The genetic structure and adaptation of Andean highlanders and Amazonian dwellers is influenced by the interplay between geography and culture

Victor Borda[1,2,3†], Isabela Alvim[1,†,] Marla M Aquino[1,†], Carolina Silva[1], Giordano B Soares-Souza[1], Thiago P Leal[1], Marilia O Scliar[1,4], Roxana Zamudio[1], Camila Zolini[1,5,6], Carlos Padilla[3], Omar Cáceres[3,7], Kelly Levano[3], Cesar Sanchez[3], Omar Trujillo[8], Pedro O. Flores-Villanueva[3], Michael Dean[9], Silvia Fuselli[10], Moara Machado[1,9], Pedro E. Romero[11], Francesca Tassi[10], Meredith Yeager[9], Timothy D. O'Connor [12,13,14], Robert H Gilman [11,15], Eduardo Tarazona-Santos[1,16,‡], Heinner Guio[3,17,‡]

[1] Departamento de Genética, Ecologia e Evolução, Instituto de Ciências Biológicas, Universidade Federal de Minas Gerais, Belo Horizonte, MG, 31270-901, Brazil.

[2] Laboratorio de Bioinformática, LABINFO. Laboratório Nacional de Computação Científica (LNCC). Petrópolis, Rio de Janeiro - Brazil

[3] Laboratorio de Biotecnología y Biología Molecular, Instituto Nacional de Salud, Lima, Perú.

[4] Human Genome and Stem Cell Research Center, Biosciences Institute, University of São Paulo, 05508-090, SP, Brazil.

[5] Beagle, Belo Horizonte, MG, 31270-901, Brazil.

[6] Mosaico Translational Genomics Initiative, Universidade Federal de Minas Gerais, Belo Horizonte, MG, 31270-901, Brazil.

[7] Carrera de Medicina Humana, Facultad de Ciencias de la Salud, Universidad Científica del Sur, Lima, Perú

[8] Centro Nacional de Salud Intercultural, Instituto Nacional de Salud, Lima, Perú

[9] Division of Cancer Epidemiology and Genetics, National Cancer Institute (NCI), National Institutes of Health (NIH), Bethesda, MD, 20892, USA.

[10] Department of Life Sciences and Biotechnology, University of Ferrara, Via L. Borsari 46, Ferrara, 44121, Italy.

[11] Universidad Peruana Cayetano Heredia, Lima, Perú.

[12] Institute for Genome Sciences, University of Maryland School of Medicine, Baltimore, MD

[13] Program for Personalized and Genomic Medicine, University of Maryland School of Medicine, Baltimore, MD

[14] Department of Medicine, University of Maryland School of Medicine, Baltimore, MD

[15] Dept of International Health, Johns Hopkins School of Public Health Baltimore, MD, 21205, USA.

[16] Instituto de Estudos Avançados Transdisciplinares, Universidade Federal de Minas Gerais, Belo Horizonte, MG, 31270-901, Brazil.

[17] Universidad Científica del Sur, Lima, Peru

† These authors equally contributed as First Authors,
‡ These authors equally contributed as Senior Authors.

Corresponding authors: Eduardo Tarazona-Santos (edutars@gmail.com), Heinner Guio (heinnerguio@gmail.com)

Keywords: Native Americans, human population genetics, natural selection

Abstract

Western South America was one of the worldwide cradles of civilization. The well known Inca Empire was the *tip of the iceberg* of a cultural and biological evolutionary process that started 14-11 thousand years ago. Genetic data from 18 Peruvian populations reveal that: (1) The between-population homogenization of the central-southern Andes and its differentiation with respect to Amazonian populations of similar latitudes do not extend northward. Instead, longitudinal gene flow between the northern coast of Peru, Andes and Amazonia accompanied cultural and socioeconomic interactions revealed by archeological studies. This pattern recapitulates the environmental and cultural differentiation between the fertile north, where altitudes are lower; and the arid south, where the Andes are higher, acting as a genetic barrier between the sharply different environments of the Andes and Amazonia (2). The genetic homogenization between the populations of the arid Andes is not only due to migration during the Inca Empire or the subsequent colonial period. It started at least during the earlier expansion of the pre-Inca Wari Empire (600-1000 YBP) (3) This demographic history allowed for cases of positive natural selection in the high and arid Andes vs. the low Amazon tropical forest: in the Andes, *HAND2-AS1* (heart and neural crest derivatives expressed 2 antisense RNA1, related with cardiovascular function) and *DUOX2* (dual oxidase 2, related to thyroid function and innate immunity) genes; in the Amazon, the gene encoding for the CD45 protein, essential for antigen recognition by T/B lymphocytes in viral-host interaction, consistent with the *host-virus arms race* hypothesis.

Main text

Western South America was one of the cradles of civilization in the Americas and the world [1]. When the Spaniard conqueror Francisco Pizarro arrived in 1532, the pan-Andean Inca Empire ruled in the Andean region and had achieved levels of socioeconomic development and population density unmatched in other parts of South America. However, the Inca Empire, which lasted for around 200 years, with its emblematic architecture such as Machu Picchu and the city of Cuzco, was just the *tip of the iceberg* of a millenary cultural and biological evolutionary process [2,3]. This process started with the peopling of the region (hereafter called *western South America*), that occurred 14–11 thousand years ago [4–6], involving the entire Andean region and its adjacent and narrow Pacific Coast.

Tarazona-Santos et al. [7] proposed that cultural exchanges and gene flow along time have led to a relative genetic, cultural, and linguistic homogeneity between the populations of western South America compared with those of eastern South America (a term that hereafter refers to the region adjacent to the eastern slope of the Andes and eastward, including the Amazonia), where populations remained more isolated from each other. For instance, only two languages (Quechua and Aymara) of the Quechumaram linguistic stock predominate in the entire Andean region, whereas in eastern South America natives speak a different and broader spectrum of languages classified into at least four linguistic families [3,7,8]. This spatial pattern of genetic diversity and its correlation with geography, environmental, linguistic and cultural diversity was confirmed, enriched and rediscussed by us and others [2,3,7–13].

