

<https://doi.org/10.1038/s44400-025-00025-z>

Genetic contributions to Alzheimer's disease and frontotemporal dementia in admixed Latin American populations

Check for updates

Juliana Acosta-Urbe^{1,2,33} ✉, Stefanie D. Piña-Escudero^{3,4,33}, J. Nicholas Cochran⁵, Jared W. Taylor⁵, P. Alejandra Castruita⁴, Caroline Jonson^{4,6,7}, Erin A. Barinaga⁵, Kevin Roberts⁵, Alexandra R. Levine², Dawwod S. George², José Alberto Ávila-Funes^{8,9}, María I. Behrens^{10,11}, Martín A. Bruno^{12,13}, Luis I. Brusco^{14,15}, Nilton Custodio¹⁶, Claudia Durán-Aniotz¹⁷, Francisco Lopera^{1,35}, Diana L. Matallana^{18,19,20}, Andrea Slachevsky^{10,21,22,23}, Leonel T. Takada²⁴, Lina M. Zapata-Restrepo^{25,26}, Dafne E. Durón-Reyes⁸, Elisa de Paula França Resende^{3,27}, Luisa F. Gómez²⁸, Nancy Gelvez²⁸, M. Beatriz Bistue^{12,13}, Maria E. Godoy¹⁷, Marcelo A. Maito¹⁷, Shireen Javandel^{3,4}, Bruce L. Miller^{3,4}, Mike A. Nalls^{6,7}, Hampton Leonard^{6,7}, Dan Vitale^{6,7}, Sara Bandres-Ciga⁷, Mathew J. Koretsky^{6,7}, Andrew B. Singleton^{7,29}, Caroline B. Pantazis⁷, Victor Valcour^{3,4,30}, Agustín Ibañez^{17,31}, Kenneth S. Kosik^{2,34}, Jennifer S. Yokoyama^{3,4,32,34} ✉ & the Multi-Partner Consortium to Expand Dementia Research in Latin America (ReDLat)*

Latin America's diverse genetic landscape provides a unique opportunity to study Alzheimer's disease (AD) and frontotemporal dementia (FTD). The Multi-Partner Consortium to Expand Dementia Research in Latin America (ReDLat) recruited 2162 participants with AD, FTD, or healthy controls from six countries: Argentina, Brazil, Chile, Colombia, Mexico, and Peru. Participants underwent genomic sequencing, and population structure analyses were conducted using Principal Component Analysis and ADMIXTURE. The study revealed a predominant mix of American, African, and European ancestries, with an additional East Asian component in Brazil. Variant curation identified 17 pathogenic variants, a pathogenic *C9orf72* expansion, and 44 variants of uncertain significance. Seventy families showed autosomal dominant inheritance, with 48 affected by AD and 22 by FTD. This represents the first large-scale genetic study of AD and FTD in Latin America, highlighting the need to consider diverse ancestries, social determinants of health, and cultural factors when assessing genetic risk for neurodegenerative diseases.

Latin America is home to a unique blend of Indigenous American, European, and African ancestries, resulting in a diverse admixed genetic landscape that emerged in the 1500s from the conquest of the Americas and the transatlantic slave trade¹. Significant East Asian immigration during the early 20th century has further contributed to the continent's population diversity observed today^{2,3}. This admixture harbors a spectrum of novel genetic variants, including some that may modulate susceptibility to neurodegenerative diseases like Alzheimer's disease (AD) and frontotemporal dementia (FTD) or may result in a higher allelic frequency of known risk-conveying variants for neurodegeneration⁴. Studying these populations, with their complex genetic architecture, offers an invaluable resource for understanding neurodegenerative diseases.

Assessing admixed populations is particularly interesting because admixture introduces a rich heterogeneity of alleles, which can be crucial for understanding genetic risk. For instance, in the Colombian population, researchers from a single center in Medellín have identified 13 different pathogenic *PSEN1* variants from different ancestral backgrounds⁴. This contrasts with the nine independent variants described in a screening study from nine centers throughout the Iberian Peninsula⁵. The presumptive excess of *PSEN1* variants may have become fixed in a relatively small effective population due to the antimicrobial benefits of high beta-amyloid levels⁴. Additionally, in the Peruvian population, a recent genome-wide association study and functional analysis suggested that the *NFASC* gene, located on chromosome 1, is associated with AD. The *NFASC* locus showed

A full list of affiliations appears at the end of the paper. *A full list of author affiliations appears at the end of the paper. ✉e-mail: acostauribe@ucsf.edu; Jennifer.Yokoyama@ucsf.edu

significant contributions from both European and African ancestries⁶. This finding emphasizes the crucial role of relatively recent founder effects among diverse ancestral backgrounds in uncovering novel variation and understanding the complex genetic foundations of AD.

Due to cultural, religious, and historical factors, Latin American families have traditionally been large. Government policies during the 1960s and 1970s further encouraged population growth, resulting in an average family size of six children^{7,8}. Additionally, geographic and socioeconomic conditions often led to extended families residing in proximity⁹. This unique demographic structure provides an exceptional opportunity to study founder effects in large families by tracing the lineage and impact of specific genetic variants across generations. For example, the *PSEN1* Glu280Ala-variant (canonically known as *PSEN1* E280A) in Colombia can be traced back approximately 500 years to the time of the Spanish invasion. In the large families that descend from the original carrier, the variant became fixed and spread widely among numerous distant relatives. This historical and genetic tracing provides valuable insights into how genetic drift occurs over time in small populations, increasing the likelihood of identifying both risk and protective genetic variants^{10–12}.

The scientific community from the region recognized the immense potential of this approach and formed the Multi-Partner Consortium to Expand Dementia Research in Latin America (ReDLat)¹³. This consortium fosters collaboration among researchers, clinicians, and institutions across six Latin American countries and the United States to leverage the unique genetic diversity and demographic characteristics that influence AD and FTD in Latin American populations. ReDLat aims to highlight the unique genetic diversity and demographic features of these populations associated with neurodegenerative diseases of cognition¹⁴. Ultimately, the consortium's work seeks to refine diagnostic approaches, develop targeted therapeutic interventions, and significantly enhance our understanding of dementia across Latin America¹⁵.

This paper is the first genetic report from this consortium, focusing on AD and FTD in admixed Latin American participants, with a particular emphasis on families. Our research efforts aim to identify genetic variants associated with these neurodegenerative diseases in the region and provide insights into some of the clinical presentations observed within the studied families, thereby enhancing our understanding of AD and FTD across this vast and diverse population. By expanding the genomic dataset in the coming years, we aim to deepen our understanding of the genetic underpinnings and clinical presentations of these neurodegenerative diseases

across diverse Latin American populations, enhancing the potential for targeted interventions and therapies.

Results

Characterization of the population

To investigate the genetic landscape of neurodegeneration in Latin America, we recruited patients with mild to moderate AD or FTD, along with healthy controls from ten memory clinics across six Latin American countries: Argentina, Brazil, Chile, Colombia, Mexico, and Peru. Recruitment occurred in two phases due to the COVID-19 pandemic, which delayed the initiation of ReDLat's prospective enrollment, originally planned for 2020. Despite differences in timing, the inclusion criteria for both the retrospective and prospective cohorts were identical. As of August 17, 2024, a total of 5834 participants had been enrolled in the study (Fig. 1).

By the time of manuscript submission, PCR-free whole genome sequencing (WGS), whole exome sequencing (WES), and/or single-nucleotide polymorphism (SNP) array genotyping had been performed in 2254 participants from both cohorts. After thorough quality control, genomic data from 2162 individuals were retained for analysis. The final dataset included 658 participants with SNP array data, of whom 174 also had WES; 1495 participants with WGS; and 9 individuals with only WES [Fig. 1]. All participants underwent medical and neuropsychological evaluation. Among those with high-quality genomic data, 999 were diagnosed with AD, 381 with FTD, and 755 were classified as cognitively healthy at the time of assessment. The dataset also included eight participants with mild cognitive impairment and 19 individuals with other neuropsychiatric diagnoses, such as Parkinson's disease, Lewy body dementia, atypical parkinsonism, neurodevelopmental disorders, cerebellar ataxia, brain tumor, cognitive impairment associated with non-brain cancer, vascular dementia, severe depressive disorder, bipolar disorder, obstructive sleep apnea, and chronic traumatic encephalopathy. These individuals were members of recruited families and had a relative enrolled in the study with AD- or FTD-related dementia.

As expected, a higher percentage of participants with FTD were under 65 years of age compared to those with AD (22.3% vs. 13.9%). Furthermore, the percentage of female participants was higher in the AD group (66.6%) compared to the FTD group (52.8%). Those in the AD group also showed a higher proportion of individuals who are heterozygous (41.5%) and homozygous (8.6%) for the *APOE* $\epsilon 4$ allele (Table 1). The distribution of *APOE* alleles of unrelated individuals varies per country, as shown in

Fig. 1 | Assembly of the ReDLat genomic dataset. WES whole exome sequencing, WGS whole genome sequencing, SNP single-nucleotide polymorphism.

