Response to Reviewer 3 Comments

Point 1. The title of the manuscript should be changed. I suggest: Genetic characterization of the Chocholteca Creole sheep population from Oaxaca, Mexico.

Response 1: The title you suggest is good but we do not see any problem with our title, we think both express the same idea, so if there is no problem, we keep ours.

Point 2. The word race should be replaced with breed.

Response 2: The word "race" has been changed in all the manuscript to "breed".

Point 3. It should be noted that currently whole genome SNP analysis is more efficient tool than microsatellites and why the authors think that STR analysis is appropriate in their study.

Response 3: We agree with the referee, but this project did not receive any external funds for its execution, as has been detailed in the "Funding" section of the draft. Furthermore, the database used to compare the Chocholteca Creole sheep with other Creole breeds has been built with microsatellites as was recommended by FAO (2007). Unfortunately, most of these breeds are endangered local populations that don´t receive any financial support to study their genetic diversity, so we and the rest of the BIOVIS Consortium researchers have carried out our studies with our own resources. Genotyping of populations with high-throughput SNP panels is still quite expensive and is unaffordable without any external funding. Furthermore, according with FAO 2011 (Molecular genetic characterization of animal genetic resources. Vol. 9. FAO Animal Production and Health Guidelines. Rome. http://www.fao.org/docrep/014/i2413e/i2413e00.htm.) one inconvenient of the use of SNP for the study of AnGR diversity is the ascertainment bias because of the population in which the SNP were discovered. The commercially available SNP chips have largely been constructed with the objective of supporting genomic selection, and thus a majority of SNP in these panels originate from international transboundary breeds. Some of these SNP may be monomorphic in local breeds, whereas loci that are polymorphic only in local breeds are likely to be excluded. In addition, the SNP in commercial panels have been selected to have high minor-allele frequency (MAF) and, consequently, greater variability in the international transboundary breeds, without considering variability in other breeds. As a consequence, diversity in the local breeds, can be underestimated and estimates of relationships among breeds can be distorted.
Anyway, we know that nowadays the use of SNP chips is needed, and we hope we have funding to address the study of this and other local breeds with SNP in a nearby future.

Point 4. Instead of Nei's Da, Reynolds and Fst distances I would recommend to use Jost's D genetic distances (Jost L. 2008. GST and its relatives do not measure differentiation. Mol Ecol 17: 4015-4026.) which were designed for microsatellites. It could be done in the R package diveRsity. The input file for this package is GENEPOP which the authors already have, so I don't think this will be difficult.

Response 4: We followed the FAO recommendations to assess genetic diversity of closely related populations (FAO, 2007. The State of the World's Animal Genetic Resources for Food and Agriculture Ed.Dafydd Pilling and Barbara Rischkowsky. Rome: Dafydd Pilling & Barbara Rischkowsky). In this report, FAO recommends the use of DA genetic distance.

In a previous survey, DA is the Genetic distance measure used in most studies of genetic diversity in sheep (Baumung, R., H. Simianer, y I. Hoffmann. 2004. «Genetic Diversity Studies in Farm Animals – a Survey». Journal of Animal Breeding and Genetics 121 (6): 361-73. https://doi.org/10.1111/j.1439-0388.2004.00479.x).
Many publications can be found were Reynolds or DA genetic distances was used to determine genetic relationships among Creole or other local breeds. For example,
cattle breeds (https://biobovis.jimdofree.com/publicaciones-publications/),
goats (https://biogoat.jimdofree.com/publicaciones-publications/),
pigs (https://biopig.jimdofree.com/publicaciones-publications/) or
horses (https://biohorse.jimdofree.com/publicaciones-publications/)
These and other papers in prestigious journals demonstrate that the results obtained with DA and Reynolds genetic distances are valid.
We will calculate Jost's D genetic distance in a future study, it will be interesting to check the differences between the results obtained with both methods in our local breeds.

Point 5. The quality of figures (Figure 2, Figure 3, Figure 4) must be improved. They all become blurry after they are zoomed in. For example: the Neighbor-Net could be saved as PDF in the SplitsTree software and then converted to high quality (600 dps) PNG in inkscape program.

Response 5: The figures 2, 3 and 4 have been improved.

Point 6. In the Figure 4, the names of the breeds should be given instead of the numbers. It would be easier to understand the results.

Response 6: The numbers have been substituted by the names of the breeds in Figure 4.

Point 7. Geographic coordinates of the Oaxaca region should be checked. I assume it is 96° and not 19°. If it is written – "between", then the coordinates indicating another border must be specified. But perhaps it is better to indicate the coordinates of the central part of the region.

Response 7: The geographic coordinate has been corrected to 96º. Such coordinates correspond to the region of "la estancia", and the word "between" has been removed.

Point 8. The studied individuals should also be checked for their relatedness. One individual from the pair of full sibs should be removed. The presence of closely related animals (full sibs) could bias the study results. For example, it could be done using software ML-Relate (https://www.montana.edu/kalinowski/software/).

Response 8: In farm local breeds, most of them endangered, it's possible to find closely related animals. Usually, genealogical records don't exist. Animals should be randomly sampled for biodiversity studies, and to try sampling only unrelated animals would cause a bias. To validate Structure results, we use a high number of breeds and some reference populations whose historic origin and evolution is well known. The publications referred in the answer to comment 4 show that our results are valid.

Point 9. In the Table 1 Criollo sheep should be written as Creole.

Response 9: The word Criollo has been changed by Creole in table 1.

Point 10. As far as I know, an automatic sequencer ABI 3130XL (Applied Biosystem) allows to conduct a capillary electrophoresis instead of the polyacrylamide gel as the authors indicated. Please, clarify this.

Response 10: The 4-capillary electrophoresis was written in the methodology.

Point 11. I recommend to include a table with genetic diversity of all the studied breeds. The following parameters are needed: observed heterozygosity (Ho), expected heterozygosity (He), inbreeding coefficient (Fis) and allelic richness (Ar). So it will be easier to compare Creole from Oaxaca to the other breeds.

Response 11: A table with the parameters of observed heterozygosity (Ho), expected heterozygosity (He), inbreeding coefficient (Fis) and allelic richness (Ar) to each sheep was added as supplementary file.

Point 12. The authors should indicate the current census size of the Chocholteca Creole sheep population from Oaxaca.

Response 12: The census size has been indicated in introduction section

Point 13. The abbreviations of the breeds should be the same as throughout the entire manuscript. Now in the Table 1 Chocholteca Creole sheep is given as CChO, in the Table 4 – Ccho, in the Figure 3 – MX.

Response 13: The abbreviation of Chocholteca Creole sheep has been written as CChO in the entire manuscript.

Point 14. I suggest to move Table 2 and Table 3 as well as Figure 2 to Supplementary material.

Response 14: We consider that table 2, 3 and figure 2 show relevant results in the genetic characterization of the Chocholteca Creole Sheep and that why they must be part of the principal manuscript.