There are pending issues: First, whether the dichotomic organization of genetic variation characterized by the between-population homogeneous Southern Andes vs. between-population heterogeneous Central Amazon, extends northward. This is important because scholars from different disciplines emphasize that western South America is not latitudinally homogeneous, differentiating a northern and in general lower and wetter fertile Andes and a southern, higher and more arid Andes [14]. These environmental and latitudinal differences are correlated with demography and culture, including different spectra of domesticated plants and animals. Indeed, the development of agriculture, of the first urban centers such as Caral [1] and its associated demographic growth, occurred earlier in the northern Fertile Andes (around 5ky ago) than in the southern arid Andes (and their associated Coast), with products such as cotton, beans, and corn domesticated in the fertile north and the potato and South American camelids in the arid south [14]. In human population

genetics studies, the region where the between-population homogeneity was ascertained by Tarazona-Santos et al.[7] was the arid Andes. Consequently, here we test **(i) whether the between-population homogenization of Western South America, and the dichotomy Arid Andes/Amazonia extends to the northward Fertile Andes associated regions?** To address this and the below questions, we used data from Harris et al.[3] for 74 indigenous individuals and an additional 289 unpublished individuals from 18 Peruvian populations, genotyped for ~2.5 million SNPs (Figure 1 and Table S1). We created three datasets with different SNP densities and populations, including data from [15–18] (Figure S1, Tables S2 and S3 and Supporting information-SI). Institutional Review Boards of participants institutions approved this research.

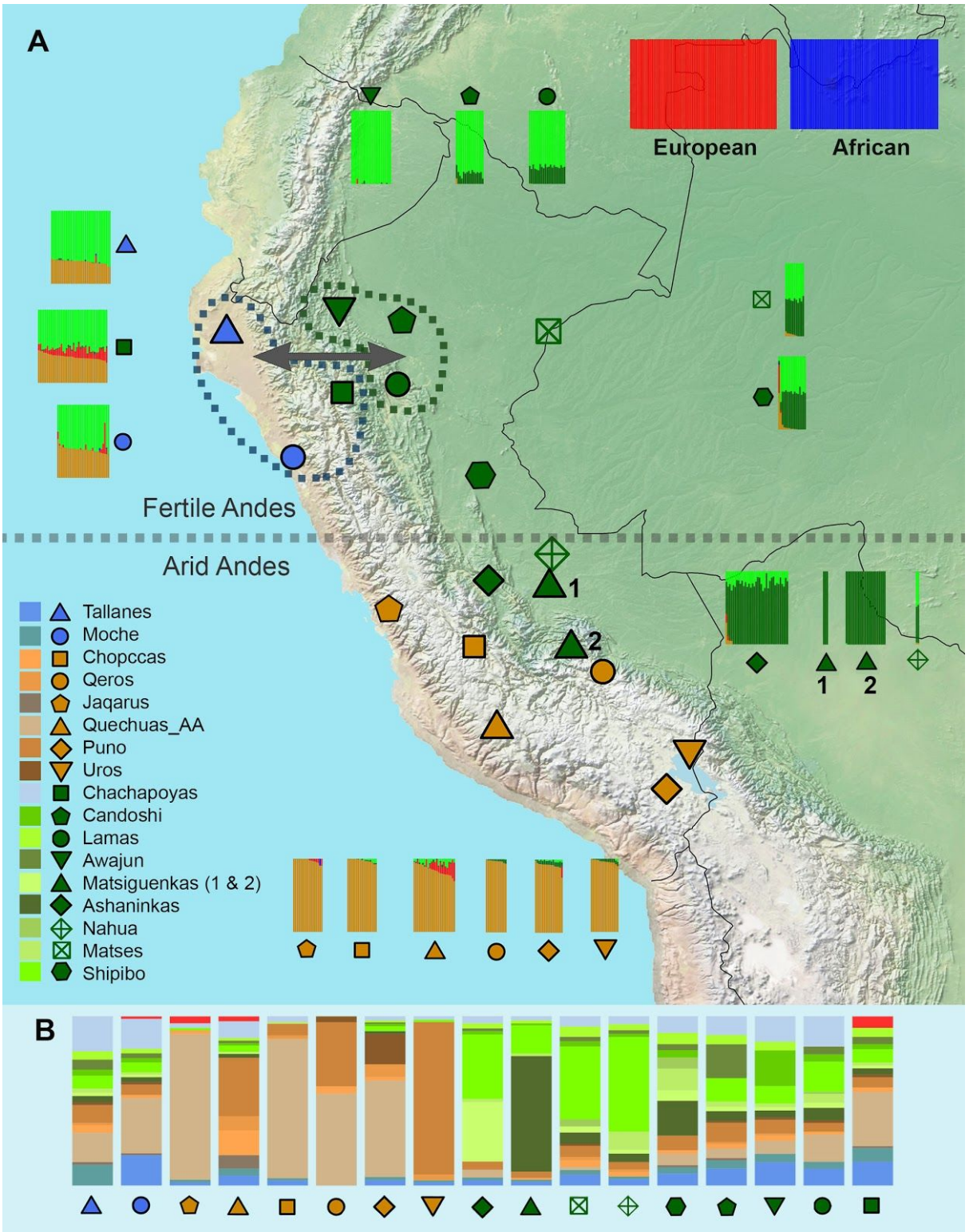

Figure 1. Genetic structure and gene flow of Western South American natives. Eighteen native groups along the Coast, Andes, and Amazon regions were sampled. A) Grey dashed line in the center shows the division between Fertile Andes and Arid Andes [14]. We showed the geographical distribution coupled with ADMIXTURE patterns for the lowest cross-validation value (K=5) using the dataset of 1.9M unlinked SNPs (including Native

Americans, Europeans, and Africans). Blue and green dashed lines delimited the groups that showed highly significant value for the gene flow test (|Z score| > 4). Matsiguenka 1= Matsiguenkas-Sepahua, Matsiguenkas 2= Matsiguenkas-Shimaa B) Haplotype based inference of ancestry profile for each Native American population, each bar corresponds to the ancestry composition for a native population. For this analysis, Matsiguenka samples were merged into one. Colors for the ancestry profile correspond to the proportion of DNA shared between the population and donor groups detailed in the legend of section A.