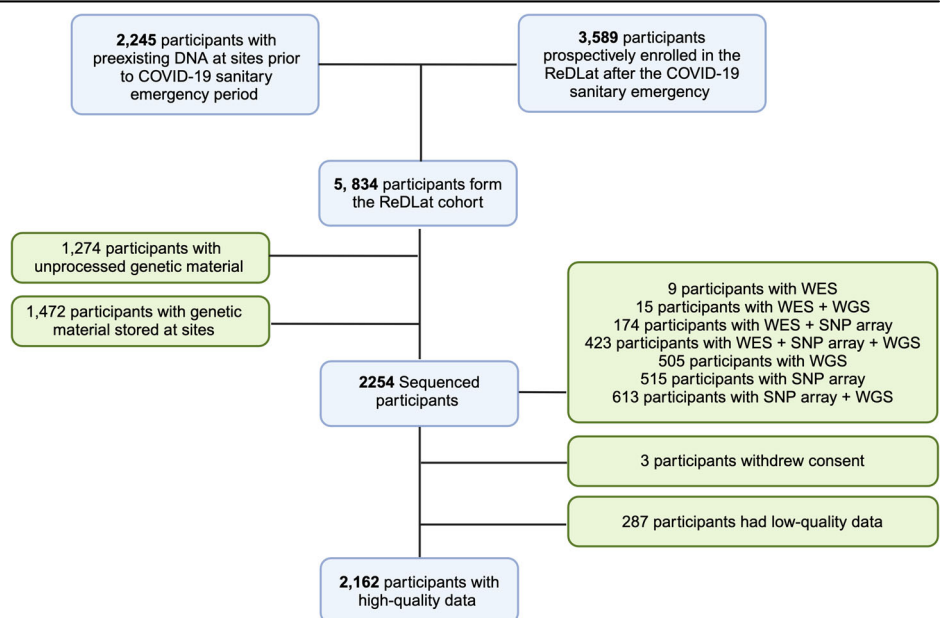


Table 1 | Clinical characteristics of included ReDLat participants

		AD <i>n</i> = 999	FTD <i>n</i> = 381	MCI <i>n</i> = 8	OTHER <i>n</i> = 19	HC <i>n</i> = 755
Age < 65years, <i>n</i> (%)		139 (13.9)	85 (22.3)	2 (25.0)	2 (10.5)	277 (36.7)
Female, <i>n</i> (%)		665 (66.6)	201 (52.8)	6 (75.0)	8 (42.1)	516 (68.3)
AD phenotype, <i>n</i> (%)	Amnestic	521 (52.2)	–	–	–	–
	lvPPA	10 (1.0)	–	–	–	–
	PCA	6 (0.6)	–	–	–	–
	fvAD	1 (0.1)	–	–	–	–
	usAD	461 (46.1)	–	–	–	–
FTD phenotype, <i>n</i> (%)	bvFTD	–	196 (51.4)	–	–	–
	nfPPA	–	25 (6.6)	–	–	–
	svPPA	–	37 (9.7)	–	–	–
	FTD-CBS	–	26 (6.8)	–	–	–
	FTD-MND	–	9 (2.4)	–	–	–
	FTD-PSP	–	9 (2.4)	–	–	–
	usFTD	–	72 (18.9)	–	–	–
	usPPA	–	7 (1.8)	–	–	–
APOE ϵ 4 carriers, <i>n</i> (%)	Heterozygous	408 (41.5)	96 (25.9)	1 (14.3)	2 (10.5)	153 (20.7)
	Homozygous	85 (8.6)	15 (4.0)	–	1 (5.3)	13 (1.8)
Country of origin	Argentina	191 (19.1)	2 (0.5)	–	1 (5.3)	74 (9.8)
	Brazil	42 (4.2)	99 (26.0)	–	–	88 (11.7)
	Chile	115 (11.5)	45 (11.8)	6 (75.0)	11 (57.9)	115 (15.2)
	Colombia	536 (53.7)	193 (50.7)	1 (12.5)	6 (31.6)	289 (38.3)
	Mexico	59 (5.9)	17 (4.5)	–	–	133 (17.6)
	Peru	56 (5.6)	25 (6.6)	1 (12.5)	1 (5.3)	56 (7.4)
Genomic data, <i>n</i> (%)	SNP Array	282 (28.2)	126 (33.1)	6 (75.0)	10 (52.6)	234 (31.0)
	WES	6 (0.6)	–	–	–	3 (0.4)
	WGS	711 (71.2)	255 (66.9)	2 (25.0)	9 (47.4)	518 (68.6)

AD Alzheimer's disease, FTD frontotemporal dementia, MCI mild cognitive impairment, HC healthy controls, *n* number, lvPPA logopenic variant primary progressive aphasia, PCA posterior cortical atrophy, fvAD frontal variant AD, usAD unspecified Alzheimer's disease, bvFTD behavioral variant frontotemporal dementia, nfPPA nonfluent variant primary progressive aphasia, svPPA semantic variant primary progressive aphasia, FTD-CBS frontotemporal dementia-corticobasal syndrome, FTD-MND frontotemporal dementia-motor neuron disease, FTD-PSP frontotemporal dementia-progressive supranuclear paralysis, usFTD unspecified variant of frontotemporal dementia, usPPA unspecified variant of primary progressive aphasia, SNP single nucleotide polymorphism, WES whole exome sequencing, WGS whole genome sequencing.

Supplementary Table 1. Homozygous APOE ϵ 2 carriers were observed only in Colombia (0.2%) and Brazil (1%). Conversely, the highest numbers of APOE ϵ 4 alleles were found in Argentina and Colombia (39% in both), although Colombia has 1% more homozygous carriers than Argentina. It is worth noting that these numbers may not be fully representative due to the Colombian sample size being considerably larger. As larger numbers of samples from all regions are collected, allele frequency estimates will be further refined.

Genetic ancestry

To assess genetic ancestry similarity among our samples, we initially generated a merged ReDLat dataset that included 2153 participants with WGS or SNP array data that passed the concordance analysis. We then used WGS from the 1000 Genomes Project (1000GP) as reference populations to estimate the global ancestry of the participants, employing Principal Component Analysis (PCA) and ADMIXTURE software to estimate global ancestry. We used the 1000GP cohort to identify variants with allelic frequency >10% and in linkage equilibrium that were present in both the 1000GP and the merged ReDLat dataset, resulting in a total of 226,524 variants used for ancestry analyses.

PCA (Fig. 2, and Supplementary Fig. 1) reveals that the ReDLat dataset shows substantial overlap with the American populations (AMR) sampled by the 1000GP, which also included Colombian, Mexican, and Peruvian participants. Though a substantial number of participants overlap with European (EUR) cohorts, there were 21 individuals clustering with the East

Asian (EAS) population, suggesting recent EAS descent subgroups within ReDLat.

When analyzing the data by country, participants from Peru, Mexico, Chile, and Argentina exhibit minimal variation along Principal Component 1 (PC1), which is associated with African ancestry. In contrast, there is greater variation along Principal Component 2 (PC2), which is associated with Amerindian ancestry. This variation is particularly pronounced among participants from Mexico and Peru, who have individuals with a majority of their ancestry being Amerindian. These findings suggest a predominant two-way ancestry pattern in these populations. There is a clear overlap between Argentinian and Brazilian samples with the European (EUR) cohort; however, the Brazilian samples show significant variation along PC1, highlighting the African component of the sample. While most Brazilian samples are distributed primarily between African (AFR) and EUR populations, a small subset clusters with EAS, indicating a distinct ancestral subgroup within Brazil. Colombia displays a clear three-way admixture pattern, as evidenced by its wide distribution along both PC1 and PC2. Additionally, Peruvian and Mexican samples from ReDLat exhibit clear overlap with their 1000GP counterparts, while ReDLat Colombian samples showed greater diversity than those in 1000GP (Fig. 2). This increased diversity is likely due to ReDLat's broader sampling across multiple regions of Colombia. Overall, the PCA analysis confirms that the ReDLat cohort accurately represents the different historical admixture patterns previously described in the corresponding countries¹.

