More genetic data representative of the genome should thus be obtained, for sure. However, I don't think the current comparison of FIT values between human and tortoise of Loire and Galtier is meaningful, since the data of Loire et al. (2013) and Romiguier et al. (2014) also suffer from such a small population samples size, that calculation of the allele (SNP) frequency and thus F-statistics are totally unreliable.

Regarding microsatellite-based Fst estimates in C. nigra and H. sapiens: see above our response to previous comment. Assessing the significance of published estimates of Fst in C. nigra is difficult because these have been obtained after so-called "mixed ancestry" individuals were removed from the data set. Comparison with analyses in other species or theoretical expectations is therefore tricky. Still, despite the upward bias, published Fst between "species" of giant Galapagos tortoise can be as low as 0.1 (Poulakakis et al 2015 PLoS ONE), similar to values reported between human populations.

Regarding our negative F estimates: this is indeed very unexpected – and one of the reasons why we decided to take a comparative approach rather than relying on the estimate itself. Clearly, our data contain more heterozygous genotypes than would be expected from natural samples, even in panmictic populations – and this is true of a substantial number of species, so that sampling variance is probably not a sufficient explanation.

Our hypotheses regarding this bias include sequencing errors, genotyping errors, hidden paralogy and cross-contamination. We recently analyzed patterns of cross-contamination in the Romiguier et al (2014 Nature) data set (Ballenghien et al 2017 BMC Biol). In this study, we developed an improved SNP- and genotype-calling algorithm that accounts for the overdispersion of read counts due to contamination or similar sources of noise. Compared to the original approach (Gayral et al 2013 PLoS Genetics), the new method typically yields lower estimates of piS and higher estimates of F (Ballenghien et al. 2017 BMC Biol), in agreement with our hypothesis

that original estimates were biased.

The new, contamination-aware estimates of piS and F were used in this revised version. The text and Figure 1 have been modified accordingly. The point estimate for F is now significantly negative in only one species out of 53 (confidence interval obtained by bootstrap resampling). The overall message of the figure is unchanged, with single-species and two-species samples being well separated, and C. nigra clearly falling in the one-species side and very close to H. sapiens. This pattern – very low piS and F in C. nigra – is strong and robust to methodological artefacts.

In conclusion, I agree with the fact that not enough loci have been used so far to properly characterize the genetic structure and properly resolve taxonomic status of the giant Galapagos turtles. Additional genetic data representative of the genome should be obtained in order to refine the estimation of population genetic parameters and to resolve ambiguities. This should be done even before conducting any conservation initiatives. I do not think the authors are currently doing a very good job at showing the limitations and issues of previous studies. The arguments based on the low genetic divergence between named species could make sense, since calling them species may not be justified. However, the shallow genetic divergence does not mean that groups are not fully isolated. This can occur if populations have split very recently. Previous studies have shown substantial differentiation between groups. Even if this is based on a dozen of loci, this suggests that gene flow between them may be very limited or may have even ceased. A more representative assessment based on a genome-wide sampling is required, but this does not invalidate previous results that showed differentiated groups. Microsatellite can still be valuable markers to track very recent evolutionary processes. The argument about population structure being similar between what is seen in tortoises and in human makes no sense, given that estimated values for the Fstatistics are based the data set of Loire et al 2013 and Romiguier et al 2014, which include too few samples in each species leading to unreliable estimations of FIT. This preclude any proper test and accurate quantification of departure from HWE. Even if the data set include ~1000 of loci, accurate estimation of allele frequencies at each locus is required to test HWE and quantify departures from the expectations. Currently each species includes between 4 and 11 individuals, probably scattered across multiple populations. Theoretical and empirical studies show that about 30 individuals per population would be required. In the case of the tortoise, only 5 samples are available, and there I only 1 or two specimens for each group, which is totally insufficient to get any reliable estimate of any of the 3 F-statistics. I would thus suggest the author to remain focus on clarifying the current limitations of previous studies (low number of loci), the interpretation issues with the species delimitation and population differentiation, make sure that the terminology is clear even for non-specialists, and discuss what should be ideally done to resolve all the ambiguities. I would restrict the usage of the data of Loire et al 2013 and Romiguier et al 2014 to only estimation of the amount of divergence, but I would not use them to estimate any of the F-statistics, since the population sample size is too small. Previous studies reported strong genetic differentiation between at least between some of the groups, which then suggest that gene exchange between them may be quite low. Thus, these certainly qualifies as distinct conservation units. Should they be called populations, sub-species or species remained to be clarified. However more data are required to estimates accurately population genetics parameters, and these should be obtained based on an adequate sampling of the genome AND of the groups. The critics about the usage of microsatellite loci and the comparison of genetic structure between human and tortoises is not relevant in my opinion and risk to greatly reduce the impact of the message the authors want to share.