By applying ADMIXTURE [19] (Figures 1, S2, S4, S6) and principal component analyses (Figures S3, S5, S7), as well as haplotype-based methods [20,21] (Figures S8-S13), we confirmed that populations in the Arid Andes are genetically homogeneous, appearing as an almost panmictic unit, and with an ancestry pattern differentiated with respect to Amazonian populations of similar latitudes (Figure 1). Conversely, populations of the northern Coast (Moches), in the north Coast/Andes interface (Tallanes) and in the northern Amazon Yunga (the rainforest transitional region between the Andes and the Lower Amazonia) (Chachapoyas), share the same ancestry profile between them (Figures 1, S8-S13), which is different from the populations from the Arid Andes. Thus, the between-population homogenization of the Arid Andes and its differentiation respect to Amazonian populations of similar latitudes do not extend northward and is not characteristic of all Western South America. Instead, the genetic structure of Western South Amerindian populations recapitulates the environmental and cultural differentiation between northern Fertile Andes and the southern Arid Andes.

A second open issue is the evolutionary relationship between Andean and Amazonian populations, particularly with the culturally, linguistically, and environmentally different neighboring populations of the Amazon Yunga. Harris et al. [3] inferred that Andean and Amazonian populations diverged around 12,000 years ago. Archaeological findings of recent decades have rejected the traditional view of the Amazonian environment as incompatible with complex pre-Columbian societies, and have revealed that the Amazonian basin has produced the earliest ceramics of South America, that endogenous agricultural complex societies have developed there, and that population sizes were larger than previously thought [22]. Population genetics studies by Barbieri et al. [23] have reported episodes of gene flow in the Amazonia, which suggest that Amazonian populations were not necessarily isolated groups. Moreover, people living in the Peruvian Coast, the Andes, and the Amazon Yunga had cultural and commercial interactions during the last millennia, sharing practices

such as sweet potato and manioc cultivation, ceramic iconography and styles (e.g. Tutishcanyo, Kotosh, Valdivia and Corrugate) and traditional coca chewing [24]. Therefore, here we address (ii) **whether gene flow accompanied the cultural and socioeconomic interactions between Andean and Amazon Yunga populations**?

Haplotype-based inferences [20,21] (Figures 1 and S11-S13) and statistical tests of treeness [25] (Figures 1 and S14-S17) show genetic signatures of gene flow between Coastal/Andean and Amazon Yunga populations in latitudes of the northern fertile Andes, but not in the southern arid Andes. Thus, longitudinal gene flow between the North Coast, Andes, and Amazonia accompanied the well documented cultural and socioeconomic interactions, recapitulating the differentiation between the fertile north, where altitudes are lower; and the arid south, where the Andes altitudes are higher and may have acted as a barrier to gene flow, imposing a sharper environmental differentiation between the Andes and the Amazon Yunga. Formal tests show that in latitudes of the fertile north, gene flow includes important ethnic groups such as the current Chachapoyas of the Amazon Yunga, as well as eastward Lower Amazonian populations such as of the Jivaro linguistic family (Awajun and Candoshi) and Lamas (Figure 1 and S18-S21).

Despite some controversy about definitions and chronology, archeologists identify a unique cultural process in Western South America, which include three temporal *Horizons*: Early, Middle, and Late, that corresponds to periods of cultural dispersion involving a wide geographic area [26] (Figure 2). In particular, the Middle and Late Horizons are associated with the expansions of the Wari (~1400 to 1000 YBP) and Inca (~524 to 466 YBP) states, respectively [27–29]. Isbell [28] has suggested that the Wari expansion has been associated with the spread of the Quechua language in the Central Andes and the Wari were pioneers in developing a road system in the Andes, called *Wari ñam*, which was later used as a base by the Incas to develop their network roads (the *Qapaq ñam*) (Lumbreras et al. 2015). The between-population homogeneity currently observed in the arid Andes implies high levels of gene flow in this region, which are commonly associated with the Inca Empire [26]. A relevant question is: **(iii) when this between-population genetic homogenization started in the context of the arid Andean chronology.** Particularly, is this is a phenomenon restricted to the time of the Inca Empire or did it extend backward to the Middle/Wari Horizon?

We analyzed the distribution of IBD-segment lengths between individuals of different arid Andean populations, which is informative for the dynamics of past gene flow [3,30], and

observed a signature of gene flow in the interval 1400 to 1000 YBP, that is within the Wari expansion in the Middle Horizon (Figure 2). Thus, the homogenization of the Central Arid Andes is not only due to migrations during the Inca Empire or later during the Spanish Viceroyalty of Peru, when migrations (often forced) occurred [31]. The Wari expansion (1400 to 1000 YBP) was also accompanied by intensive gene flow whose signature is still present in the between-population genetic homogeneity of the arid central Andes region. Because IBD analysis on current individuals does not allow for inferences of gene flow that occurred more than 75 generations ago [30], ancient DNA analysis at the population level will be necessary to infer if the between-population homogenization of the Andes started even earlier.