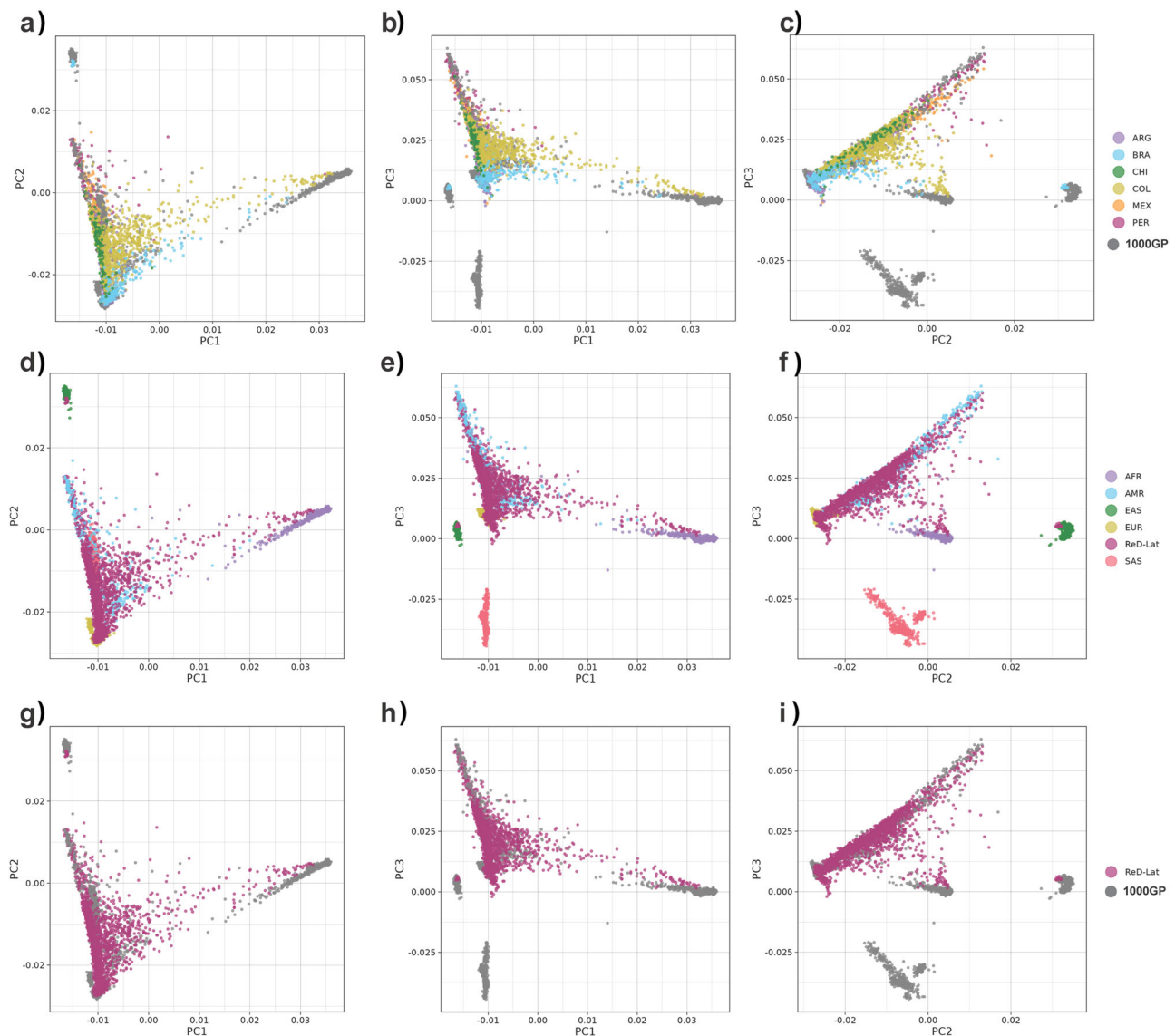


Fig. 2 | Principal component analysis (PCA) of the ReDLat dataset and individuals from the 1000 Genome project (1000GP). Panels represent the main principal component (PC) comparisons. **a, d, g:** PC1 vs. PC2. **b, e, h:** PC1 vs. PC3. **c, f, i:** PC2 vs. PC3. Different color schemes were used to differentiate sample origin. The 1000GP includes African (AFR), European (EUR), South Asian (SAS), East Asian (EAS) and Admixed American (AMR) individuals. ReDLat subcohorts include

individuals from Argentina (ARG), Brazil (BRA), Chile (CHI), Colombia (COL), Mexico (MEX), and Peru (PER). Panels **a, b, c** show color-coded ReDLat cohort participants projected onto the 1000GP reference samples PCA (gray). ReDLat cohort genomes are colored according to their country of origin. **d, e, f** show the ReDLat cohort (magenta) plotted against the color-coded 1000GP subpopulations. Panels **g, h, i** show ReDLat cohort (magenta) overlaid on 1000GP cohort (gray).

To calculate global ancestry proportions (Q -values), which are adjusted p -values accounting for multiple testing and controlling the false discovery rate, we projected the ReDLat samples onto the 1000GP dataset ADMIXTURE results at multiple clustering values (Supplementary Fig. 2). At $K = 5$, where K represents the number of ancestral populations in the clustering analysis, we observed a continental separation of ancestral origins and were able to differentiate the Amerindian component (Fig. 3). Peru is the only country where Amerindian ancestry exceeds European ancestry, followed by Mexico, where these two ancestries show similar distribution. In Argentina, Brazil, Colombia, and Chile, European ancestry is the most prevalent among the participants, with a median value of 86.9% (mean value of 79.7%, standard deviation 17.5) for Argentina and 84.2% (mean value of 73.8%, standard deviation 26.9) for Brazil. African ancestry is present in Colombia and Brazil at lower levels; we observe a continuum of this ancestry, with individuals having over 90% and 75% African descent, down to the mean levels for both countries (around 10%), suggesting ongoing admixture over generations. In contrast, this continuum is not observed in

the EAS component of the Brazilian samples, suggesting a recent diaspora without intercontinental admixture (Supplementary Fig. 3).

Variant pathogenicity analysis

To identify Mendelian forms of neurodegeneration in our cohort, we analyzed data from 1678 participants who had high-quality WGS or WES data to detect pathogenic variants. Following standard practices in complex systems analysis, we employed both “bottom-up” and “top-down” approaches.

Our bottom-up approach was a “gene-to-family” search, in which we initially assessed the genes most commonly associated with adult-onset neurodegeneration for pathogenic variants (see “Methods”). We identified a total of 17 pathogenic variants, a pathogenic *C9orf72* expansion, and 44 variants of uncertain significance (VUS) (Table 2 and Supplementary Table 2). We note that the *C9orf72* expansions were identified using ExpansionHunter software from PCR-free WGS but were not explicitly confirmed by Southern blot, although this predictive tool is very accurate¹⁵.

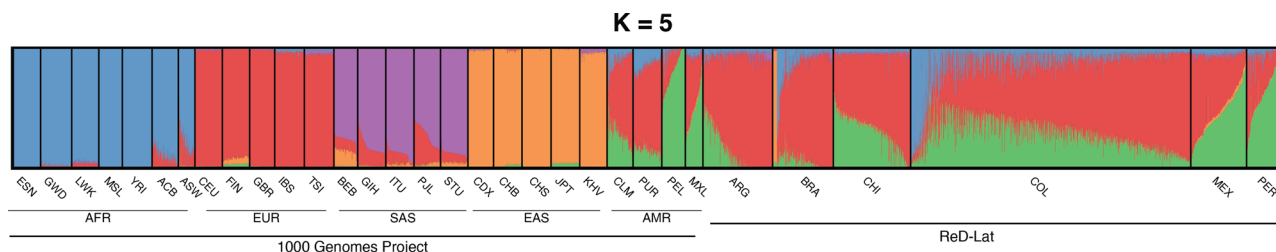


Fig. 3 | Global ancestry proportions of the ReDLat cohort represented by ADMIXTURE *Q* values assuming 5 ancestral populations (*K*). The 1000 Genomes project includes African (AFR): GWD Gambian in Western Divisions in the Gambia, LWK Luhya in Webuye, MSL Mende in Sierra Leone, YRI Yoruba in Ibadan, Nigeria, ACB African Caribbean in Barbados, ASW African Ancestry in the Southwest US. European (EUR): CEU Utah Residents (CEPH) with Northern and Western European Ancestry, FIN Finnish in Finland, GBR British in England and Scotland, IBS Iberian

Population in Spain, TSI Tuscany in Italy. South Asian (SAS): BEB Bengali in Bangladesh, GHI Gujarati Indians in Houston, ITU Indian Telugu in the UK, PJL Punjabi in Lahore, STU Sri Lankan Tamil in the UK. East Asian (EAS): CDX Chinese Dai in Xishuangbanna, China, CHB Han Chinese in Beijing, CHS Han Chinese South, JPT Japanese, KHV Kinh Vietnamese. Admixed American (AMR): CLM Colombian from Medellin, PUR Puerto Rican from Puerto Rico. ReDLat subcohorts: ARG Argentina, BRA Brazil, CHI Chile, COL Colombia, MEX Mexico, PER Peru.

Table 2 | Pathogenic variants found in primary AD/FTD genes

Gene	Coding Change	AA Change	ExAC	CADD	REVEL	Family history	ACMG Classification	Country of origin	Proband's Phenotype(s)
APP	c.2020 G > C	Glu674Gln	<0.00001	25.2	0.635	A-Dom	Likely pathogenic	Mexico	Amnesic AD
C9orf72	c.-45 + 163 GGGGCC [> 24]	A-Dom	Pathogenic	Chile Brazil	bvFTD
GRN	c.21 G > A	Trp7Ter	.	36	.	Sporadic	Pathogenic	Colombia	usAD
	c.58dupT	Cys20Leufs*45	.	.	.	A-Dom	Pathogenic	Chile	bvFTD
	c.328 C > T	Arg110Ter	<0.00001	29.4	.	Positive	Pathogenic	Colombia	FTD-PSP
	c.462+1 G > A	.	.	26.3	.	Positive	Pathogenic	Peru	bvFTD
	c.767_768insCC	Gln257Profs*27	.	.	.	Unknown	Pathogenic	Brazil	bvFTD
	c.1098 T > A	Cys366Ter	.	36	.	Positive	Pathogenic	Colombia	svPPA
MAPT	c.796 C > G	Leu266Val	.	25.8	0.662	A-Dom	Pathogenic	Brazil	bvFTD
	c.915 T > C	Ser305Ser	.	.	.	A-Dom	Pathogenic	Colombia	bvFTD
	c.1280 C > T	Thr427Met	0.00001	33	0.508	Unknown	Pathogenic	Argentina	Amnesic AD
PSEN1	c.250 A > G	Met84Val	.	23.9	0.923	Unknown	VUS	Argentina	Amnesic AD
	c.280 G > A	Val94Met	<0.00001	25.1	0.847	A-Dom	Pathogenic	Colombia	usAD
	c.356 C > T	Thr119Ile	.	24.4	0.795	A-Dom	Pathogenic	Colombia Argentina	bvFTD
	c.415 A > G	Met139Val	.	23.4	0.879	A-Dom	Pathogenic	Argentina	Amnesic AD
	c.428 T > C	Ile143Thr	.	26.8	0.985	A-Dom	Pathogenic	Colombia	Amnesic AD
	c.519 G > T	Leu173Phe	.	25.8	0.923	A-Dom	Pathogenic	Colombia	Amnesic AD
	c.1223 T > C	Ile408Thr	.	27.4	0.877	Unknown	VUS	Peru	Amnesic AD
VCP	c.283 C > T	Arg50Cys	.	27.1	0.606	A-Dom	Pathogenic for essential tremor	Colombia	Amnesic AD
TARDBP	c.1147 A > G	Ile383Val	<0.00001	0.308	0.377	A-Dom	Pathogenic	Colombia	usFTD

AA amino acid, A-Dom autosomal dominant; three affected individuals in two generations. Positive family history; at least one first or second-degree relative with neurodegeneration. usAD unspecified Alzheimer's disease, bvFTD behavioral variant frontotemporal dementia, FTD-PSP frontotemporal dementia-progressive supranuclear paralysis, svPPA semantic variant primary progressive aphasia, AD Alzheimer's disease, VUS variant of unknown significance. PVS1: The PVS1 rule is related to null variants, which usually result in a loss-of-function effect. These variants include stop-gain, frameshift insertions or deletions, splice donor or acceptor, or a start-loss.