Thanks for your thorough, thoughtful review. Among other things, we hope having clarified that:

- we are not criticizing the usage of microsatellite loci;
- published estimates of Fst in *C. nigra* are to be taken with caution given the step of pre-filtering of individuals;

- our F estimator was indeed biased downwards due to experimental artefacts; correcting for these did not affect our conclusions.

Decision by Michael C. Fontaine, posted 13 April 2017

Revision needed

Dear Dr. Loire and Galtier,

We have received two thoughtful reviews of your pre-print manuscript entitled "Preserving microsatellites? Conservation genetics of the giant Galápagos tortoise." Both reviewers agree that this is a valuable and interesting letter which raise important points and issues regarding the ongoing conservation practices of the Galapagos Giant tortoises based in insufficient genetic data. However, both reviewers also raised important issues and reserves regarding the current version. Both reviewers felt that the authors should do a better job at explaining the context, the actual issues, and limitations of previous studies. Furthermore the results provided in this current version as an attempt to illustrate the poor representativeness of a dozen of loci to delineate the genetic structure and taxonomic delimitation are based on another data set which also raise a lot of questions. The two reviewers have made extensive recommendations which could greatly improve the current version and I encourage the authors to address them carefully in revised version in order to be consider for a recommendation by PCI Evol Biol.

Reviewed by anonymous reviewer 1, 13 March 2017

This paper serves to highlight the previous transcriptome work of the authors (Lore et al 2013) demonstrating that there is low genetic diversity in Galapagos tortoises and make comment on the current conservation management based primarily microsatellite data. The authors have re-used transcriptome data from 5 wild caught individual zoo animals (Lore et al 2013) who's collection data was not known but genetically assigned to 3 species/subspecies (Russello et al. 2010). They reconfirm they have found very low level of genetic diversity and the taxonomy of the group needs to be addressed. They compare the diversity within Galapagos tortoises to a range of other species diversity and finally compare the difference to a pair-wise comparison of transcriptome data from across the tree of life. They question the usefulness of the current management plan for Galapagos turtles based on microsatellite diversity.

I found this article thought provoking based on an interesting group, however found the discussion of the taxonomy, current management plan and evidence of natural hybridization in the system requires more work, the limited data and the conclusions drawn overreach the evidence as currently presented.

Comments: There is a general lack of background regarding the taxonomy of the Galapagos turtles. The recognition of separate species or interbreeding subspecies is at the heart of the authors argument. More therefore could be written regarding the taxonomic status as currently recognised Whether the taxa are species or subspecies does not affect their need for conservation. However, their ability to hybridize and whether or not that should or shouldn't be prevented when populations are so small is what is the issue is here. The management program I am assuming from the authors summary removes "hybrids" from the breeding pool; with the assumption, that hybridization is due to anthropogenic interference rather than natural causes. (This needs more references to demonstrate this is what is happening).

However, lineage fusion is also being argued for the diversity found in C. becki (Garrick et al 2014). C. becki one of the 3 species chosen in their sample, seems to be well established as a taxon that has mixed ancestry (e.g. (Poulakakis et al. 2012)). 50% of the sampled individuals have mixed ancestry (Garrick et al. 2012). Only 29% of the PBL population were found to be pure-bred individuals (Garrick et al. 2014), but crucially in support of the authors argument lineage fusion is thought to predate human mediated translocations. More should be written about this species to support the authors' view point.

The heart of this paper is based on the provenance of the 5 samples and the robustness of their assignment to different species/subspecies tortoise. I recommend these data be presented in a table in the supplementaries.

