Short Communication

Apolipoprotein E Polymorphisms in Andean Population of Jujuy, Argentina

Arturo Nicolás Trigo[a], Marina Muzzio[b], Marcelo Isidro Figueroa[a,c,*],
Emma Laura Alfaro-Gómez[c,d], Graciela Bailliet[b], Hernán Javier Dopazo[e]
and José Edgardo Dipierri[a]
[a]*Genetics Unit, Maternal and Child Hospital "Dr. Hector Quintana'', Jujuy, Argentina*
[b]*Multidisciplinary Institute of Cellular Biology (IMBICE – CONICET – UNLP – CIC), La Plata, Argentina*
[c]*Institute of Andean Ecoregions (INECOA – UNJu – CONICET), Jujuy, Argentina*
[d]*Institute of Biology of the Altitude (INBIAL – UNJu), Jujuy, Argentina*
[e]*Institute of Ecology, Genetics and Evolution of Buenos Aires (IEGEBA – UBA – CONICET), Ciudad Autónoma de Buenos Aires, Argentina*

Abstract. The pleiotropic nature of the apolipoprotein E (*APOE*) gene is associated with complex diseases in different populations. We analyzed *APOE* polymorphisms in 76 individuals from Jujuy – Argentina using NGS technology. The observed genotypes align with the expected Hardy–Weinberg equilibrium. *APOE3* was the most common allele, followed by *APOE4* and *APOE2*. The allele distribution pattern is consistent with findings in previously studied populations of Native Americans and Asians. The E4 allele's low frequency, always observed in a heterozygous state, raises questions regarding its relevance in explaining dementia and longevity associated with this marker in the Central Andes.

Keywords: Allele frequency, Alzheimer's disease, Andean population, APOE

INTRODUCTION

The *APOE* gene is linked to lipoproteins of both very low and high density, as well as chylomicrons, and it plays a crucial role in lipid metabolism [1]. There are three major isoforms, namely E2, E3, and E4, which are identified through isoelectric focusing and coded for by three alleles (epsilon 2, 3, and 4). These isoforms vary in their amino acid sequence at site 2, residue 112 (site A), and residue 158 (site B). The *APOE* alleles have shown pleiotropic

effects in various populations, being linked to a plethora of complex diseases and traits, including neurodegenerative disorders, proteinopathies, lipid metabolism, cardiovascular and cerebrovascular disease, and longevity [2–4]. Furthermore, recent studies indicate that individuals homozygous for *APOE4* face an elevated risk of severe COVID-19 infection [5]. Notwithstanding, it is important to note that not all alleles have an equal likelihood of developing the described associations. There is evidence of a gradual increase in risk and loss of protein function from *APOE2* to *APOE3* and *APOE4*, with some controversy surrounding *APOE2* having a protective effect. Furthermore, the probability of association increases in accordance with the number of *APOE4* alleles

present, ranging from 4 in heterozygotes to 12 in homozygotes [3].

APOE allele distribution is heterogeneous globally, with *APOE3* being the most common allele across all populations, followed by *APOE4*. On a continental level, Africa and Oceania present the highest E2 and E4 frequencies, while E3 is the predominant allele in Asia [6, 7]. However, there is a scarcity of studies concerning *APOE* polymorphism distribution in indigenous populations in the Andes [8–10].

Jujuy's population, located in the Andes, exhibits a high frequency of autochthonous surnames in their onomastics and an elevated Native American ancestry [12, 13]. This study analyses the distribution of *APOE* allele in an admixed population from Jujuy and compares it with other Native American groups from the Central Andes.

MATERIALS AND METHODS

In this study, exome data on 76 pediatric patients, unrelated to one another, was examined for genomic studies on various genetic diseases initially unrelated to *APOE*. The patients were invited by the Medical Genetics Unit of the Maternal and Child Public Hospital of Jujuy and each gave written consent to participate. The study has been approved by the Provincial Bioethics Committee.