In families with AD, we identified several previously described *PSEN1* variants: c.356 C > T (p.Thr119Ile), c.415 A > G (p.Met139Val), c.428 T > C (p.Ile143Thr), and c.519 G > T (p.Leu173Phe)^{4,16,17} (Supplementary Fig. 4). Notably, Thr119Ile and Ile143Thr are of European origin and exhibit identity by descent among Colombian carriers for Ile143Thr and between Colombian and Argentine carriers for Thr119Ile⁴. In contrast, the *PSEN1* c.415 A > G (p.Met139Val) variant, identified in a large family from Argentina, was determined to be of Amerindian origin using the local ancestry inference software RFMix (see "Methods"). This haplotype was also present in two individuals from the 1000 Genomes Project, originating

from Peru and Colombia (Supplementary Fig. 5). The *PSEN1* Met139Val variant has been reported across diverse ancestral backgrounds^{18–20} and is associated with either AD or atypical dementia, characterized by amnesic and behavioral symptoms. These symptoms may also include spastic paraparesis, psychosis, seizures, and myoclonus²¹ (see Supplementary Note 1 for the clinical description of this family). APP duplications were assessed using the methods described by Bowling et al.²². No duplication detected by multiple copy number variation callers overlapped the APP gene.

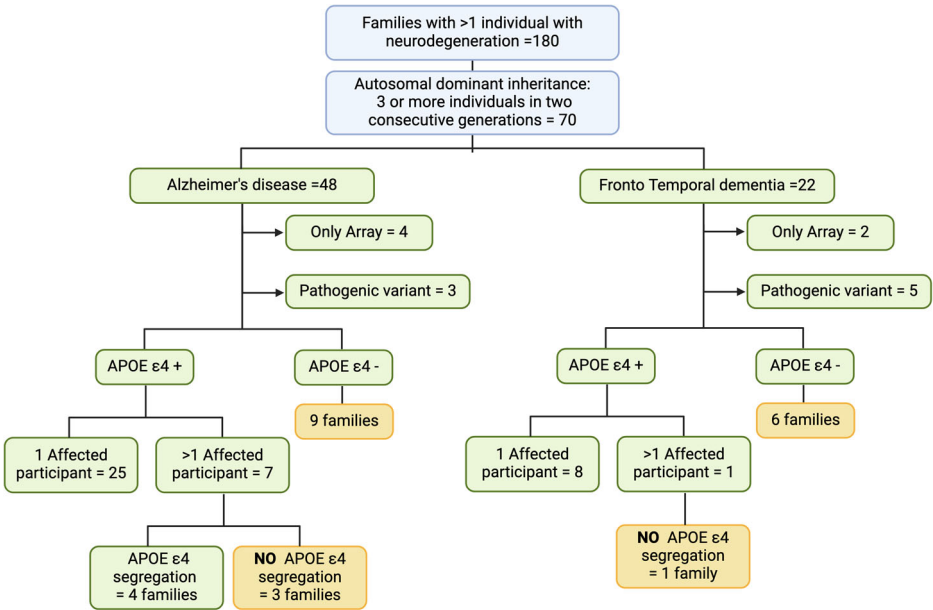
In families with FTD, the gene most commonly associated with hereditary forms of the illness in our cohort was *GRN*, followed by *MAPT* [See

Table 3 | Ancestral origin of pathogenic variants found in more than one carrier

Gene	Coding change	AA change	Number of independent families	Number of carriers with WGS/WES	Ancestral haplotype
GRN	c.767_768insCC	Gln257Profs*27	2	2	European
	c.1098 T > A	Cys366Ter	1	3	European
MAPT	c.915 T > C	Ser305Ser	1	3	European
PSEN1	c.356 C > T	Thr119Ile	4	16	European
	c.415 A > G	Met139Val	1	8	Amerindian
TARDBP	c.1147 A > G	Ile383Val	1	2	European

AA amino acid, WGS whole genome sequencing, WES whole exome sequencing. The ‘number of independent families’ clusters individuals into ‘families’ based on pedigree reconstruction from the participant interview.

Fig. 4 | Cosegregation analysis of *APOE* $\epsilon 4$ and neurodegeneration in families with multiple affected individuals.¹



Supplementary Notes 2 and 3 for the clinical description of these families]. Pathogenic *C9orf72* repeat expansions were observed in several families from geographically non-adjacent areas. The variants in *GRN*, *MAPT*, *C9orf72*, and *TARDBP* that were present in more than one individual were all of European descent (Table 3). The pedigrees of families with pathogenic variants are shown in the Supplementary Figs. 6–9. Allele frequencies of the pathogenic variants identified in the ReDLat cohort, compared with their respective frequencies in gnomAD v4.1.0, are reported in Supplementary Table 3.

After the initial analysis of these primary genes, we expanded our search to include secondary genes associated with adult-onset neurodegeneration. Utilizing the OMIM database, we identified a set of genes where single-nucleotide variants or short insertions/deletions could be disease-causing (Supplementary Table 4)^{19,23–25}. In our analysis, we found four additional pathogenic variants (Supplementary Tables 5 and 6). Three of these variants are present in families with autosomal dominant disease and are located in the *PRNP* and *NOTCH3* genes. Most notably, we identified an FTD patient without motor symptoms carrying a pathogenic variant in *SOD1* c.388 G > A (p.Phe21Leu), which was previously reported in another FTD patient from the same geographical region and is believed to have originated from an Amerindian haplotype⁴.

In contrast, our top-to-bottom approach consisted of a “family-to-gene” search, where we analyzed the pedigrees of all recruited participants with WES or WGS data available (766 individuals grouped in 592 families). After excluding individuals recruited as “healthy”, we identified 426

independent families and classified them based on the presence of affected individuals (Fig. 4). In this cohort, 70 families exhibited autosomal dominant inheritance of neurodegenerative diseases, as evidenced by the presence of three affected individuals in two consecutive generations. The families were later classified according to the diagnosis of the proband, with 48 families identified as having AD and 22 as FTD. These families were subsequently categorized based on the age at disease onset: ‘late onset’ was assigned to families where all affected members presented dementia at ages older than 65 years; ‘early onset’ applied to those where all affected individuals were 65 years or younger at the time of dementia onset; and ‘mixed onset’ described families that included members with both early and late-onset disease (Supplementary Table 7). Though many of the carriers of pathogenic variants belonged to the ReDLat retrospective cohort, 14 of the recruited families were carriers of pathogenic variants, and the majority had a positive family history of neurodegeneration.

In a further analysis of families exhibiting autosomal dominant patterns, we assessed for the presence of at least one allele of *APOE* $\epsilon 4$. Among the families diagnosed with AD, nine were negative for *APOE* $\epsilon 4$ alleles, while 32 had at least one *APOE* $\epsilon 4$ carrier. We were able to determine the *APOE* $\epsilon 4$ allele status for more than one participant in seven families, and only four families showed co-segregation of the *APOE* $\epsilon 4$ allele with the illness. In contrast, when assessing the families with FTD, six were negative for *APOE* $\epsilon 4$ alleles, and in the one family where we could determine the *APOE* $\epsilon 4$ allele status for more than one participant, no co-segregation was observed (Fig. 4).

Discussion

This initial release of genomic data from the ReDLat cohort provides early insights into the genetic underpinnings of neurodegeneration within a Latin American population, supported by genomic analyses of established variants associated with AD and FTD. Our genetic ancestry analysis, leveraging data from the 1000 Genomes Project, revealed tricontinental admixture patterns across most regions and an East Asian component in Brazil, reflecting historical migration and admixture events. The study notably identifies a significant prevalence of autosomal dominant inheritance patterns in AD and FTD, characterized by distinct age-of-onset categorizations, geographic distribution of genetic variants, and a stronger presence of the *APOE* $\epsilon 4$ allele in AD families. These patterns also include newly discovered variants in the *PSEN1* and *APP* genes for AD, which play critical roles in the disease's pathogenesis and the recurrence of a *SOD1* variant presenting as FTD without motor symptoms, suggesting a novel disease phenotype association. Families with as-yet unidentified variation remain strong candidates for future novel gene discovery as additional family members are recruited for gene-mapping linkage studies.

Indeed, there is considerable potential for novel genetic discovery in diverse cohorts such as ReDLat, both in terms of risk for AD and related dementias (ADRD) and resilience against it, in both families and sporadic cases. Previous work in the region has unveiled more than 13 *PSEN1* pathogenic variants in Colombia, including the E280A kindred that spans more than 5000 descendants of a founder couple¹¹. Leveraging larger, diverse cohorts—as well as genetic families with substantial clinical heterogeneity—represents a unique opportunity for the discovery of resilience factors for ADRD, which may serve as strong targets for disease intervention. Notably, two genetic variants previously described as protective for AD, *APOE* Christchurch Arg154Ser (rs121918393)²⁶ and *RELN* His3447Arg²⁷ (rs201731543), were observed in our cohort. *APOE* Arg154Ser was found in one patient with FTD and another with AD, with ages at onset of 66 and 80 years, respectively. *RELN* His3447Arg was observed in three individuals with AD, all with dementia onset in their early 70 s. The lack of replication for *RELN* His3447Arg as a protective variant was also observed in the TANGl cohort⁴, where we identified four carriers who had early-onset dementia; three with FTD (onset at 50, 60, and 65 years) and one with AD (onset at 59 years). All carriers of the *APOE* and *RELN* variants described here were heterozygous and did not carry pathogenic mutations in *PSEN1* or other known disease-causing genes. This finding suggests that protective effects conferred by the *RELN* variants may be restricted to individuals carrying the *PSEN1* E280A. Although *APOE* Arg154Ser heterozygosity was not associated with a protective effect in a patient with FTD and an elderly individual with clinical AD, ongoing evaluation in non-*PSEN1* E280A carriers is necessary to elucidate its potential role in resilience.