There seems to be conflicting results for species assignment using microsatellite data for 3 of the 5 samples between studies (Russello et al. 2007; Russello et al 2010). Russello et al. (2007) states the differences between MtDNA and microsat assignment for these samples is due to them having mixed ancestry, the different methods also assign them differently. After adding data from a new population, he reports different assignment again for the same samples, with lower q (Russello et al. 2010). Could the authors add a sentence or two, to assure provenance. ZUZ10/ZUZ20 are reported as mixed up when the authors sequenced their transcriptome and assigned them GA05H and GA05G in the original transcriptome paper (Lore et al 2013). It is therefore not clear which microsatellite data belongs to which transcriptome. The same samples that were assigned to PBL in 2007 but were assigned to different populations in 2010, both were thought to mixed ancestry. All 4 samples had q-values below 0.79, there seems to be no data on the 5th sample.

Can the authors comment on these samples population assignment. Is it unclear? 1. Can you provide evidence that the samples are not from mixed ancestry? The micosats suggest all 4 samples are. And if they are, comment on how that would affect your conclusions about panmixa? 2. The fifth sample added to the data set in 2013, has a MtDNA assignment but I could not find a microsatellite profile for comparison. 3. Does mixed ancestry explain the genetic differences found for GAO5H mentioned in the text? 4. Do the authors have evidence that the microsatellites are not robust at identifying populations with confidence? It would strengthen the authors point.

Coding data is usually constrained and this will not necessarily reflect the historical process. The authors have a clear expectation that the coding proteins would be divergent and demonstrate measures of selection, and write as if surprised by these low Fit levels, despite the many other papers demonstrating low nuclear divergence in the turtles (e.g. (Burns et al. 2003)).

Speciation can be driven by small areas in the genomic landscape and not observed by a statistical average across all variation (Fit). The genetic differences between the pied and collared flycatcher for example closely related species but still able to hybridize, were found to be small, restricted to a small part of the overall genome that control the production of gender cells, chromosome structures, rather than gene adaptation (Ellegren et al. 2012). It may not be until speciation is further along in the evolutionary process that divergence will be evidenced in the proteome. The lack of diversity may therefore be because it is transcriptome data rather than genomes-scans.

The authors argue that tortoises have not accumulated diversity of well-established species, and compare it to a range of pair-wise species comparisons, some of which are highly divergent, which has the effect of inflating this point. The pair wise comparison of (Microtus arvalis + Microtus glareolus); these actually represent two different genera; “glareolus” is Myodes glareolus (the European common bank vole).

I find the gap in the graph between the single species and pair wise data needs a little bit further explanation. Could sampling pairs that are incipient species fill this gap? Transcriptomes of other relatively young radiating species, maybe a more relevant comparison to include; e.g. cichlids (Brawand et al. 2014), rather than different subgenera of hares. If Fit is greater in young divergent taxa (which it is for cichlids) this would add more support to your conclusion. Or comparisons with Darwin’s Finches?

Management implications: If a view is taken that the turtles represent 1 species and multiple subspecies, should their present management plan be changed? The authors, I am sure would not want to retain unnatural crosses brought about by translocation in a system. Apart from stopping the current inbreeding and sterilization program, the authors have not offered an alternative management plan, and their opinion is not based on data that is expansive (taxon sampling) or conclusive (genomic wide sampling). Putting this in context with the C. becki “natural” fusion would help strengthen the arguments to leave populations alone (if that is thier view).

Does it automatically follow that we should assume the populations are not in a process of separation and therefore manage them as a single population based on a single Fit measure ?

The group clearly appears to be a species undergoing radiation, with mtDNA and microsatellites showing structuring that is significant between populations ((Russello et al. 2010, Garrick et al. 2012). The tortoises are famed for demonstrating speciation in progress as Darwin himself observed morphological differences related

to the environmental selection.

Arguments could be made about what constitutes a good species, what species concept is being followed, but non were made. Apart from morphological, MtDNA and microsat evidence is there any evidence of incompatibility between subspecies; e.g. lower fitness, breeding avoidance? If not, this would add support to your argument (see Garrick et al 2014).

I am sure the authors could add more discussion regarding the founder effect, bottle-neck, life history and demographic effects on Fit; compare the data to other island endemics that have experienced recent bottleneck e.g. lemurs are in the dataset.

The authors write: "…current practice whereby so called "purebred" individual are crossed is pointless". Could you reference this please, I could not find reference to this in the management plan nor, any of the papers quoted in the article.