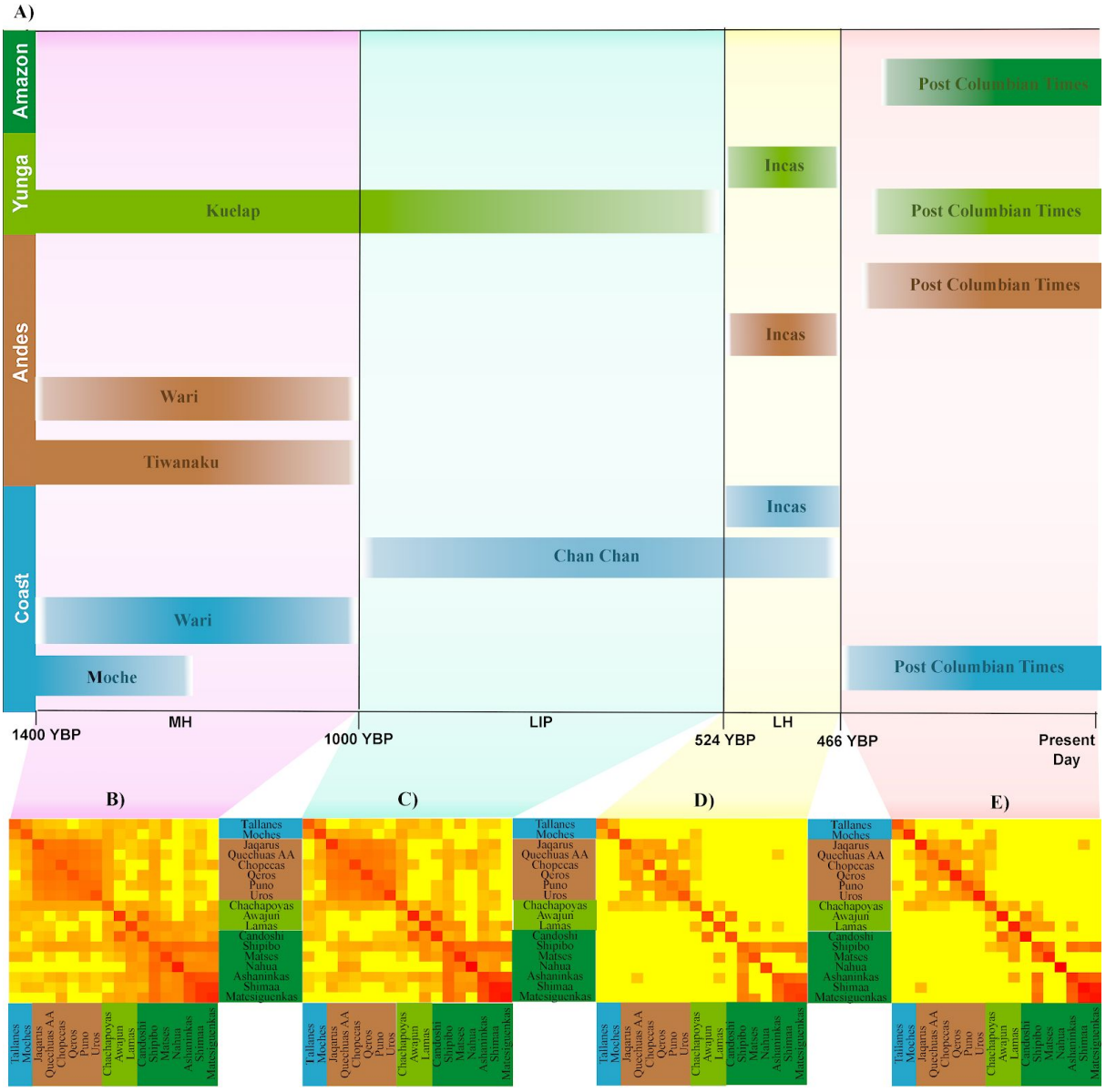

Figure 2. Evolution of IBD sharing between the Pacific Coast, Central Andes, Amazon Yunga and Amazon and its relationship with archaeological chronology of the Andes. A) Figure 2 of Scliar et al.[2] adapted, showing key historical events (cultures and

archeological sites) of Peruvian history in each region. YBP: Years Before Present, LH: Late Horizon, LIP: Late Intermediate Period, MH: Middle Horizon. B), C) and D) are heat-maps of the average pairwise relatedness (Baharian et al 2016) among Native Americans of the Natives 1.9M dataset. Each heatmap represents an interval of IBD segment lengths, which correspond to interval times (Palamara et al. 2012). B) Interval from 3.2-4.2 cM corresponding to 50-36 generations ago. C)Interval from 4.2-7.8 cM corresponding to 36 to 19 generations ago. D) Interval 7.8-9.3 cM corresponding to 19-16 generations ago. E) The last interval for all segments longer than 9.3 cM corresponds to 16-0 generations ago.

Native Americans had to adapt to different and contrasting environments and stress. The high and arid Andes is characterized by high UV radiation, cold, dryness, and hypoxia (a stress that does not allow for cultural adaptations and requires biological changes) [32,33]. The Amazon has a low incidence of light, a warm and humid climate typical of the rainforest and high biodiversity, including human pathogens [34]. Populations from the high and arid Andes and from the Amazon (Figure 1) settled in these contrasting environments more than 5000 years ago [35] and show little evidence of gene flow between them (i.e. that would homogenize allele frequencies, potentially concealing the effect of diversifying natural selection). Thus, we performed genome-wide scans in these two groups of populations using two tests of positive natural selection: (i) Population Branch Statistics (PBSn) comparing arid Andeans (Chopccas, Quechuas, Qeros, Puno, Jaqarus, and Uros) vs. Amazonian populations (Ashaninkas, Matsiguenkas, Matses and Nahua) with CDX population (Chinese Dai in Xishuangbanna, China) from 1000 Genomes as an outgroup [36] and (ii) long-range haplotypes (xpEHH)[37] estimated with the same two groups of populations (Supplementary text, Section 4, Figure 3, S25-S27). The complete lists of SNPs with high PBSn and xpEHH statistics for Andean and Amazonian populations are in Tables S4-S7.

The gene with the consensually strongest signal of adaptation (both from PBSn and xpEHH statistics) to the Andean environment (Figure 3, Table S4) is the long non-coding RNA gene *HAND2-AS1* (heart and neural crest derivatives expressed 2 RNA antisense 1, chromosome 4), that modulates cardiogenesis by regulating the expression of the HAND2 gene [38,39]. HAND2-AS1 is located in the antisense 5' region of HAND2 and contains 2 enhancers for this gene. A natural selection genome-wide scan [36] identified three genes related to the cardiovascular system in Andeans, including TBX5, which works together with HAND2 in reprogramming fibroblasts to cardiac-like myocytes [40,41]. This information suggests (but does not demonstrate) that HAND2-AS1 signature of natural selection is related with

cardiovascular adaptations. Andeans have cardiovascular adaptations to high altitude that differ from those of lowlanders exposed to hypoxia and from other highlanders, showing higher pulmonary vasoconstrictor response to hypoxia and lower resting middle cerebral flow velocity than Tibetans, and higher uterine artery blood flow than Europeans raised in high altitude and than lowlanders [42].

DUOX2 (dual oxidase 2, chromosome 15) is the gene with the highest signal of adaptation to the Andean environment by PBSn analysis (Figure 3). It was reported as a natural selection target in the Andes by [43,44]. *DUOX2* encodes a trans-membrane component of an NADPH oxidase, which produces hydrogen peroxide (H_2O_2), and is essential for the synthesis of the thyroid hormone and for the production of microbicidal hypothiocyanite anion (OSCN–) during mucosal innate immunity response against bacterial and viral infections in the airways and intestines [45,46]. Mutations in *DUOX2* produce inherited hypothyroidism [47]. Here we report for the first time that: (i) The PBS signal for *DUOX2* comprises several SNPs, including 2 missense mutations (rs269868: C>T:Ser1067Leu, C allele frequencies: Amazon: 0.01, Andean: 0.53; rs57659670: T>C:His678Arg, C allele frequencies: Amazon: 0.01, Andean: 0.53), thus the ancestral allele has been positively selected in the Andes; (ii) bioinformatics analysis reveals that rs269868 is located in an A-loop aa1064-1078, which is a region of interaction of *DUOX2* with its coactivator *DUOXA2*. Mutations in this region of the protein can affect the stability and maturation of the dimer and, consequently, the conversion of the intermediate product $O_2^{\square}$ to the final product H_2O_2 and their released proportions [48]. It is not clear if the *DUOX2* natural selection signal is related to thyroid function or innate immunity. Before the introduction of the public health program of supplementing manufactured salt with iodine, one of the environmental stresses of the Andes for human populations was iodine deficiency, which impairs thyroid hormone synthesis, increasing the risk of developing hypothyroidism, goiter, obstetric complications and cognitive impairment [49,50].