As previously noted by Browning et al.²⁸, historical factors like colonization, migration, and bottlenecks have significantly shaped the genetic landscape of Latin American populations. During and after the period of colonization, many Latin Americans lived in small, often isolated villages. This created a population structure characterized by multiple mini-bottlenecks as descendants of a small number of founders in each village tended to remain in the same location for many generations. As families expanded, the specific rare alleles in each place became common, even surpassing the allelic frequencies of the same variant in the ancestral population, as is the case for the *SQSTM1* FTD risk-conferring variant Pro392Leu (rs104893941)^{29–31}. This phenomenon resulted in a genetic map of the region that closely corresponds to the geographic map, where multiple individuals share the same deleterious variants and are identical by descent. The long stretches of identical haplotypes created by the bottlenecks and increased ancestral diversity within isolated populations are advantageous for researchers seeking to identify rare variants associated with diseases like AD and FTD or the interaction between genetic variation and ancestral haplotypes³². These factors underscore the value of family studies in Latin America, offering unique insights into genetic patterns and the potential for discovering new genetic contributions to disease²⁸.

The first-wave study cohort reported here has several limitations. First, we have chosen not to conduct unbiased discovery efforts, such as genome-wide association studies and burden analyses, in this cohort due to the extensive family structure and relatively small sample size of the cross-sectional cohort collected to date. Second, despite the tangible advancement in global representation offered by this cohort, participants are still enriched for higher socioeconomic status due to the urban-centric recruitment. This drawback is being actively addressed through ongoing enrollments and community outreach efforts in more rural areas. Third, as with any clinic-based enrollment cohort, there is a possibility of ascertainment bias among recruited participants because the study recruits from clinical practices specializing in cognitive disorders, which may lead to an overrepresentation of more extreme clinical phenotypes. The findings to date represent significant advances in understanding the etiology of Alzheimer's and Frontotemporal dementia in this region. Continued enrollment in this project will provide additional valuable insights through future studies that map the genetic underpinnings of disease risk in large families, genetic risk burden in cases, and offer well-powered cohorts for case-control studies to identify common risk variants. Moreover, the robust family structure already observed in ReDLat provides a unique opportunity to map genetic modifiers and assess the impact of local genomic ancestry. As global population representation continues to expand, it will be critical to evaluate the generalizability of genetic risk factors for AD and FTD across diverse ancestral backgrounds, within the context of distinct social determinants of health, and accounting for modifiable risk factors that may influence disease risk and resilience across distinct cultures.

Methods

Participant recruitment and inclusion criteria

Clinical diagnosis was determined by site investigators through consensus conferences at each site, adhering to the current diagnostic criteria for AD and FTD^{33–35}. Healthy controls were recruited at the same locations, meeting the following criteria: clinical dementia rating (CDR)³⁶ of 0, a mini-mental state examination (MMSE)³⁷ score greater than 25, or having been evaluated by a neuropsychologist who confirmed normal cognition in participants with few years of formal education. Family members of participants with AD or FTD, aged 18 years or older, were included if there were two or more individuals with neurodegenerative illnesses in the family, or were related to a study participant with a known dementia-associated genetic mutation and had undergone genetic counseling. All participants (diagnosed patients, healthy controls, and family members) demonstrated minimum fluency in the language of assessment (Spanish or Portuguese), had adequate vision and hearing for cognitive testing as determined by the investigator, and were required to have a study partner (informant) with at least 6 months of knowledge about their daily activities and cognitive/functional status. All participants (prospective and retrospective) had to be capable of providing informed consent or have a legally authorized representative.

Written informed consent was obtained from all participants or their legally authorized proxies for all evaluations and assessments conducted, following a detailed explanation of the procedures, associated risks, and potential benefits. This process adhered to the ethical guidelines of each participating country, the Code of Ethics of the World Medical Association, the Declaration of Helsinki, and the Belmont Report. Assent was also secured from participants themselves, ensuring their willingness to participate. The consent process included explicit permission to publish the findings. The study and informed consent procedures were approved by the Institutional Review Board of each participating medical institution.

Participating institutions and their respective Federalwide Assurance (FWA) numbers included the following: Argentina—INECO-Centro de Psicología Médica San Martín de Tours (FWA00028264); Brazil—Hospital das Clínicas da Faculdade de Medicina da Universidade de São Paulo (FWA00001035); Chile—Hospital Clínico Universidad de Chile (FWA00029089) and Universidad Adolfo Ibáñez (FWA00030846); Colombia—Comité de Bioética del Instituto de Investigaciones Médicas, Facultad de Medicina, Universidad de Antioquia (FWA00028864),

Pontificia Universidad Javeriana—Hospital Universitario San Ignacio (FWA00001113), and Fundación Valle de Lili (FWA00029865); Mexico—Instituto Nacional de Ciencias Médicas y Nutrición Salvador Zubirán (FWA00014416); Peru—Hospital Nacional Docente Madre Niño San Bartolomé (FWA00010121); and the United States—University of California, San Francisco—Memory and Aging Center (FWA00000068).

Clinical characterization

As part of the recruitment process, participants were interviewed about their family history of neurodegeneration. Information was collected via self-report from both patients and their study partners using the Genetic Pedigree Software-Progeny^{®38}. A positive family history was defined as having at least one first- or second-degree relative with dementia or another neurodegenerative disorder. Families with three or more affected individuals in two consecutive generations were then labeled as ‘strong family aggregation’. Medical history and a full neuropsychological examination were conducted as described in Ibañez et al.¹³ Retrospective participants were assessed based on a re-evaluation of the available clinical data for each individual. The cognitive tests were harmonized as described in Maito et al.³⁹.

Genetic sequencing and data processing

Sample acquisition and processing. Standardized phlebotomy with EDTA tubes was used for sample collection. Genomic DNA was extracted using Wizard® Genomic DNA Purification Kit (Promega), QIAamp® DNA Mini Kit (Qiagen), or similar salting-out methods. Samples were shipped quarterly from the various participating sites in Latin America to the United States. HudsonAlpha Institute for Biotechnology (Alabama, USA) performed Single Nucleotide Polymorphism (SNP) Arrays, WES, and/or WGS of the samples. Additional whole genome sequencing was performed at Psomagen, Inc. (Maryland, USA)

SNP arrays. Variants were genotyped using the NeuroBooster array from Illumina, designed to capture variants relevant to neurological conditions⁴⁰. Quality control (QC) of the SNP Array data was conducted using Genotools v1 default settings⁴¹. Prior to imputation, the QC’ed output files from Genotools were processed with the *no_qc_imputation_prep.sh* script. This script ran the datasets through the Wrayner script to compare them against all TOPMed freeze 8 variants. Excluded variants were then flipped to rescue additional variants, after which the dataset was processed again through the Wrayner script to compare against the PASS TOPMed freeze 8 variants. Data was subsequently imputed using the TOPMed Imputation Panel and Server v1.3.3, following a previously developed pipeline for multi-ancestral sample sets as described in Vitale et al.⁴¹.

Whole exome sequencing. DNA was processed using Integrated DNA Technologies xGen Exome Hyb Panel v2, and sequenced on the NovaSeq 6000 platform using paired-end 100-base pair reads to a target depth of 100×.

Whole genome sequencing. Samples at Psomagen were prepared using the TruSeq DNA PCR-Free library prep method to avoid PCR amplification bias. Samples at HudsonAlpha (64 of the pass-QC genomes) underwent a custom PCR-free preparation involving Covaris shearing (fragmenting the DNA), end repair (preparing the DNA fragments for ligation), and adapter ligation, all without PCR amplification. All libraries from both sites were then normalized using KAPA qPCR and sequenced on the Illumina NovaSeq 6000 platform to a target depth of 30×. The sequencing was paired-end with a read length of 150 bp (Illumina 150bpPE).

Alignment and variant calling. The raw sequence data (fastq files) were aligned to the hg38 reference genome using the Sentieon v202112.05 implementation of the BWA MEM algorithm at HudsonAlpha. Sentieon

v202112.05 utilities were used to sort the reads, mark duplicate sequences, and recalibrate the base quality scores. Variant calling, which identifies differences between the sample DNA and the reference genome, was performed using GATK4 tools implemented by Sentieon v202112.05. This step was conducted across all samples in a batch for exomes and one for genomes to maintain consistency. Finally, variant quality score recalibration (VQSR) was applied to filter out false positive variant calls, ensuring high-quality data. This comprehensive approach achieved an average recall rate of 99.22% when compared to the Genome in a Bottle high-confidence truth set, indicating a high level of accuracy in detecting genetic variants.

Genomic analyses

Genomic data quality control.

