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.

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').
*** = assignments different for the same sample between studies or genetic data.**

Mitochondrial DNA control region

	#	Residence since	Haplotype	Population	Island	GENBANK
GA05F	ZUZ01	1949	18	CRU	Santa Cruz	AF548221/ GQ220702
GA05H or GA05G	ZUZ10	1962	52	PBL	Isabela	AF548255
GA05H or GA05G	ZUZ20	1962	52	PBL*	Isabela	AF548255
GA05E	ZUZ30	1962	61	VA*	Isabela	AF548264
GA05A			Rotterdam Zoo (Netherlands)	PBL	clade c,	SRS509366

Microsatellite data

Russell et al 2010

		Pitchard 2000		Rannala & Mountain 1997		
	#	Popul	q		L_1	Island
GA05F	ZUZ01	CRU	0.796	CRU	15.66	Santa Cruz
GA05H or GA05G	ZUZ10	PBL	0.33	PBL	14.07	Isabela
GA05H or GA05G	ZUZ20	AGO*	0.693	PBL*	12.95	Santiago
GA05E	ZUZ30	CR*	0.47	VD*	14.66	Isabela

Russell et al 2007

		Pitchard 2000			Rannala & Mountain (1997)	
	#	Popul	q	Island		L_1
GA05F	ZUZ01	CR*	0.910	Santa Cruz	CRU	17.74
GA05H or GA05G	ZUZ10	PBL	0.852	Isabela	PBL	15.89
GA05H or GA05G	ZUZ20	PBL	0.503	Santiago	PBL*	14.58
GA05E	ZUZ30	CR*	0.950	Isabela	VD*	16.67