Genomic DNA was extracted from whole blood (300 μl) utilizing a QIAamp DNA Blood purification kit as per QIAGEN GmbH's directions. Library preparation, exome, and neighboring intronic regions capture were conducted with Agilent's SureSelect All Exon V6 kit. Full exome sequencing by NGS was performed using an Illumina NovaSeq PE150 platform. The protocol followed for mapping, alignment, and variant calling was based on the Best Practices of the Eli and Edythe L. Broad Institute of Harvard and MIT considering GRCh37 as the reference genome [14].

The identification of alleles in each patient was carried out by utilizing two *APOE* markers, rs429358 and rs7412. These markers determine the presence of E2 alleles (rs429358 T, rs7412 T), E3 alleles (rs429358 T, rs7412 C), and E4 alleles (rs429358 C, rs7412 C). Abbreviations were explained when first used. Genotypic and allelic frequencies were computed, and Hardy-Weinberg equilibrium was evaluated via Arlequin v3.5 software [15]. Genetic differentiation among samples was appraised utilizing an exact G test from Genepop v4 software [16].

To facilitate exploratory analysis, allelic frequencies from other Native American populations were included [9, 10], and the dataset was evaluated through cluster analysis employing Euclidean distances for the E2, E3 and E4 allelic frequencies with R software [17].

RESULTS

The recorded genotypes conform to those anticipated by the Hardy-Weinberg equilibrium. Our findings are presented in Table 1, where we collated them with comparable populations from the Central Andes. Despite conducting the exact G test, we did not identify any statistically significant difference. Nevertheless, increasing the sample size could reveal such differences. *APOE3* dominated as the primary allele, with *APOE4* and *APOE2* manifesting in low frequencies. The *APOE3* allele was present in a homozygosis state with a frequency of 0.89. Figure 1 presents cluster analysis results on *APOE* allele frequencies in various Native American populations, indicating frequencies measured in this study to be comparable to those found in other Andean populations. This study also notes that the lower frequency of the E2 and E4 alleles is not unique to the Andean population.

DISCUSSION

The distribution of alleles conformed to the pattern observed in other Andean populations and South American Native Americans [8, 9]. These populations are characterized by a low frequency of *APOE4*, and the absence of the potentially most detrimental E4/E4 genotype, with the exception of Gavião (7%), Zoró (17%), Wai-Wai (7%), and Myá-Guaraní (4.5%), which are non-Andean populations.

The infrequent occurrence of E4, even in a heterozygous state, suggests that this allele could be significant in explaining the epidemiology of Alzheimer's disease, coronary disease, and longevity among the Native Andean population. In contrast, this genotype appears to be linked with these illnesses in communities of diverse ethnicities. We must consider that, in Andean populations, there have been a limited number of molecular or epidemiological studies on these conditions and their relationship with *APOE*, making the analysis more intricate [18]. According to the GBD 2016 Dementia Collaborators group (2019), the worldwide propor-

Table 1
Genotypic and allelic frequencies of *APOE3* in the population of Jujuy and its comparison with populations from the Central Andes