- Variant Call Format (VCF) files were filtered according to established criteria to ensure high-quality data. For whole genome and exome sequences, variants with genotype quality greater than 20 and read depth scores above 10 were retained. The filtered VCF was then annotated with gene names, variant types, and amino acid changes for all exonic variants using GRCh38.99 with SnpEff, dbSNP release 156, CADD 1.6, TOPMed Bravo Freeze 8 allele frequencies, and ClinVar through BCFtools and Annovar^{42–46}. Variants with genotyping rates below 95% by individual and 95% by variant were removed. Chromosomal sex was further validated via genetic data by splitting the pseudoautosomal regions of the X chromosome and analyzing the heterozygosity of the X chromosome, as well as the count of variants present on the Y chromosome. A detailed pipeline and scripts are available at <https://github.com/TauConsortium/redlat-genetics>.
- Relatedness: Family history was documented through elaboration of detailed pedigrees for each recruited participant. Disclosed relatedness was compared to expected genetic relatedness using KING⁴⁷. Individuals with cryptic relatedness (kinship coefficient <0.125 without a familial relationship documented on the pedigrees) or discrepancies between disclosed and genetic relatedness were removed.
- Combining Datasets: To combine the arrays and WGS data, all variant IDs in both datasets were first annotated with “chrom:pos:ref:alt”. The VCFs were then intersected using BCFtools v1.9 isec, producing a list of intersecting variants between the two datasets. Both VCFs were filtered to include only these intersecting variants and then merged using BCFtools v1.9 merge⁴⁸. After merging, the VCF was annotated with dbSNP 156⁴⁹.
- Concordance Check: To ensure concordance between the imputed arrays and WGS, concordance was checked for individuals with data from both methods. The imputed array data was filtered for varying allele frequencies (AF) using BCFtools v1.9⁴⁸. Concordance was assessed for each AF-filtered VCF using SnpSift concordance. The final concordance for each individual was determined by dividing the number of correct variant calls by the total number of intersecting variant calls. The complete script for the concordance check is available at the project’s GitHub repository (see data sharing). Samples with concordance below 0.95 at sites with AF less than 0.0001 were excluded.

Population stratification. To capture the ancestral diversity of the cohort and represent all participating countries, we merged the WGS data with imputed SNP array data for individuals lacking WGS data. The ReDLat dataset was then combined with high-depth WGS data from the 1000 Genomes Project (1000GP)⁵⁰ and filtered to include only biallelic variants with a genotyping rate of >95%. The resulting ReDLat-1000GP dataset was used for the following analyses:

- Principal component analysis (PCA) was conducted using the smartpca package from EIGENSOFT (version 8.0.0)⁵¹. PCA was performed on samples from 1000GP, and the ReDLat samples were subsequently projected onto the principal components.

b. Global ancestry was estimated using ADMIXTURE (version 1.3)⁵². We performed an unsupervised ancestry analysis on the 1000GP data, modeling ancestry from two to eight populations (K). The ancestry fractions for the ReDLat samples were calculated using the allelic ancestry proportions derived from the 1000GP analysis.

Variant pathogenicity analysis. Data from WGS and WES were merged for a joint analysis of pathogenic variation. We used the Online Mendelian In Man (OMIM) database to search for genes associated with autosomal dominant, autosomal recessive, or X-linked forms of adult-onset dementia⁵³.

We manually curated protein-altering variants in the ten genes most commonly associated with adult-onset neurodegeneration: *APP*, *CHMP2B*, *FUS*, *GRN*, *MAPT*, *PSEN1*, *PSEN2*, *TARDBP*, *TBK1*, and *VCP*. Expansions in *C9orf72* were identified using ExpansionHunter v5.0.0^{54,55}, copy number variations in *APP* were called using DELLY, ERDS^{56,57}, MANTA⁵⁸, and CNVnator⁵⁹, following the methodology used in Bowling et al.²². Variants located in introns, the 3' untranslated region (3' UTR), the 5' untranslated region (5' UTR), and synonymous variants within exons were included if their in silico splice-predicting scores (dbSNV_RF_SCORE and dbSNV_ADA_SCORE) were both greater than 0.6, since these variants were considered likely to have an impact on splicing, making them relevant for further study⁶⁰. Exonic non-synonymous variants (missense, nonsense, and frameshift) were analyzed following guidelines from the American College of Medical Genetics and Genomics (ACMG)⁵⁵ and the Guerreiro algorithm for *PSEN1* and *PSEN2* genes⁵. The variants identified in the remaining genes listed in Supplementary Table 4 were queried in the ClinVar⁶¹ database and reported if previously classified as pathogenic or likely pathogenic. Variants identified through this process were then tested by hand for familial segregation to confirm their association with the disease within families. ExAC and gnomAD v4.1.0 databases were used to assess the allelic frequency of variants of interest in global populations.

Local ancestry of pathogenic variants. To determine the haplotypic origin of pathogenic variants, we first phased the ReDLat WGS using SHAPEIT5 (v5.1.1)⁶². Pedigree information and 1000GP phased genomes were included as family and population references to improve phase accuracy. Since heterozygous variants in a single individual cannot be reliably assigned to a specific haplotype, we restricted our assessment of local ancestry to pathogenic variants present in two or more individuals.

After phasing, we estimated the local ancestry of pathogenic variants using RFMix (v2.03-r0)⁶³. We constructed the reference population panel by merging high-coverage sequence data from the Human Genome Diversity Project (HGDP)⁶⁴ and 1000GP. Due to the absence of Amerindian individuals in the 1000GP dataset, we used admixed American samples from 1000GP that were over 99.9% Amerindian in the ADMIXTURE analysis at $K = 3$ as Amerindian reference samples. We then extracted representative individuals of African, Amerindian, European, and East Asian ancestry to build reference cohorts of similar size (100–150 individuals). Finally, we divided the ReDLat WGS into subgroups of individuals with similar global ancestry based on their ADMIXTURE results and ran RFMix with the following settings: terminal node size of five, five expectation-maximization iterations, and both with and without the —reanalyze-reference option.

Data availability

The data used in this publication have been deposited in the FAIR Data Portal within the AD Workbench, supported by the Alzheimer's Disease Data Initiative, and are accessible at <https://www.alzheimersdata.org/ad-workbench>. Allelic counts for variants in twenty genes associated with neurodegeneration can be explored through the *Tau Consortium Variant Visualizer*⁶⁵, an interactive tool designed to plot and compare variant frequencies across subcohorts of the ReDLat dataset. <https://doi.org/10.5062/F4BR8QFB>.

Code availability

All the scripts used for the study have been compiled into a single GitHub repository: <https://github.com/TauConsortium/redlat-genetics>. The scripts for the *Tau Consortium Variant Visualizer*⁶⁵ can be found at <https://github.com/TauConsortium/variant-visualizer>.

Received: 15 November 2024; Accepted: 14 July 2025;

Published online: 21 January 2026

References

- Adhikari, K., Chacón-Duque, J. C., Mendoza-Revilla, J., Fuentes-Guajardo, M. & Ruiz-Linares, A. The Genetic Diversity of the Americas. *Annu. Rev. Genomics Hum. Genet.* **18**, 277–296 (2017).
- Kunimoto, I. Japanese Migration to Latin America. in *Japan, the United States, and Latin America: Toward a Trilateral Relationship in the Western Hemisphere?* (eds Stallings, B. & Székely, G.) 99–121 (Palgrave Macmillan UK, London, 1993). https://doi.org/10.1007/978-1-349-13128-0_4.
- Hassan, W. S. Arabs and the Americas: a multilingual and multigenerational legacy. *Rev. Lit. Arts Am.* **52**, 166–169 (2019).
- Acosta-Urbe, J. et al. A neurodegenerative disease landscape of rare mutations in Colombia due to founder effects. *Genome Med.* **14**, 27 (2022).
- Guerreiro, R. J. et al. Genetic screening of Alzheimer's disease genes in Iberian and African samples yields novel mutations in presenilins and APP. *Neurobiol. Aging* **31**, 725–731 (2010).
- Akgun, B. et al. A genome-wide association study in Peruvians suggests new risk loci for Alzheimer disease. Preprint at *medRxiv* <https://doi.org/10.1101/2023.11.29.23299201> (2023).
- Arriagada, I. *Familias latinoamericanas: diagnóstico y políticas públicas en los inicios del nuevo siglo*. (Naciones Unidas, CEPAL, Div. de Desarrollo Social, Santiago de Chile, 2001).
- Van Bavel, J. The world population explosion: causes, backgrounds and projections for the future. *Facts Views Vis. Obgyn* **5**, 281–291 (2013).
- Arriagada, I. Familias latinoamericanas: cambiantes, diversas y desiguales. *Papeles Población* **13**, 9–22 (2007).
- Acosta-Baena, N. et al. Pre-dementia clinical stages in presenilin 1 E280A familial early-onset Alzheimer's disease: a retrospective cohort study. *Lancet Neurol.* **10**, 213–220 (2011).
- Lalli, M. A. et al. Origin of the PSEN1 E280A mutation causing early-onset Alzheimer's disease. *Alzheimers Dement* **10**, S277–S283.e10 (2014).
- Cochran, J. N. et al. Genetic associations with age at dementia onset in the PSEN1 E280A Colombian kindred. *Alzheimers Dement* **19**, 3835–3847 (2023).
- Ibanez, A. et al. The multi-partner consortium to expand dementia research in Latin America (ReDLat): driving multicentric research and implementation science. *Front. Neurol.* **12**, 631722 (2021).
- Moguilner, S. et al. Brain clocks capture diversity and disparities in aging and dementia across geographically diverse populations. *Nat. Med.* <https://doi.org/10.1038/s41591-024-03209-x> (2024).
- Dolzhenko, E. et al. Detection of long repeat expansions from PCR-free whole-genome sequence data. *Genome Res.* **27**, 1895–1903 (2017).
- Arango, D. et al. Systematic genetic study of Alzheimer disease in Latin America: mutation frequencies of the amyloid β precursor protein and presenilin genes in Colombia. *Am. J. Med. Genet.* **103**, 138–143 (2001).
- Itzcovich, T. et al. A novel mutation in PSEN1 (p.T119I) in an Argentine family with early- and late-onset Alzheimer's disease. *Neurobiol. Aging* **85**, 155.e9–155.e12 (2020).
- Alzheimer's Disease, C. ollaborative Group. The structure of the presenilin 1 (S182) gene and identification of six novel mutations in early onset AD families. *Nat. Genet.* **11**, 219–222 (1995).

