Reply to Moura *et al*

The authors appreciate the comments of Moura et al[1] in their letter and agree that when the markers are available, Amerindian genotypes probably better serve as a pseudoancestor group for the Brazilian population than the Chinese when this is the only comparison or consideration to be made. We further agree that in some, but not all instances, use of a Han Chinese data set in this way may result in inflation of 'Amerindian' ancestry relative to the Native American Microsatellite (NAM) panel of the HGDP-CEPH Diversity project or the AMI panel used in their analyses. However, there is no evidence that this would change any of the conclusions in the paper. We observe the following: (1) There *was* a comparison of ancestry estimates using pseudoancestors from the NAM panel with those from the CHB (Figure 2c and d). (2) Ancestry estimates are relative to the specific samples included and the markers used, so the influence of a pseudoancestor group will depend on the composition of the admixed sample being analyzed. (3) The degree of differentiation between groups in the paper by Pena et al cannot be taken as representative of these areas. (4) What degree of error is important depends on the magnitude of the error and on the question being asked.

For our study, the markers used to genotype the Brazilian population were selected in 2004–2006 for their relevance to a study on dengue along with ~40 selected specifically as AIMs. Twice as many of the selected markers were present in the Hapmap CHB panel (237 SNPs) as the NAM set (108 SNPs). We did, however, directly compare the use of the CHB and NAM panels in our Figure 2. Visually the differences appear small and would not change any of the conclusions about association with self-identified ethnicity. We do find a 2–3-fold higher 'Amerindian' ancestry for Whites in Fortaleza when using the CHB population as pseudoancestors compared with the NAM. Because the accuracy of ancestry estimates improves with the number of markers used (informativeness for ancestry being equal[2]), the larger number of markers available with the CHB panel improves the accuracy for comparison with the other source populations included in the analysis. Use of the NAM panel throughout would not have changed any of the conclusions in the paper.

It should be kept in mind that all ancestry estimates are comparative and can be dialed up or down depending on which individuals are used in the comparison, the number of markers and the information content of the markers (when the number is small). There are no absolute ancestry proportions, except in simulation. We all pointedly use the term 'pseudoancestors' in recognition that we do not know what the genotypes of the original ancestors were. All estimates using current populations as source populations have an unknown degree of error relative to the true allelic composition. Further, Amerindians, unless from very isolated populations, tend to underestimate their degree of European Ancestry. That being said, most Amerindian populations clearly cluster more closely to each other than to East Asian populations, but they also show East–West differentiation and greater intragroup differentiation than other continental populations.[3]

Moura et al refer to results from populations in Columbia, Peru and Puerto Rico as 'the literature' for American populations, but this ignores studies performed in Brazil itself, to which our results are more similar.[4–6] The one Brazilian study cited was that of Pena et al.[7] We noted in our text that, in contrast to our own sample, this paper uses a sample unlikely to be representative of Ceará state or Fortaleza and includes no pseudoancestors of any kind. This is known to distort ancestry in admixed populations.[8] We mimicked this result (Table 2, no or little difference between categories) in our sample by excluding all pseudoancestors.

Finally, we would like to take this opportunity to emphasize that this discussion is important for interpretation of genetic studies where genetic ancestry estimates are numerical guides and not gold standards, but it is irrelevant with respect to most social issues. There is a need for better Amerindian open-access data sets. They should have similar coverage to the other Hapmap populations, should include South Americans and should have little admixture with other continental groups. These, however, are not easy populations to collect.

CONFLICT OF INTEREST

The authors declare no conflict of interest.

Thiago Magalhães da Silva[1], MR Sandhya Rani[2],
Gustavo Nunes de Oliveira Costa[1], Maurício L Barreto[1,3] and
Ronald E Blanton*[,2]
[1]*Federal University of Bahia, Institute for Collective Health, Salvador,
Brazil;*
[2]*Case Western Reserve University, Center for Global Health, Cleveland,
OH, USA;*
[3]*Oswaldo Cruz Foundation, Gonçalo Moniz Research Center, Salvador,
Brazil*
E-mail: reb6@case.edu

1 de Moura RR, de Queiroz Balbino V, Crovella S, Brandão LAC: On the use of Chinese population as a proxy of Amerindian ancestors in genetic admixture studies with Latin American populations. *Eur J Hum Genet* 2015 (this issue).

2 Tsai HJ, Choudhry S, Naqvi M, Rodriguez-Cintron W, Burchard EG, Ziv E: Comparison of three methods to estimate genetic ancestry and control for stratification in genetic association studies among admixed populations. *Hum Genet* 2005; **118**: 424–433.

3 Wang S, Lewis CM, Jakobsson M et al: Genetic variation and population structure in native Americans. *PLoS Genet* 2007; **3**: e185.

4 Manta FS, Pereira R, Caiafa A, Silva DA, Gusmao L, Carvalho EF: Analysis of genetic ancestry in the admixed Brazilian population from Rio de Janeiro using 46 autosomal ancestry-informative markers. *Ann Hum Biol* 2013; **40**: 94–98.

5 Saloum de Neves Manta F, Pereira R, Vianna R et al: Revisiting the genetic ancestry of Brazilians using autosomal AIM-Indels. *PLoS One* 2013; **8**: e75145.

6 Santos NP, Ribeiro-Rodrigues EM, Ribeiro-Dos-Santos AK et al: Assessing individual interethnic admixture and population substructure using a 48-insertion-deletion (INSEL) ancestry-informative marker (AIM) panel. *Hum Mutat* 2009; **31**: 184–190.

7 Pena SD, Di Pietro G, Fuchshuber-Moraes M et al: The genomic ancestry of individuals from different geographical regions of Brazil is more uniform than expected. *PLoS One* 2011; **6**: e17063.

8 Tang H, Peng J, Wang P, Risch NJ: Estimation of individual admixture: analytical and study design considerations. *Genet Epidemiol* 2005; **28**: 289–301.