| Populations	N	Genotypes						Allelic frequencies			References
		E2/E2	E2/E3	E2/E4	E3/E3	E3/E4	E4/E4	E2	E3	E4	
Quechua	40	0	0	0	0.95	0.05	0	0.000	0.975	0.025	Reales et al. (2017) [27]
Aymara	39	0	0	0.02	0.90	0.08	0	0.013	0.936	0.051	Gayà-Vidal et al. (2012) [11]
Quechua	38	0	0	0	0.89	0.11	0	0.000	0.947	0.053	
Bolivia	77	0	0	0.02	0.89	0.09	0	0.007	0.942	0.052	
Peru (Ayacucho)	37	0	0.03			0.86	0.11	0	0.014	0.932	0.054	Marca et al. (2011) [28]
Peru (Junin)	21	0	0	0	0.90	0.09	0	0	0.952	0.054	
This study	76	0	0.04	00	0.88	0.08	0	0.020	0.941	0.039	–
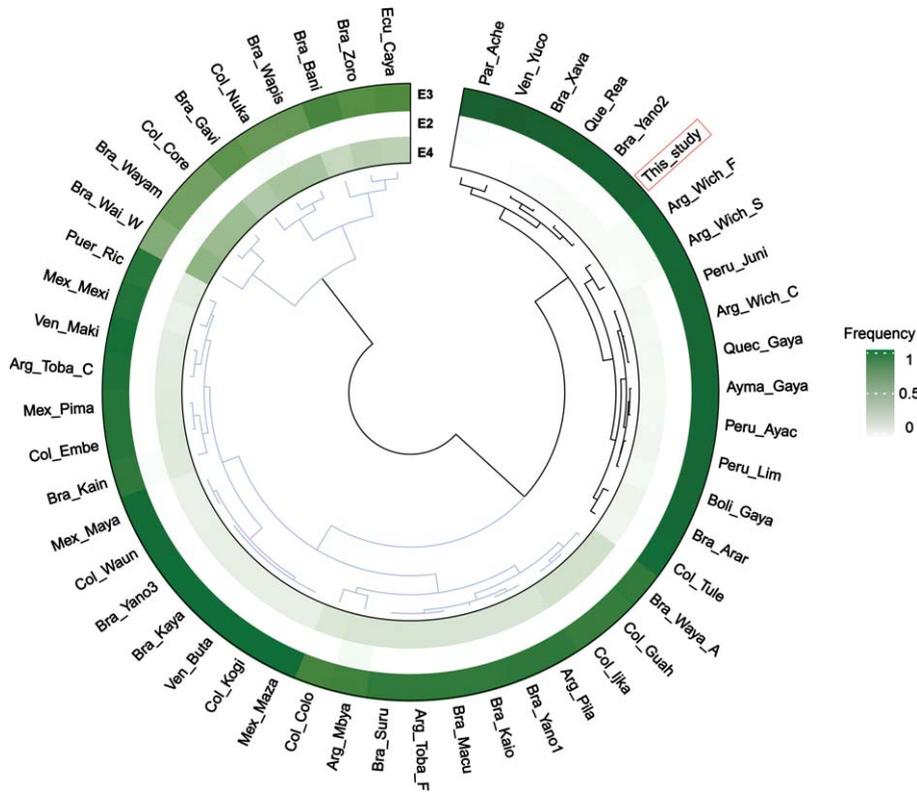
Fig. 1. Clustering analysis of *APOE* allele frequencies in different Native American populations [9]: Arg_Mbya (Argentina – Mbyá–Guarani); Bra_Kain (Brazil – Kaingang); Arg_Toba_F (Argentina – Toba Formosa); Arg_Toba_C (Argentina – Toba Chaco); Arg_Wich_C (Argentina – Wichí Chaco); Arg_Pila (Argentina – Pilagá); Arg_Wich_F (Argentina – Wichí Formosa); Par_Ache (Paraguay – Aché, Paraguay); Bra_Kaio (Brazil – Kaiowá/Ñandeva–Guarani); Arg_Wich_S (Argentina – Wichí Salta); Bra_Xava (Brazil – Xavante); Bra_Suru (Brazil – Suruí); Bra_Zoro (Brazil – Zoró); Bra_Gavi (Brazil – Gavião); Bra_Kaya (Brazil – Kayapo); Bra_Arar (Brazil – Arara); Bra_Waya_A (Brazil – Wayana–Apalai); Bra_Wayam (Brazil – Wayampi); Bra_Wai_W (Brazil – Wai Wai); Ecu_Caya (Ecuador – Cayapa); Bra_Bani (Brazil – Baniwa); Col_Core (Colombia – Coreguaje); Col_Embe (Colombia – Embera); Col_Nuka (Colombia – Nukak); Col_Tule (Colombia – Tule); Bra_Yano1 (Brazil – Yanomami 1); Bra_Yano2 (Brazil – Yanomami 2); Bra_Yano3 (Brazil – Yanomami 3); Bra_Wapis (Brazil – Wapishana); Col_Waun (Colombia – Waunana); Bra_Macu (Brazil – Macushi); Col_Guah (Colombia – Guahibo); Ven_Maki (Venezuela – Makiritare); Ven_Yuco (Venezuela – Yuco); Ven_Buta (Venezuela – Butaregua); Col_Ijka (Colombia – Ijka); Col_Kogi (Colombia – Kogui); Mex_Maza (Mexico – Mazatecan); Mex_Maya (Mexico – Maya); Mex_Pima (Mexico – Pima/Maricopa/Papago); Que_Rea (Quechua [28]); Ayma_Gaya (Aymara [12]); Quec_Gaya (Quechua [12]); Boli_Gaya (Bolivia [12]); Peru_Ayac (Ayacucho [29]); Peru_Juni (Peru (Junin) [29]). This_study (Argentina – This study). Current American populations [10]: Col_Colo (Colombians from Medellin, Colombia); Mex_Mexi (Mexican Ancestry from L.A. USA); Peru_Lim (Peruvians from Lima, Peru); Puer_Ric (Puerto Ricans from Puerto Rico).