Apart from referring to a news and views article, there is limited evidence presented of what the current management plan is for the Galapagos turtles. For those less familiar with this system it would help if the authors could provide references to the management plan itself, especially that demonstrates that it is being conducted primarily on microsatellite data a central theme to the paper. The microsatellites confirm morphological described subspecies/species also delimited by MtDNA (Ciofi et al. 2006). I understand they are not using microsatellites to delineate taxa, they found microsatellites to delineate the already morphologically described taxa and using these markers to identify hybrids (Russello et al. 2007; 2010; Garrick et al 2014) and make breeding decisions in captivity (Milinkovitch et al 2004).

The Galapagos Islands conservation program states they are "continuing using advances in genetic to refine tortoise conservation". Is NGS therefore already in progress as part of the management plan?

It would seem there are large numbers of blood samples of tortoises (>1700 samples from Volcano Wolf alone, Garrick et al. 2014), more data is therefore feasible to get and is required to make such strong statements on management plans.

Minor recommendations: There are 54 animal species reported but, 52 in graph and with 4< individuals and in the supplementary table 1. There are more than 54 animals if the single and pair data are compared.

116 species listed in the supplementary table S1, but there are 44 species pairs in Galtier et al 2016 (on Git hub). Where is the data from? Access to the data, could be improved with accession numbers, and a list of what individuals were pooled as the populations' esp. in the pair wise comparison in Table S1. This will also hopefully include some sampling localities, as it is difficult to understand what sampling regime has been used e.g. which 4 samples were chosen to represent humans, or where the ants were sampled, from a single colony, are they actually clones? Are all transcriptomes of a similar tissue type? Or mixed? I would like further clarification on the data used; i.e. how many loci were in each set, what proportion loci analysed were used, what is the portion that shows no variation, all this could be added to Table S1? A table of the genetic variation between the 5 samples would be useful. I am sure this won't change the overall conclusion reached, but a little more transparency would make reading the paper easier.

This transcritpome data is also taken from only 1 tissue type (blood) while this is understandable on live animals, there is no discussion regarding the limitations of expression and therefore data from 1 tissue type. Neither do we get to know what type of tissue the other transcriptomes are from.

The authors write; "C. nigra harbours lower genetic polymorphism and population structuring than most vertebrates or other animal species…" However most of the vertebrate samples are mammals (16/22), (more comparative reptiles (n=1) would be appropriate in this dataset) and 32% have lower Fit values than the turtles.

A low Fit, originally proposed as a measure of inbreeding, does not per se mean there is no population structuring. Ideally the two components of within and between population diversity need to be assessed but with only 5 individuals, makes what can be done inconclusive. Fit is below 0, which means there are slightly more heterozygotes than expected than a panamictic scenario; the cause of this could be discussed- mixed ancestry samples, or deleterious mutations?

Figure 1: the square on the human in the graph needs to be shifted so it does not overlap with the turtle.

Main conclusion: Conserving microsatellites? I think the inference about the conservation program is premature. I fully support the authors' desires to encourage NextGen data to be completed across the Galapagos tortoise group. However, they have demonstrated that transcriptomes from blood may not be the most useful dataset, but there may be other genomic islands of diversification for example in the sex chromosomes or non-coding regions as found in other groups.

The tortoises are clearly a recent incipient speciating group, undergoing a radiation and suffered from multiple anthropogenic threats. Determining definitively whether these populations are species or subspecies, or populations, requires more data. No one will argue with the need to preserve all subspecies and populations, but the question remains on how to treat mixed ancestry individuals as natural or non-natural outcomes of habitat and species interference.

The authors here, suggest a halt to the practise of preventing these "hybirds" to contribute to the mating population. I have empathy with their view that sterilization seems a very extreme measure, as does crossing "purebreds", but I am not persuaded yet, by their argument as presented that hybrids from the islands should not be removed. Writing further on the significance of lineage fusion (Garrick et al. 2014), the ability to identify recent (<200 years vs older hybridization events) would strengthen their argument.

I have empathy with the authors, but unfortunately these data, and current literature review of other evidence are not enough to say these turtles are panmictic. A more measured approach to writing an article about these issues I do think however has merit.