Natural selection studies in Amazon populations are scarce. Studies targeting rainforest populations in Africa and Asia have found natural selection signals in genes related to height and immune response [51]. In the Amazon region, the strongest natural selection PBSn signal is in a long non-coding RNA gene on chromosome 18 with unknown function (Table S5). The second highest signal (that also shows a significant long-range haplotype signal) comes from the region around the gene *PTPRC*, which encodes the protein CD45, essential in antigen recognition by T and B lymphocytes, particularly in pathogen-host interaction, in particular for virus such as Human adenovirus type 19 [52], HIV-1-induced cell apoptosis [53,54]

and susceptibility to Hepatitis C [55,56] and herpes simplex virus 1 [57]. Interestingly, HSV-1 this herpes virus has a high incidence in isolated Amerindians from the Peruvian and Brazilian Amazon [58–60], with the elevated diversity of the virus and an endemic subtype that suggest an ancient endemic infection [61]. This result is consistent with the hypothesis of CD45 evolution driven by the host-virus arms race model [62].

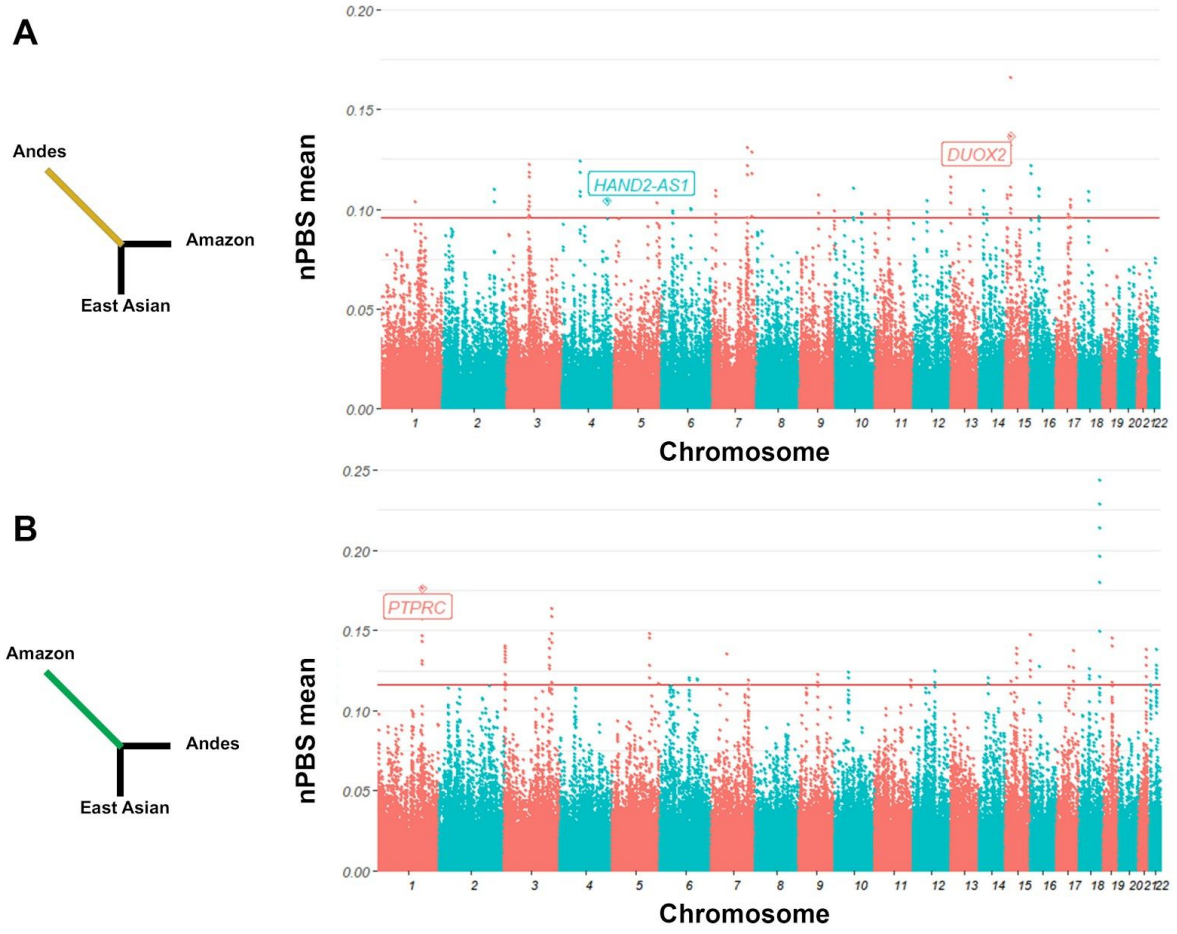

Figure 3. Manhattan plot of the Population Branch Specific (PBSn) estimated as a sliding windows of 20 SNPs with 5 SNPs of overlap in Andean (A) and Amazon (B) populations. We tagged PBSn hits that correspond to genes and show a significant xpEHH statistics. The horizontal red line shows 95.95th percentile of PBSn values.