tions of age-standardized death rates for Alzheimer's disease experienced a 3.6 change (1.1 to 5.6) between 1990 and 2016. In Argentina, Bolivia and Peru, the change was –2.6 (–10.5 to 5.7), 2.9 (–11.7 to 2.3) and –4.0 (–19.2 to 12.8), respectively. During this time frame, the global age–standardized prevalence rate experienced a percentage change of 1.7 (1.0 to 2.04). Conversely, the prevalence rate in Argentina, Bolivia, and Peru decreased by –0.4 (–7 to 0.7), –3.5 (–6.2 to –0.5), and –7.5 (–11.6 to –3.4), respectively. According to the 2019 Global Burden of Disease Database, the worldwide standardized incidence and prevalence rates of Alzheimer's disease were 7.24 (6.22–8.23) and 51.62 (44.28–59.02) cases per million, respectively. These figures were 0.045 (0.04–0.05) and 0.29 (0.25–0.34) for Andean Latin America, respectively [19, 20].

Population differences have also been detected in relation to the burden of cardiovascular diseases. For example, global ischemic heart disease age-standardized incidence in 2017 was 135.6 (122.2 to 150.4), while in Andean Latin American population was 78.8 (68.5 to 90.1) [21].

Contradictory findings have emerged from case/control studies assessing the impact of *APOE* in Alzheimer's disease in admixed Latin American populations. Moreno et al. [22] found no significant differences in the proportion of Native American ancestry between Colombian carriers and non-carriers of *APOE4*, despite both groups having a similar allele frequency. Conversely, Marca et al. [23] revealed that local Amerindian ancestry considerably increased the risk of Alzheimer's disease in *APOE4* carriers in Peru. It is worth noting, however, that investigating dementia in Native American groups and obtaining reliable statistics regarding this issue presents substantial impediments due to cultural and linguistic dissimilarities.

A metanalysis by Garatachea et al. [24] revealed that extreme longevity (>100 years) is negatively associated to the E4 allele and the genotypes E4/E4, E3/E4, and E2/E4 while positively to E2/E3. There have been reports of extreme longevity in some Andean populations of Peru (Huancavelica) and questioned in Ecuador (Vilcabamba) [25, 26]. According to the 2010 Bolivian census, population > 65 years old was 9.8% in the departments of the Andean and Sub-Andean regions, in contrast with 5.8% in the Llanos Regions (lowlands).

The *APOE* system is revealed as a useful marker to analyze new and insufficiently explored traits such as longevity and dementia regarding the adaptation

to an extreme environment such as high altitude and lifestyle.

CONFLICT OF INTEREST

The authors have no conflict of interest to report.