19. Rippon, G. A. et al. Presenilin 1 mutation in an African American family presenting with atypical Alzheimer dementia. *Arch. Neurol.* **60**, 884–888 (2003).
20. Jia, L. et al. PSEN1, PSEN2, and APP mutations in 404 Chinese pedigrees with familial Alzheimer's disease. *Alzheimer's. Dement.* **16**, 178–191 (2020).
21. Ryan, N. S. et al. Clinical phenotype and genetic associations in autosomal dominant familial Alzheimer's disease: a case series. *Lancet Neurol.* **15**, 1326–1335 (2016).
22. Bowling, K. M. et al. Genome sequencing as a first-line diagnostic test for hospitalized infants. *Genet. Med.* **24**, 851–861 (2022).
23. Clark, R. F. et al. The structure of the presenilin 1 (S182) gene and identification of six novel mutations in early onset AD families. *Nat. Genet.* **11**, 219–222 (1995).
24. Sandbrink, R. et al. Missense mutations of the PS-1/S182 gene in german early-onset Alzheimer's disease patients. *Ann. Neurol.* **40**, 265–266 (1996).
25. Islam, S. et al. Presenilin is essential for ApoE Secretion, a novel role of presenilin involved in Alzheimer's disease pathogenesis. *J. Neurosci.* **42**, 1574–1586 (2022).
26. Arboleda-Velasquez, J. F. et al. Resistance to autosomal dominant Alzheimer's disease in an APOE3 Christchurch homozygote: a case report. *Nat. Med.* **25**, 1680–1683 (2019).
27. Resilience to autosomal dominant Alzheimer's disease in a Reelin-COLBOS heterozygous man—PubMed. <https://pubmed.ncbi.nlm.nih.gov/37188781/>.
28. Browning, S. R. et al. Ancestry-specific recent effective population size in the Americas. *PLOS Genet.* **14**, e1007385 (2018).
29. Karczewski, K. J. et al. Variation across 141,456 human exomes and genomes reveals the spectrum of loss-of-function intolerance across human protein-coding genes. *bioRxiv* <https://doi.org/10.1101/531210> (2019).
30. van der Zee, J. et al. Rare mutations in SQSTM1 modify susceptibility to frontotemporal lobar degeneration. *Acta Neuropathol.* **128**, 397–410 (2014).
31. gnomAD. Variant: 5-179836445-C-T. gnomAD (2024).
32. Blue, E. E., Horimoto, A. R. V. R., Mukherjee, S., Wijsman, E. M. & Thornton, T. A. Local ancestry at APOE modifies Alzheimer's disease risk in Caribbean Hispanics. *Alzheimer's Dement* **15**, 1524–1532 (2019).
33. Jack Jr, C. R. et al. Revised criteria for diagnosis and staging of Alzheimer's disease: Alzheimer's Association Workgroup. *Alzheimer's Dementia.* **20**, 5143–5169 (2024).
34. Rascovsky, K. et al. Sensitivity of revised diagnostic criteria for the behavioural variant of frontotemporal dementia. *Brain* **134**, 2456–2477 (2011).
35. Gorno-Tempini, M. L. et al. Classification of primary progressive aphasia and its variants. *Neurology* **76**, 1006–1014 (2011).
36. Arnold, B. R., Cuellar, I. & Guzman, N. Statistical and clinical evaluation of the Mattis Dementia Rating Scale-Spanish adaptation: an initial investigation. *J. Gerontol. B Psychol. Sci. Soc. Sci.* **53**, 364–369 (1998).
37. Folstein, M. F., Folstein, S. E. & McHugh, P. R. Mini-mental state'. A practical method for grading the cognitive state of patients for the clinician. *J. Psychiatr. Res.* **12**, 189–198 (1975).
38. Progeny Genetics, L. *Pedigrees Were Constructed and Drawn Using the Genetic Data Management System, Progeny Clinical – Version N from Progeny Genetics.* (Copyright 2019. Reprinted with permission of Progeny Genetics LLC, Delray Beach, FL).
39. Maito, M. A. et al. Classification of Alzheimer's disease and frontotemporal dementia using routine clinical and cognitive measures across multicentric underrepresented samples: a cross sectional observational study. *Lancet Reg. Health Am.* **17**, 100387 (2023).
40. Bandres-Ciga, S. et al. NeuroBooster array: a genome-wide genotyping platform to study neurological disorders across diverse populations. *medRxiv* <https://doi.org/10.1101/2023.11.06.23298176>.
41. Vitale, D. et al. GenoTools: an open-source Python package for efficient genotype data quality control and analysis. *G3 Genes| Genomes|Genetics* **520**, jkae268 (2025).
42. Cingolani, P. et al. A program for annotating and predicting the effects of single nucleotide polymorphisms, SnpEff: SNPs in the genome of *Drosophila melanogaster* strain w1118; iso-2; iso-3. *Fly ((Austin))* **6**, 80–92 (2012).
43. Taliun, D. et al. Sequencing of 53,831 diverse genomes from the NHLBI TOPMed Program. *Nature* **590**, 290–299 (2021).
44. Danecek, P. et al. Twelve years of SAMtools and BCFtools. *GigaScience* **10**, giab008 (2021).
45. ClinVar. <https://www.ncbi.nlm.nih.gov/clinvar/>.
46. Wang, K., Li, M. & Hakonarson, H. ANNOVAR: functional annotation of genetic variants from high-throughput sequencing data. *Nucleic Acids Res.* **38**, e164 (2010).
47. Manichaikul, A. et al. Robust relationship inference in genome-wide association studies. *Bioinformatics* **26**, 2867–2873 (2010).
48. Li, H. A statistical framework for SNP calling, mutation discovery, association mapping and population genetical parameter estimation from sequencing data. *Bioinformatics* **27**, 2987–2993 (2011).
49. Sherry, S. T. et al. dbSNP: the NCBI database of genetic variation. *Nucleic Acids Res.* **29**, 308–311 (2001).
50. Byrsk-Bishop, M. et al. High-coverage whole-genome sequencing of the expanded 1000 Genomes Project cohort including 602 trios. *Cell* **185**, 3426–3440.e19 (2022).
51. Patterson, N., Price, A. L. & Reich, D. Population structure and eigenanalysis. *PLOS Genet.* **2**, e190 (2006).
52. Alexander, D. H., Novembre, J. & Lange, K. Fast model-based estimation of ancestry in unrelated individuals. *Genome Res.* **19**, 1655–1664 (2009).
53. Home—OMIM. <https://www.omim.org/>.
54. Dolzhenko, E. et al. ExpansionHunter: a sequence-graph-based tool to analyze variation in short tandem repeat regions. *Bioinformatics* **35**, 4754–4756 (2019).
55. Richards, S. et al. Standards and guidelines for the interpretation of sequence variants: a joint consensus recommendation of the American College of Medical Genetics and Genomics and the Association for Molecular Pathology. *Genet Med.* **17**, 405–423 (2015).
56. Zhu, M. et al. Using ERDS to infer copy-number variants in high-coverage genomes. *Am. J. Hum. Genet.* **91**, 408–421 (2012).
57. Rausch, T. et al. DELLY: structural variant discovery by integrated paired-end and split-read analysis. *Bioinformatics* **28**, i333–i339 (2012).
58. Chen, X. et al. Manta: rapid detection of structural variants and indels for germline and cancer sequencing applications. *Bioinformatics* **32**, 1220–1222 (2016).
59. Abyzov, A., Urban, A. E., Snyder, M. & Gerstein, M. CNVnator: An approach to discover, genotype, and characterize typical and atypical CNVs from family and population genome sequencing. *Genome Res.* **21**, 974–984 (2011).
60. Li, Q. & Wang, K. InterVar: clinical interpretation of genetic variants by the 2015 ACMG-AMP guidelines. *Am. J. Hum. Genet.* **100**, 267–280 (2017).
61. ALZFORUM | NETWORKING FOR A CURE. <https://www.alzforum.org/>.
62. Hofmeister, R. J., Ribeiro, D. M., Rubinacci, S. & Delaneau, O. Accurate rare variant phasing of whole-genome and whole-exome sequencing data in the UK Biobank. *Nat. Genet.* **55**, 1243–1249 (2023).
63. Maples, B. K., Gravel, S., Kenny, E. E. & Bustamante, C. D. RFMix: a discriminative modeling approach for rapid and robust local-ancestry inference. *Am. J. Hum. Genet.* **93**, 278–288 (2013).
64. Bergström, A. et al. Insights into human genetic variation and population history from 929 diverse genomes. *Science* **367**, eaay5012 (2020).