In conclusion, in Western South America, there is an environmental and cultural differentiation between the fertile north of the Andes, where altitudes are lower; and the arid south of the Andes, where these mountains are higher, defining sharp environmental differences between the Andes and Amazonia. This has influenced the genetic structure of

Western South Amerindian populations. Indeed, the between-population homogenization of the central-southern Andes and its differentiation with respect to Amazonian populations of similar latitudes do not extend northward. Longitudinal gene flow between the northern Coast of Peru, Andes, and Amazonia accompanied cultural and socioeconomic interactions revealed by archeology, but in the central-southern Andes, these mountains have acted as a genetic barrier to gene flow. We provide new insights on the dynamics of the genetic homogenization between the populations of the Arid Andes, which are not only due to migrations during the Inca Empire or the subsequent colonial period, but started at least during the earlier expansion of the pre-Inca Wari Empire (600-1000 YBP). This evolutionary journey of Western South Amerindians was accompanied by episodes of adaptive natural selection to the high and arid Andes vs. the low Amazon tropical forest: *HAND2-AS1* (related with cardiovascular function) and *DUOX2* (related to thyroid function and innate immunity) in the Andes. In the Amazon forest environment, the gene encoding for the protein CD45, essential for antigen recognition by T/B lymphocytes and viral-host interactions, shows a signature of positive natural selection, consistent with the host-virus arms race hypothesis. This and other studies, continue to show how Andean highlanders and Amazonian dwellers provide examples of how the interplay between geography and culture influence the genetic structure and adaptation of human populations.

Acknowledgments

We thank the Peruvian populations for their participation. We thank the members of the Laboratório de Diversidad Genética Humana (LDGH), Mateus Gouveia, Kelly Nunes, Garrett Hellenthal, Mark Lipson, Marcia Beltrame, Fabrício Santos, Claudio Struchiner, Ricardo Santos, Luis Guillermo Lumbreras, Sandra Romero-Hidalgo, Víctor Acuña-Alonzo, and Miguel Ortega for discussions or technical assistance, the staff of the Laboratorio de Biotecnologia y Biologia Molecular of Instituto Nacional de Salud (Peru) for collaborating with the Peruvian Genome Project and conducting the genotyping. Funding support: This work was supported by the Peruvian National Institute of Health (INS), the Brazilian Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) and the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES, programs PROEX and PRINT). V.B. is a CAPES/PEC-PG fellow (88882.195664/2018-01). Data sets were processed in sagarana HPC cluster from the Centro de Laboratórios Multiusuários from the ICB-UFMG.

References

1. Solis, R.S., Haas, J., and Creamer, W. (2001). Dating Caral, a preceramic site in the Supe Valley on the central coast of Peru. Science *292*, 723–726.

2. Scliar, M.O., Gouveia, M.H., Benazzo, A., Ghirotto, S., Fagundes, N.J.R., Leal, T.P., Magalhães, W.C.S., Pereira, L., Rodrigues, M.R., Soares-Souza, G.B., et al. (2014). Bayesian inferences suggest that Amazon Yunga Natives diverged from Andeans less than 5000 ybp: implications for South American prehistory. BMC Evol. Biol. *14*, 174.

3. Harris, D.N., Song, W., Shetty, A.C., Levano, K.S., Cáceres, O., Padilla, C., Borda, V., Tarazona, D., Trujillo, O., Sanchez, C., et al. (2018). Evolutionary genomic dynamics of Peruvians before, during, and after the Inca Empire. Proc. Natl. Acad. Sci. U. S. A. 201720798.

4. Lahaye, C., Guérin, G., Boëda, E., Fontugne, M., Hatté, C., Frouin, M., Clemente-Conte, I., Pino, M., Felice, G.D., Guidon, N., et al. (2015). New insights into a late-Pleistocene human occupation in America: The Vale da Pedra Furada complete chronological study. Quat. Geochronol. *30*, 445–451.

5. Dillehay, T.D., Ramírez, C., Pino, M., Collins, M.B., Rossen, J., and Pino-Navarro, J.D. (2008). Monte Verde: seaweed, food, medicine, and the peopling of South America. Science *320*, 784–786.

6. Dillehay, T.D., Ocampo, C., Saavedra, J., Sawakuchi, A.O., Vega, R.M., Pino, M., Collins, M.B., Scott Cummings, L., Arregui, I., Villagran, X.S., et al. (2015). New Archaeological Evidence for an Early Human Presence at Monte Verde, Chile. PLoS One *10*, e0141923.

7. Tarazona-Santos, E., Carvalho-Silva, D.R., Pettener, D., Luiselli, D., De Stefano, G.F., Labarga, C.M., Rickards, O., Tyler-Smith, C., Pena, S.D., and Santos, F.R. (2001). Genetic differentiation in South Amerindians is related to environmental and cultural diversity: evidence from the Y chromosome. Am. J. Hum. Genet. *68*, 1485–1496.

8. Campbell, L. (2000). American Indian Languages: The Historical Linguistics of Native America (Oxford University Press).

9. Fuselli, S., Tarazona-Santos, E., Dupanloup, I., Soto, A., Luiselli, D., and Pettener, D. (2003). Mitochondrial DNA diversity in South America and the genetic history of Andean highlanders. Mol. Biol. Evol. *20*, 1682–1691.

10. Lewis, C.M., and Long, J.C. (2008). Native South American genetic structure and prehistory inferred from hierarchical modeling of mtDNA. Mol. Biol. Evol.

11. Wang, S., Lewis, C.M., Jakobsson, M., Ramachandran, S., Ray, N., Bedoya, G., Rojas, W., Parra, M.V., Molina, J.A., Gallo, C., et al. (2007). Genetic variation and population structure in native Americans. PLoS Genet. *3*, e185.

12. Sandoval, J.R., Lacerda, D.R., and Jota, M.S.A. (2013). The genetic history of indigenous populations of the Peruvian and Bolivian Altiplano: the legacy of the Uros. PLoS.

13. Gnecchi-Ruscone, G.A., Sarno, S., De Fanti, S., Gianvincenzo, L., Giuliani, C., Boattini, A., Bortolini, E., Di Corcia, T., Sanchez Mellado, C., Dàvila Francia, T.J., et al. (2019). Dissecting the Pre-Columbian Genomic Ancestry of Native Americans along the

Andes-Amazonia Divide. Mol. Biol. Evol. *36*, 1254–1269.

14. Lumbreras, L.G. (2015). Los orígenes de la civilización en el Perú (IAEAS, Instituto Andino de Estudios Arqueológico-Sociales).

15. 1000 Genomes Project Consortium, Abecasis, G.R., Auton, A., Brooks, L.D., DePristo, M.A., Durbin, R.M., Handsaker, R.E., Kang, H.M., Marth, G.T., and McVean, G.A. (2012). An integrated map of genetic variation from 1,092 human genomes. Nature *491*, 56–65.