DATA AVALABILITY

Data sharing is not available as it was not explicitly stated in the consent form signed by the participants.

REFERENCES

[1]   Blum CB (2016) Type III hyperlipoproteinemia: Still worth considering? *Prog Cardiovasc Dis* **59**, 119-124.

[2]   Belloy ME, Napolioni V, Greicius MD (2019) A quarter century of APOE and Alzheimer's disease: Progress to date and the path forward. *Neuron* **101**, 820-838.

[3]   Abondio P, Sazzini M, Garagnani P, Boattini A, Monti D, Franceschi C, Luiselli D, Giuliani C (2019) The genetic variability of APOE in different human populations and its implications for longevity. *Genes* **10**, 222.

[4]   Sebastiani P, Gurinovich A, Nygaard M, Sasaki T, Sweigart B, Bae H, Andersen SL, Villa F, Atzmon G, Christensen K, Arai Y, Barzilai N, Puca A, Christiansen L, Hirose N, Perls TT (2019) *APOE* alleles and extreme human longevity. *J Gerontol A Biol Sci Med Sci* **74**, 44-51.

[5]   Kuo C-L, Pilling LC, Atkins JL, Masoli JAH, Delgado J, Kuchel GA, Melzer D (2020) ApoE e4e4 genotype and mortality with COVID-19 in UK Biobank. *J Gerontol A Biol Sci Med Sci* **75**, 1801-1803.

[6]   Egert S, Rimbach G, Huebbe P (2012) ApoE genotype: From geographic distribution to function and responsiveness to dietary factors. *Proc Nutr Soc* **71**, 410-424.

[7]   Singh PP, Singh M, Mastana SS (2006) APOE distribution in world populations with new data from India and the UK. *Ann Hum Biol* **33**, 279-308.

[8]   De Andrade FM, Coimbra Jr CEA, Santos RV, Goicoechea A, Carnese FR, Salzano FM, Hutz MH (2000) High heterogeneity of Apolipoprotein E gene frequencies in South American Indians. *Ann Hum Biol* **27**, 29-34.

[9]   Demarchi DA, Salzano FM, Eugenia Altuna M, Fiegenbaum M, Hill K, Hurtado AM, Tsunetto LT, Petzl-Erler ML, Hutz MH (2005) *APOE* polymorphism distribution among Native Americans and related populations. *Ann Hum Biol* **32**, 351-365.

[10] Abondio P, Bruno F, Luiselli D (2023) Apolipoprotein E (APOE) haplotypes in healthy subjects from worldwide macroareas: A population genetics perspective for cardio-

vascular disease, neurodegeneration, and dementia. *Curr Issues Mol Biol* **45**, 2817-2831.

[11] Gayà-Vidal M, Athanasiadis G, Carreras-Torres R, Via M, Esteban E, Villena M, Vasquez R, Dugoujon J-M, Moral P (2012) Apolipoprotein E/C1/C4/C2 gene cluster diversity in two native Andean populations: Aymaras and Quechuas: APOE/C1/C4/C2 gene cluster diversity in Andeans. *Ann Hum Genet* **76**, 283-295.

[12] Luisi P, García A, Berros JM, Motti JMB, Demarchi DA, Alfaro E, Aquilano E, Argüelles C, Avena S, Bailliet G, Beltramo J, Bravi CM, Cuello M, Dejean C, Dipierri JE, Jurado Medina LS, Lanata JL, Muzzio M, Parolin ML, Pauro M, Paz Sepúlveda PB, Rodríguez Golpe D, Santos MR, Schwab M, Silvero N, Zubrzycki J, Ramallo V, Dopazo H (2020) Fine-scale genomic analyses of admixed individuals reveal unrecognized genetic ancestry components in Argentina. *PLoS One* **15**, e0233808.