Brawand, D., C. E. et al. 2014. The genomic substrate for adaptive radiation in African cichlid fish. Nature 513:375-381. Burns, C., C. Ciofi, L. Beheregaray, T. Fritts, J. Gibbs, C. Márquez, and A. Caccone. 2003. The origin of captive Galápagos tortoises based on DNA analysis: Implications for the management of natural populations. Animal conservation 6, :329-337. Ciofi, C., G. Wilson, and L. Beheregaray. 2006. Phylogeographic history and gene flow among giant Galapagos tortoises on southern Isabela Island. . Genetics 172. Ellegren, H., L. et. al. 2012. The genomic landscape of species divergence in Ficedula flycatchers. Nature 491:756-760. Garrick, R. C., et. al. 2012. Genetic rediscovery of an 'extinct' Galápagos giant tortoise species. Current Biology 22:R10-R11. Garrick, R. C., et. al. 2014. Lineage fusion in Galápagos giant tortoises. Molecular Ecology 23:5276-5290. Poulakakis, N., M. Russello, D. Geist, and A. Caccone. 2012. Unravelling the peculiarities of island life: vicariance, dispersal and the diversification of the extinct and extant giant Galápagos tortoises. Pages 160-173 Molecular Ecology. Russello, M. A., C. Hyseni, J. P. Gibbs, S. Cruz, C. Marquez, W. Tapia, P. Velensky, J. R. Powell, and A. Caccone. 2007. Lineage identification of Galápagos tortoises in captivity worldwide. Animal conservation 10:304-311. Russello, M. A., N. Poulakakis, J. P. Gibbs, W. Tapia, E. Benavides, J. R. Powell, and A. Caccone. 2010. DNA from the Past Informs Ex Situ Conservation for the Future: An "Extinct" Species of Galápagos Tortoise Identified in Captivity. Plos One 5:e8683.

Reviewed by anonymous reviewer 2, 14 March 2017

The letter by Loire and Galtier provides an interesting provocative reflection about the "current" strategies/practices in conservation biology relying on classical (some may say outdated) approaches in conservation genetics. They present this opinion in the framework of the conservation of the giant Galapagos tortoises, where multiple species have been suggested based on genetic data such as short sequences of the mitochondrial genome and a dozen of microsatellite loci, but also backed up by morphological and geographical differences. Here, the author underline how poorly reliable such species definition based on:

1- the low number of genetic loci used to identify and characterize the distinct species (mainly mitochondrial and ~10 microsatellite loci); 2- the low mitochondrial sequence divergence between the recognized species, which is typical of the amount of divergence seen within species in reptiles; 3- that hybrid genotypes or unexpected mito-types were considered as originating from non-natural causes and were removed from previous analyses, thus artificially increasing genetic distinctiveness; 4- and finally, that links between morphological and genetic variation was anything but clear, with the "saddleback" morphotype being observed in unrelated species.

Given the limited and conflicting amount of evidence used to delimit those species, the authors question the drastic/spectacular conservation initiatives that are currently being taken to preserve the threatened tortoise species, which involve (among other) translocating animals, conduct breeding program to reconstruct original gene pools (some being extinct), based on the limited amount of genetic information currently available. The authors call for a much more extensive genetic survey based on genome-scale data in order to reconsider the taxonomic status of the species currently recognized.

Globally, I fully agree with this assessment. In such circumstances, conservation assessment should make use of the now affordable genome-scale dataset, especially when conflicting evidence occurs. This would ensure that population and species identification is reliable enough to justify and lead the conservation initiatives, that estimation of connectivity (gene flow) between distinct groups and admixture proportion are accurately estimated and accounted for in management programs, and that characterization of the gene pools are actually representative to the entire genome, not just a few microsatellite and mitochondrial loci. However, I have a few reserves regarding the way this letter is currently written, and how the argumentation is build, before I can recommend this manuscript.

1- This letter looks like a late reaction to the comment in Nature of Nichols (2015) and a critic about how species identification has been identified based on the papers initiated in 2002 (by Caccone et al. and Ciofi et al. 2002) and updated in 2015 by Poulakakis et al. using the same methodology (low number of loci, morphology and geographic data). I am thus wondering why the authors decided to write this letter now, almost 2 years after the publications of those previous papers which presented the conservation initiatives undertaken at that time, and 4 years after the publication of the paper of Loire et al 2013. I understand that this paper was already discussing (I suppose) the poor representativeness of the genetic data collected to establish the conservation units, and was ignored by the scientific community leading these conservation initiatives. To me, this is not clear why in the present letter is coming so late. I think the author should clarify why they are reacting now. The issues raised in the current letter are indeed timely needed and should be heard by the conservation politics and managers. May be the authors could also provide an update of the status of the ongoing conservation initiatives, instead of just referring to the 2015 comments (of this information are available, of course).