16. Reich, D., Patterson, N., Campbell, D., Tandon, A., Mazieres, S., Ray, N., Parra, M.V., Rojas, W., Duque, C., Mesa, N., et al. (2012). Reconstructing Native American population history. Nature *488*, 370–374.

17. Raghavan, M., Steinrücken, M., Harris, K., Schiffels, S., Rasmussen, S., DeGiorgio, M., Albrechtsen, A., Valdiosera, C., Ávila-Arcos, M.C., Malaspinas, A.-S., et al. (2015). POPULATION GENETICS. Genomic evidence for the Pleistocene and recent population history of Native Americans. Science *349*, aab3884.

18. Mallick, S., Li, H., Lipson, M., Mathieson, I., Gymrek, M., Racimo, F., Zhao, M., Chennagiri, N., Nordenfelt, S., Tandon, A., et al. (2016). The Simons Genome Diversity Project: 300 genomes from 142 diverse populations. Nature *538*, 201–206.

19. Alexander, D.H., Novembre, J., and Lange, K. (2009). Fast model-based estimation of ancestry in unrelated individuals. Genome Res. *19*, 1655–1664.

20. Lawson, D.J., Hellenthal, G., Myers, S., and Falush, D. (2012). Inference of population structure using dense haplotype data. PLoS Genet. *8*, e1002453.

21. Hellenthal, G., Busby, G.B.J., Band, G., Wilson, J.F., Capelli, C., Falush, D., and Myers, S. (2014). A genetic atlas of human admixture history. Science *343*, 747–751.

22. Roosevelt, A. (1999). The Maritime, Highland, Forest Dynamic and the Origins Of Complex Culture. In The Cambridge History of the Native Peoples of the Americas, (Cambridge University Press), pp. 264–349.

23. Barbieri, C., Barquera, R., Arias, L., Sandoval, J.R., Acosta, O., Zurita, C., Aguilar-Campos, A., Tito-Álvarez, A.M., Serrano-Osuna, R., Gray, R., et al. (2019). The current genomic landscape of western South America: Andes, Amazonia and Pacific Coast. Mol. Biol. Evol.

24. Silverman, H., and Isbell, W. (2008). Handbook of South American Archaeology (Springer Science & Business Media).

25. Patterson, N., Moorjani, P., Luo, Y., Mallick, S., Rohland, N., Zhan, Y., Genschoreck, T., Webster, T., and Reich, D. (2012). Ancient admixture in human history. Genetics *192*, 1065–1093.

26. Haas, J., Pozorski, S., and Pozorski, T. (1987). The Origins and Development of the Andean State (Cambridge University Press).

27. Lanning, E.P. (1967). Peru before the Incas.

28. Isbell, W.H. (2008). Wari and Tiwanaku: International Identities in the Central Andean Middle Horizon. In The Handbook of South American Archaeology, H. Silverman, and W.H.

Isbell, eds. (New York, NY: Springer New York), pp. 731–759.

29. Valverde, G., Barreto Romero, M.I., Flores Espinoza, I., Cooper, A., Fehren-Schmitz, L., Llamas, B., and Haak, W. (2016). Ancient DNA Analysis Suggests Negligible Impact of the Wari Empire Expansion in Peru's Central Coast during the Middle Horizon. PLoS One *11*, e0155508.

30. Palamara, P.F., Lencz, T., Darvasi, A., and Pe'er, I. (2012). Length distributions of identity by descent reveal fine-scale demographic history. Am. J. Hum. Genet. *91*, 809–822.

31. Cook, N.D. (1990). Migration in colonial Peru: an overview. Migration in Colonial Spanish America 41–61.

32. Tarazona-Santos, E., Lavine, M., Pastor, S., Fiori, G., and Pettener, D. (2000). Hematological and pulmonary responses to high altitude in Quechuas: a multivariate approach. American Journal of Physical Anthropology: The Official Publication of the American Association of Physical Anthropologists *111*, 165–176.

33. Moore, L.G. (2017). Measuring high-altitude adaptation. J. Appl. Physiol. *123*, 1371–1385.

34. Amorim, C.E.G., Daub, J.T., Salzano, F.M., Foll, M., and Excoffier, L. (2015). Detection of convergent genome-wide signals of adaptation to tropical forests in humans. PLoS One *10*, e0121557.

35. Eriksen, L. (2011). Nature and Culture in Prehistoric Amazonia : Using G.I.S. to reconstruct ancient ethnogenetic processes from archaeology, linguistics, geography, and ethnohistory. Lund University.

36. Crawford, J.E., Amaru, R., Song, J., Julian, C.G., Racimo, F., Cheng, J.Y., Guo, X., Yao, J., Ambale-Venkatesh, B., Lima, J.A., et al. (2017). Natural Selection on Genes Related to Cardiovascular Health in High-Altitude Adapted Andeans. Am. J. Hum. Genet. *101*, 752–767.

37. Sabeti, P.C., Varilly, P., Fry, B., Lohmueller, J., Hostetter, E., Cotsapas, C., Xie, X., Byrne, E.H., McCarroll, S.A., Gaudet, R., et al. (2007). Genome-wide detection and characterization of positive selection in human populations. Nature *449*, 913–918.

38. Anderson, K.M., Anderson, D.M., McAnally, J.R., Shelton, J.M., Bassel-Duby, R., and Olson, E.N. (2016). Transcription of the non-coding RNA upperhand controls Hand2 expression and heart development. Nature *539*, 433–436.

39. Cheng, X., and Jiang, H. (2019). Long non-coding RNA HAND2-AS1 downregulation predicts poor survival of patients with end-stage dilated cardiomyopathy. J. Int. Med. Res. *47*, 3690–3698.

40. Hashimoto, H., Wang, Z., Garry, G.A., Malladi, V.S., Botten, G.A., Ye, W., Zhou, H., Osterwalder, M., Dickel, D.E., Visel, A., et al. (2019). Cardiac Reprogramming Factors Synergistically Activate Genome-wide Cardiogenic Stage-Specific Enhancers. Cell Stem Cell *25*, 69–86.e5.

41. Fernandez-Perez, A., Sathe, A.A., Bhakta, M., Leggett, K., Xing, C., and Munshi, N.V. (2019). Hand2 Selectively Reorganizes Chromatin Accessibility to Induce Pacemaker-like

Transcriptional Reprogramming. Cell Rep. *27*, 2354–2369.e7.