[13] Muzzio M, Motti JMB, Paz Sepulveda PB, Yee M, Cooke T, Santos MR, Ramallo V, Alfaro EL, Dipierri JE, Bailliet G, Bravi CM, Bustamante CD, Kenny EE (2018) Population structure in Argentina. *PLoS One* **13**, e0196325.

[14] Van der Auwera GA, O'Connor BD (2020) *Genomics in the cloud: Using Docker, GATK, and WDL in Terra*. O'Reilly Media.

[15] Excoffier L, Lischer HEL (2010) Arlequin suite ver 3.5: A new series of programs to perform population genetics analyses under Linux and Windows. *Mol Ecol Resour* **10**, 564-567.

[16] Raymond M, Rousset F (1995) An exact test for population differentiation. *Evolution* **49**, 1280.

[17] R Core Team (2023) *R: A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria.

[18] Paz-y-Miño C, García-Cárdenas JM (2016) *Alzheimer's disease in Latin America and Caribe*. SMG Books, Dover.

[19] GBD 2016 Dementia Collaborators (2019) Global, regional, and national burden of Alzheimer's disease and other dementias, 1990–2016: A systematic analysis for the Global Burden of Disease Study 2016. *Lancet Neurol* **18**, 88-106.

[20] Li X, Feng X, Sun X, Hou N, Han F, Liu Y (2022) Global, regional, and national burden of Alzheimer's disease and other dementias, 1990-2019. *Front Aging Neurosci* **14**, 937486.

[21] Wang F, Yu Y, Mubarik S, Zhang Y, Liu X, Cheng Y, Cao J (2021) Global burden of ischemic heart disease and attributable risk factors, 1990–2017: A secondary analysis based on the global burden of disease study 2017. *Clin Epidemiol* **13**, 859-870.

[22] Moreno DJ, Pino S, Ríos Á, Lopera F, Ostos H, Via M, Bedoya, G (2017) Genetic ancestry and susceptibility to late-onset Alzheimer disease (LOAD) in the admixed Colombian population. *Alzheimer Dis Assoc Disord* **31**, 225-231.

[23] Marca-Ysabel MV, Rajabli F, Cornejo-Olivas M, Whitehead PG, Hofmann NK, Illanes Manrique MZ, Veliz Otani DM, Milla Neyra AK, Castro Suarez S, Meza Vega M, Adams LD, Mena PR, Rosario I, Cuccaro ML, Vance JM, Beecham GW, Custodio N, Montesinos R, Mazzetti Soler PE, Pericak-Vance MA (2021) Dissecting the role of Amerindian genetic ancestry and the ApoE ε4 allele on Alzheimer disease in an admixed Peruvian population. *Neurobiol Aging* **101**, 298.e11-298.e15.

[24] Garatachea N, Marín PJ, Santos-Lozano A, Sanchis-Gomar F, Emanuele E, Lucia A (2015) The ApoE gene is related with exceptional longevity: A systematic review and meta-analysis. *Rejuvenation Res* **18**, 3-13.

[25] Mazess RB, Mathisen RW (1982) Lack of unusual longevity in Vilcabamba, Ecuador. *Hum Biol* **54**, 517-524.

[26] Sandoval Ú, Sandoval T (2018) Longevidad y pobreza extrema en Huancavelica, Perú. *Anthropia*, 51-63.

[27] Reales G, Rovaris DL, Jacovas VC, Hünemeier T, Sandoval JR, Salazar-Granara A, Demarchi DA, Tarazona-Santos E, Felkl AB, Serafini MA, Salzano FM, Bisso-Machado R, Comas D, Paixão-Côrtes VR, Bortolini MC (2017) A tale of agriculturalists and hunter-gatherers: Exploring the thrifty genotype hypothesis in native South Americans. *Am J Phys Anthropol* **163**, 591-601.

[28] Marca V, Acosta O, Cornejo-Olivas M, Ortega O, Huerta D, Mazzetti P (2014) Polimorfismo genético de la apolipoproteína E en una población peruana. *Rev Peru Med Exp Salud Publica* **28**, 589-594.