2- The authors state that mitochondrial and microsatellite data do not fully agree in the definition of the conservation units without providing any details on the actual issues and discrepancies. However, this is a central argument justifying the doubt raised by the authors about the validity of the currently recognized conservation units, populations, and species identification. Yet the authors do not really enter in any details and just quote Poulakakis et al (2012) as a reference, and stop short after that. The authors must make a better job at describing the actual conflict between loci, the potential or real issued in the interpretations and definitions of the conservation units, etc. If the species are actually not reproductively isolated pools, but are still very distinct populations, would that change the conservation initiatives?

3- The authors also present the microsatellite loci as poorly representative of the nuclear diversity. However, I have the feeling that the authors are mixing different aspects leading to a lot of confusion.

a. The heart of the author's arguments is the low number of loci used to identify distinct gene pools, population and species rather than the type of markers. If I am mistaken, then the authors provide much more details about why microsatellites are not appropriate. Both mitochondrial and microsatellite data were used to recognized distinct populations and species in this system, not just microsatellites. So why stigmatizing the usage of microsatellites (see below)?

b. The title of the paper even question whether we should keep using microsatellite loci at all in conservation genetics. Later in the text the authors question how representative the diversity at microsatellite level are compared to the genetic diversity of the genome. These points make little sense. Indeed, this is the low number of loci which is the real problem, not the type of loci. Hundreds and even thousands of microsatellite loci occur in genomes, but conservation geneticists usually use a dozen of them in their studies for practical reasons. Hopefully this is about to changes with the new NGS approaches (Gymrek 2017). Yet microsatellite variation is part of the diversity of genomes, among other source of variation such as SNPs, inversions, CNVs,

etc. Furthermore, microsatellites display some appealing features making them very useful and informative even today. Their very fast evolutionary rate can be useful to track recent divergence, admixture proportion, very recent gene flow between population, and very recent changes in demographic histories provided that enough loci are used. Studies on model species showed how fast evolving these markers are (Gymrek 2017), how performant they can be to address key questions related to population structure and evolution within species, e.g. population structure, admixture and gene flow, such as in human (Rosenberg et al. 2002), chicken (Rosenberg et al. 2001). Furthermore these markers also show a remarkable clockwise evolution in human and potentially also in other species, making them very suitable for retracing the evolutionary history over recent time scale (Sun et al. 2009), and can also be linked with gene expression, making them more and more interesting as well (Gymrek et al. 2016). The main problem with microsatellite data is related to the difficulty and labor intensive work they require to optimize and score them accurately with classic approaches, which limited the number that can be handled. However, new NGS-based technics can solve this kind of limitations. Thus, these types of makers are still potentially very useful for tracking very recent population split, interruption of gene flow between groups, even if the level of divergence between these groups is low.

c. Although the very fast evolutionary rate of microsatellite loci can be very useful to address some questions happening at very recent time scale, they show a lot of limits related to homoplasy at larger evolutionary timescale, making them less suitable for addressing population and species evolution at larger timescale, of course. Thus, these markers might indeed not be so adequate, but it does not invalidate their usage when divergence is low. However, the authors do not really discuss any of these points, and just question their relevance without providing any details.

Now, I do not think the issues the authors want to raise is really about which type of markers should be used, but whether enough coverage of the genome has been achieved to resolve the taxonomic status of the different populations of tortoises. In the present letter, this is really that a dozen loci which is poorly representative of the genetic variation across the genome of the tortoises. This argument would have been equally true with only a dozen of SNPs scattered across the genome. I would thus suggest to change the title, removing the question “Preserving microsatellite?” and adjust the argumentation around the number of loci instead of stigmatizing the use of microsatellite loci.