42. Julian, C.G., and Moore, L.G. (2019). Human Genetic Adaptation to High Altitude: Evidence from the Andes. Genes *10*,.

43. Zhou, D., Udpa, N., Ronen, R., Stobdan, T., Liang, J., Appenzeller, O., Zhao, H.W., Yin, Y., Du, Y., Guo, L., et al. (2013). Whole-genome sequencing uncovers the genetic basis of chronic mountain sickness in Andean highlanders. Am. J. Hum. Genet. *93*, 452–462.

44. Jacovas, V.C., Couto-Silva, C.M., Nunes, K., Lemes, R.B., de Oliveira, M.Z., Salzano, F.M., Bortolini, M.C., and Hünemeier, T. (2018). Selection scan reveals three new loci related to high altitude adaptation in Native Andeans. Sci. Rep. *8*, 12733.

45. van der Vliet, A., Danyal, K., and Heppner, D.E. (2018). Dual oxidase: a novel therapeutic target in allergic disease. Br. J. Pharmacol. *175*, 1401–1418.

46. Deken, X.D., De Deken, X., Corvilain, B., Dumont, J.E., and Miot, F. (2014). Roles of DUOX-Mediated Hydrogen Peroxide in Metabolism, Host Defense, and Signaling. Antioxidants & Redox Signaling *20*, 2776–2793.

47. Maruo, Y., Nagasaki, K., Matsui, K., Mimura, Y., Mori, A., Fukami, M., and Takeuchi, Y. (2016). Natural course of congenital hypothyroidism by dual oxidase 2 mutations from the neonatal period through puberty. Eur. J. Endocrinol. *174*, 453–463.

48. Ueyama, T., Sakuma, M., Ninoyu, Y., Hamada, T., Dupuy, C., Geiszt, M., Leto, T.L., and Saito, N. (2015). The extracellular A-loop of dual oxidases affects the specificity of reactive oxygen species release. J. Biol. Chem. *290*, 6495–6506.

49. Pretell, E.A., Pearce, E.N., Moreno, S.A., Dary, O., Kupka, R., Gizak, M., Gorstein, J., Grajeda, R., and Zimmermann, M.B. (2017). Elimination of iodine deficiency disorders from the Americas: a public health triumph. Lancet Diabetes Endocrinol *5*, 412–414.

50. Pan, L., Fu, Z., Yin, P., and Chen, D. (2019). Pre-existing medical disorders as risk factors for preeclampsia: an exploratory case-control study. Hypertens. Pregnancy *38*, 245–251.

51. Fan, S., Hansen, M.E.B., Lo, Y., and Tishkoff, S.A. (2016). Going global by adapting local: A review of recent human adaptation. Science *354*, 54–59.

52. Windheim, M., Southcombe, J.H., Kremmer, E., Chaplin, L., Urlaub, D., Falk, C.S., Claus, M., Mihm, J., Braithwaite, M., Dennehy, K., et al. (2013). A unique secreted adenovirus E3 protein binds to the leukocyte common antigen CD45 and modulates leukocyte functions. Proc. Natl. Acad. Sci. U. S. A. *110*, E4884–E4893.

53. Anand, A.R., and Ganju, R.K. (2006). HIV-1 gp120-mediated apoptosis of T cells is regulated by the membrane tyrosine phosphatase CD45. J. Biol. Chem. *281*, 12289–12299.

54. Meer, S., Perner, Y., McAlpine, E.D., and Willem, P. (2019). Extraoral plasmablastic lymphomas in a high human immunodeficiency virus endemic area. Histopathology.

55. Dawes, R., Hennig, B., Irving, W., Petrova, S., Boxall, S., Ward, V., Wallace, D., Macallan, D.C., Thursz, M., Hill, A., et al. (2006). Altered CD45 expression in C77G carriers influences immune function and outcome of hepatitis C infection. J. Med. Genet. *43*, 678–684.

56. Hsiao, J.-L., Ko, W.-S., Shih, C.-J., and Chiou, Y.-L. (2017). The Changed Proportion of CD45RA+/CD45RO+ T Cells in Chronic Hepatitis C Patients During Pegylated Interferon-α with Ribavirin Therapy. J. Interferon Cytokine Res. *37*, 303–309.

57. Caignard, G., Leiva-Torres, G.A., Leney-Greene, M., Charbonneau, B., Dumaine, A., Fodil-Cornu, N., Pyzik, M., Cingolani, P., Schwartzentruber, J., Dupaul-Chicoine, J., et al. (2013). Genome-wide mouse mutagenesis reveals CD45-mediated T cell function as critical in protective immunity to HSV-1. PLoS Pathog. *9*, e1003637.

58. Cunha, A.M.G., Caterino-de-Araujo, A., Costa, S.C.B., Santos-Fortuna, E., Boa-Sorte, N.C.A., Gonçalves, M.S., Costa, F.F., and Galvão-Castro, B. (2005). Increasing seroprevalence of human herpesvirus 8 (HHV-8) with age confirms HHV-8 endemicity in Amazon Amerindians from Brazil. J. Gen. Virol. *86*, 2433–2437.

59. de Souza, V.A.U.F., Sumita, L.M., Nascimento, M.-C., Oliveira, J., Mascheretti, M., Quiroga, M., Freire, W.S., Tateno, A., Boulos, M., Mayaud, P., et al. (2007). Human herpesvirus-8 infection and oral shedding in Amerindian and non-Amerindian populations in the Brazilian Amazon region. J. Infect. Dis. *196*, 844–852.

60. Nascimento, M.C., Sumita, L.M., Souza, V.U., Weiss, H.A., Oliveira, J., Mascheretti, M., Quiroga, M., Vela, R.A.R., Sabino, E., Pannuti, C.S., et al. (2009). Seroprevalence of Kaposi sarcoma-associated herpesvirus and other serologic markers in the Brazilian Amazon. Emerg. Infect. Dis. *15*, 663–667.

61. Ishak, R., Machado, L.F.A., Cayres-Vallinoto, I., Guimarães Ishak, M. de O., and Vallinoto, A.C.R. (2017). Infectious Agents As Markers of Human Migration toward the Amazon Region of Brazil. Front. Microbiol. *8*, 1663.

62. Thiel, N., Zischke, J., Elbasani, E., Kay-Fedorov, P., and Messerle, M. (2015). Viral interference with functions of the cellular receptor tyrosine phosphatase CD45. Viruses *7*, 1540–1557.