4- There seems to be some confusion around the definition and identification of conservation units, populations and species in the previous papers and the authors in the present MS does not provide much clarification. The authors could shed some clarity on the biological units (populations and species) and what this could mean in term of conservation units for the present case of the groups giant Galapagos tortoises.

a. Indeed, based on previous mitochondrial data, the authors state that the level of divergence between named species observed are not higher than typical amounts of within species variation. They also show this with their own transcriptome data previously published in Loire et al (2013). However, the authors could certainly be more explicit and phrase it in term of level of divergence along the speciation continuum, and that this level of divergence is typical of what we see between populations within a given species. I suppose that the 4 to 11 individuals in each species considered in Romiguier et al (2014) included individuals from distinct populations. If correct, the authors should insist on that fact. Otherwise, the argument is falling short if the individuals considered in these studies only come from a single population or if this information is not known.

b. In addition, there is a common confusion among conservation geneticists (and even more among conservation biologists in general) about the difference between genetic divergence between species and genetic differentiation between populations. The authors should make sure that this terminology is clear and accessible in their MS. Indeed, the divergence between the named species seems quite low and comparable to the divergence we see between populations within species. Thus, I agree that calling these groups as distinct species may seems exaggerated. However, this is a semantic discussion of what do we call distinct populations and distinct species, and how should be defined the conservation units, based on what variables and where do we put the threshold on genetic divergence? Answering these questions are certainly quite hard, subjective and at even philosophical. However, the authors limit their argumentation mainly to genetic divergence and

make an attempt on genetic differentiation (although see my comment below about that), but does not enter into any details and consideration of how should we define conservation units, based on what criteria. If the authors want to have a real impact on the decisions and be heard by the manager, they have to provide more details and be more specific.

c. Since viable hybrids seems to exist, the distinct named species are apparently still interfertile. Thus calling them distinct species seems exaggerated (following the classic evolutionary definition of a species). That being said, this does not mean that these groups (populations) are not very differentiated populations. Actually, Poulakakakis et al 2015 showed, based on a descent sampling for the two groups (20 to 30 individuals for each) they have studied with a dozen of microsatellite loci and mitochondrial data, that the 2 populations were very differentiated from each other with remarkably high FST and RST values for microsatellite data (respectively 0.20 and 0.30 in average, Fig.5 of Poulakakakis et al 2015). This already quite high for such type of loci (Jakobsson et al. 2012). In contrast, Loire and Galtier contest these results using FIT values estimated for thousands of SNP loci from only 5 individuals (one or 2 in each named species). They show in their MS that FIT values are very low, and actually negative, and conclude that population structure is actually almost inexistent. Although I agree that the Loire et al 2013 data set offers a representative genomic sampling with ~1000 of loci, 1 or 2 individuals per named species (or population) and 5 individual total does not offer at all a representative sampling of the allele frequency spectrum in each group, which is required to estimated such FIT values (FST, or FIS). Tracking departure from Hardy Weindberg Equilibirum (HWE) based on 1 or 2 individuals per population is completely inappropriate, even with ~1000 loci, because there is no way the data can be informative on the distribution of the homozygotes and heterozygotes genotypes at given locus and test for departure from HWE expectations. FST, FIS and FIT are based on allele frequency calculation, which cannot be estimated with 1 or 2 individuals per group and a total of 5.

d. Furthermore, the authors try to provide an ad-hock comparison of the FIT value between tortoise and human, in order to provide a comparative picture of the amount of genetic structure seen in the two species. They show that the negative FIT values found in human are comparable with those found in tortoises, arguing how absurd the is currently the species delimitation. However, the authors should note that FST and FIT values between human populations estimated in the literature are far from being negative, questioning how their estimated value is representative of differentiation among human population. Sticking with microsatellite loci for comparative purposes with the Galapagos tortoise case, Rosenberg et al (Rosenberg et al. 2002; 2005) showed that genome-wide estimates of FST values based on hundreds microsatellite loci across the genome can reach up to 0.12 between the most differentiated human populations, and is not at all negative as the authors seem to suggest. The authors should probably inspect the variance of their own FIT estimator before making any definitive conclusion. I suspect the variance would be very large given the very low (population) sample size. This also suggests that the averaged FST of 0.30 previously reported in Poulakakis et al (2015) between some turtle groups is extremely high and suggest very strong population differentiation. This is especially true since the maximum theoretical expectation for such highly polymorphic type of markers is usually bounded to a maximum of ~0.35 (Jakobsson et al. 2012). So even if the divergence between the groups is low, suggesting that population split very recently and did not accumulated much fixed differences, genetic differentiation between them is substantial. Now whether or not these should be called distinct populations, (sub)-species is a never-ending taxonomic debate in the divergence continuum of speciation, but they do certainly qualify as distinct conservation units. Now again, a dozen of loci might not be representative enough to estimates precisely population genetic parameters across the genome. More genetic data representative of the genome should thus be obtained, for sure. However, I don't think the current comparison of FIT values between human and tortoise of Loire and Galtier is meaningful, since the data of Loire et al. (2013) and Romiguier et al. (2014) also suffer from such a small population samples size, that calculation of the allele (SNP) frequency and thus F-statistics are totally unreliable.

In conclusion, I agree with the fact that not enough loci have been used so far to properly characterize the genetic structure and properly resolve taxonomic status of the giant Galapagos turtles. Additional genetic

data representative of the genome should be obtained in order to refine the estimation of population genetic parameters and to resolve ambiguities. This should be done even before conducting any conservation initiatives. I do not think the authors are currently doing a very good job at showing the limitations and issues of previous studies. The arguments based on the low genetic divergence between named species could make sense, since calling them species may not be justified. However, the shallow genetic divergence does not mean that groups are not fully isolated. This can occur if populations have split very recently. Previous studies have shown substantial differentiation between groups. Even if this is based on a dozen of loci, this suggests that gene flow between them may be very limited or may have even ceased. A more representative assessment based on a genome-wide sampling is required, but this does not invalidate previous results that showed differentiated groups. Microsatellite can still be valuable markers to track very recent evolutionary processes. The argument about population structure being similar between what is seen in tortoises and in human makes no sense, given that estimated values for the F-statistics are based the data set of Loire et al 2013 and Romiguier et al 2014, which include too few samples in each species leading to unreliable estimations of FIT. This preclude any proper test and accurate quantification of departure from HWE. Even if the data set include ~1000 of loci, accurate estimation of allele frequencies at each locus is required to test HWE and quantify departures from the expectations. Currently each species includes between 4 and 11 individuals, probably scattered across multiple populations. Theoretical and empirical studies show that about 30 individuals per population would be required. In the case of the tortoise, only 5 samples are available, and there I only 1 or two specimens for each group, which is totally insufficient to get any reliable estimate of any of the 3 F-statistics. I would thus suggest the author to remain focus on clarifying the current limitations of previous studies (low number of loci), the interpretation issues with the species delimitation and population differentiation, make sure that the terminology is clear even for non-specialists, and discuss what should be ideally done to resolve all the ambiguities. I would restrict the usage of the data of Loire et al 2013 and Romiguier et al 2014 to only estimation of the amount of divergence, but I would not use them to estimate any of the F-statistics, since the population sample size is too small. Previous studies reported strong genetic differentiation between at least between some of the groups, which then suggest that gene exchange between them may be quite low. Thus, these certainly qualifies as distinct conservation units. Should they be called populations, sub-species or species remained to be clarified. However more data are required to estimates accurately population genetics parameters, and these should be obtained based on an adequate sampling of the genome AND of the groups. The critics about the usage of microsatellite loci and the comparison of genetic structure between human and tortoises is not relevant in my opinion and risk to greatly reduce the impact of the message the authors want to share.

Ref. cited

Gymrek M (2017) A genomic view of short tandem repeats. Current opinion in genetics & development, 44, 9–16.

Gymrek M, Willems T, Guilmatre A et al. (2016) Abundant contribution of short tandem repeats to gene expression variation in humans. Nature Genetics, 48, 22–29.

Jakobsson M, Edge MD, Rosenberg NA (2012) The Relationship Between FST and the Frequency of the Most Frequent Allele. Genetics.

Rosenberg NA, Burke T, Elo K et al. (2001) Empirical evaluation of genetic clustering methods using multilocus genotypes from 20 chicken breeds. Genetics, 159, 699–713.

Rosenberg NA, Mahajan S, Ramachandran S et al. (2005) Clines, Clusters, and the Effect of Study Design on the Inference of Human Population Structure. PLoS Genetics, 1, e70.

Rosenberg NA, Pritchard JK, Weber JL et al. (2002) Genetic structure of human populations. Science, 298, 2381–2385.

Sun JX, Mullikin JC, Patterson N, Reich DE (2009) Microsatellites are molecular clocks that support accurate inferences about history. Molecular Biology and Evolution, 26, 1017–1027.