65. Lu, D. & Acosta-Urbe, J. *Tau Consortium Variant Visualizer*. <https://doi.org/10.5281/zenodo.15652736> (2025).

Acknowledgements

We dedicate this manuscript to Professor Francisco Lopera, a pioneer in the study of neurodegeneration and Alzheimer's disease in Latin America. We thank the individuals and the families who participated in this study. We are grateful to Dr. Sharon R. Browning for her assistance with determining local ancestry, Dylan Lu for developing a variant visualization tool for our data and to all the researchers in the participating centers, whose collaboration made ReDLat possible. The Multi-Partner Consortium to Expand Dementia Research in Latin America (ReDLat) was supported in part by the National Institutes of Health; Intramural Research Program of the NIH; National Institute on Aging (NIA) [R01 AG057234, R01 AG075775, R01 AG021051, R01 AG083799]; NIH Department of Health and Human Services [ZO1 AG000534 and AG000548]; National Institute of Neurological Disorders and Stroke (NINDS); Fogarty International Center; Alzheimer's Association (SG-20-725707); Rainwater Charitable Foundation–Tau Consortium; Bluefield Project to Cure Frontotemporal Dementia; and the Global Brain Health Institute. In addition to the grants reported above, A.S. is supported by ANID/FONDAP [15150012] and ANID/FONDECYT [Regular 1231839]. C.D.A. is supported by ANID/FONDECYT [Regular 1210622] and ANID/PIA/ANILLOS [ACT210096]. D.L.M. received funding from Bluefield Project to Cure Frontotemporal Dementia and Alector. A.I. is supported by ReDLat grants mentioned above and by ANID/FONDECYT Regular [1210195, 1210176, and 1220995]; ANID/FONDAP [15150012]; ANID/PIA/ANILLOS [ACT210096]; and FONDEF [ID2010152 and ID2210029]. J.S.Y. receives funding from NIH-NIA [R01AG062588, P30AG062422, P01AG019724, U19AG079774]; NIH-NINDS [U54NS123985]; NIH-NIDA [75N95022C00031]; AFTD Susan Marcus Memorial Fund; Larry L. Hillblom Foundation; Bluefield Project to Cure Frontotemporal Dementia; French Foundation; Mary Oakley Foundation; Alector and Transposon Therapeutics. This work utilized the computational resources of the NIH STRIDES Initiative (<https://cloud.nih.gov>) through the Other Transaction agreement [Azure: OT2OD032100, Google Cloud Platform: OT2OD027060, Amazon Web Services: OT2OD027852] and the NIH high-performance computing Biowulf cluster (<https://hpc.nih.gov>). M.A.N., H.L., C.J., D.V. and M.J.K.'s participation in this project was part of a competitive contract awarded to DataTecnica LLC by the National Institutes of Health to support open science research.

Author contributions

J.A.-U., S.D.P.-E., J.S.Y., and K.S.K. conceived and designed the project, with contributions from J.N.C., J.W.T., P.A.C., C.J., D.L.M., and M.A.N. A.R.L. and D.S.G. managed pedigree information and conducted the heritability assessment. E.A.B. handled sample processing, while K.R. generated SNP-array data at HudsonAlpha.J.A.-F., M.I.B., M.A.B., L.I.B., N.C., C.D.-A., F.L., A.S., L.T.T., L.M.Z.-R., D.E.D.-R., E.P.F.R., L.F.G., N.G., M.B.B., M.E.G., M.A.M., S.J., B.L.M., C.B.P., A.B.S., V.V., and A.I.

contributed to participant recruitment, evaluation, sample acquisition, genomic data generation, as well as project management and coordination. H.L., D.V., S.B.-C., and M.J.K. assisted with genetic data analyses. F.L. passed away in September 2024; although he was an active collaborator in data acquisition, analyses, and interpretation, he was not able to review the final version of the paper. All remaining authors read and approved the final paper. J.A.-U. and S.D.P.-E. contributed equally to this work, as did J.S.Y. and K.S.K.

Competing interests

M.A.N. owns stock from Character Bio Inc. and Neuron23 Inc. K.S.K. and J.S.Y. are members of the scientific advisory board of the Epstein Family Alzheimer's Research Collaboration. K.S.K. is a member of the board of directors of the Tau Consortium and consults with Expansion Therapeutics and ADRx Pharma. J.S.Y. is Editor-in-Chief for npj Dementia. J.S.Y. was not involved in the journal's review of, or decisions related to, this paper.

Additional information

Supplementary information The online version contains supplementary material available at <https://doi.org/10.1038/s44400-025-00025-z>.

Correspondence and requests for materials should be addressed to Juliana Acosta-Urbe or Jennifer S. Yokoyama.

Reprints and permissions information is available at <http://www.nature.com/reprints>

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Open Access This article is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by-nc-nd/4.0/>.

© The Author(s) 2025

¹Grupo de Neurociencias de Antioquia (GNA), Facultad de Medicina, Universidad de Antioquia, Medellín, Colombia. ²Neuroscience Research Institute and Department of Molecular, Cellular and Developmental Biology, University of California Santa Barbara, Santa Barbara, USACA. ³Global Brain Health Institute (GBHI), Department of Neurology, University of California San Francisco, San Francisco, USACA. ⁴Memory and Aging Center, Department of Neurology, Weill Institute for Neurosciences, University of California San Francisco, San Francisco, USACA. ⁵HudsonAlpha Institute for Biotechnology, Huntsville, USAAL. ⁶DataTecnica LLC, Washington, USADC. ⁷Center for Alzheimer's and Related Dementias (CARD), National Institute on Aging and National Institute of Neurological Disorders and Stroke, National Institutes of Health, Bethesda, USAMD. ⁸Instituto Nacional de Ciencias Médicas y Nutrición Salvador Zubirán, Ciudad de México, México. ⁹Bordeaux Population Health Research Center, University of Bordeaux, Bordeaux, France. ¹⁰Departamento de Neurología y Psiquiatría, Clínica Alemana-Universidad del Desarrollo, Santiago, Chile. ¹¹Centro de Investigación Clínica Avanza (CICA), Departamento de Neurología y Neurocirugía and Departamento de Neurociencia, Facultad de Medicina, Universidad de Chile, Santiago, Chile. ¹²Instituto de Ciencias Biomédicas (ICBM), Facultad de Ciencias Médicas, Universidad Católica de Cuyo, San Juan, Argentina. ¹³Consejo Nacional de Investigaciones Científicas y Técnicas (CONICET), San Juan, Argentina. ¹⁴Asociación Alzheimer Argentina (ALZAR), Buenos Aires, Argentina. ¹⁵Universidad de Buenos Aires, Buenos Aires, Argentina. ¹⁶Instituto Peruano de Neurociencias, Lima, Perú. ¹⁷Latin American Institute for Brain Health (BrainLat), Universidad Adolfo Ibáñez, Santiago, Chile. ¹⁸Instituto de Envejecimiento, Facultad de Medicina, Pontificia Universidad Javeriana, Bogotá, Colombia. ¹⁹Centro de Memoria y Cognición Intellectus, Hospital Universitario San Ignacio, Bogotá, Colombia. ²⁰Departamento de Salud Mental, Hospital Universitario Fundación Santa Fe, Bogotá, Colombia. ²¹Laboratorio de

Neuropsicología y Neurociencias Cognitivas (LANNEC) del Programa de Fisiopatología del Instituto de Ciencias Biomédicas (ICBM), Departamento de Ciencias Neurológicas Oriente y Departamento de Neurociencias de la Facultad de Medicina de la Universidad de Chile, Santiago, Chile. ²²Centro de Gerociencia, Salud Mental y Metabolismo (GERO), Santiago, Chile. ²³Centro de Memoria y Neuropsiquiatría (CMYN), Departamento de Neurología. Hospital del Salvador, Salvador, Chile. ²⁴Hospital das Clínicas, Faculdade de Medicina da Universidade de São Paulo, São Paulo, Brasil. ²⁵Facultad de Ciencias de la Salud, Universidad Icesi, Cali, Colombia. ²⁶Fundación Valle de Lili, Cali, Colombia. ²⁷Faculdade de Medicina da Universidade Federal de Minas Gerais, Belo Horizonte, Brasil. ²⁸Instituto de genética humana, Facultad de Medicina, Pontificia Universidad Javeriana, Bogotá, Colombia. ²⁹National Institute on Aging, National Institutes of Health, Bethesda, USAMD. ³⁰Division of Geriatric Medicine, University of California San Francisco, San Francisco, USACA. ³¹Global Brain Health Institute (GBHI), Trinity College Dublin, Dublin, Ireland. ³²Department of Radiology and Biomedical Imaging, University of California, San Francisco, USACA. ³³These authors contributed equally: Juliana Acosta-Urbe, Stefanie D. Piña-Escudero. ³⁴These authors jointly supervised this work: Kenneth S. Kosik, Jennifer S. Yokoyama. ³⁵Deceased author: Francisco Lopera.

✉ e-mail: acostauribe@ucsb.edu; Jennifer.Yokoyama@ucsf.edu

the Multi-Partner Consortium to Expand Dementia Research in Latin America (ReDLat)

Juliana Acosta-Urbe^{1,2,33}✉, Stefanie D. Piña-Escudero^{3,4,33}, J. Nicholas Cochran⁵, Jared W. Taylor⁵, P. Alejandra Castruita⁴, Caroline Jonson^{4,6,7}, Erin A. Barinaga⁵, Kevin Roberts⁵, Alexandra R. Levine², Dawwod S. George², José Alberto Ávila-Funes^{8,9}, María I. Behrens^{10,11}, Martín A. Bruno^{12,13}, Luis I. Brusco^{14,15}, Nilton Custodio¹⁶, Claudia Durán-Aniotz¹⁷, Francisco Lopera^{1,35}, Diana L. Matallana^{18,19,20}, Andrea Slachevsky^{10,21,22,23}, Leonel T. Takada²⁴, Lina M. Zapata-Restrepo^{25,26}, Dafne E. Durón-Reyes⁸, Elisa de Paula França Resende^{3,27}, Luisa F. Gómez²⁸, Nancy Gelvez²⁸, M. Beatriz Bistue^{12,13}, Maria E. Godoy¹⁷, Marcelo A. Maito¹⁷, Shireen Javandel^{3,4}, Bruce L. Miller^{3,4}, Mike A. Nalls^{6,7}, Hampton Leonard^{6,7}, Dan Vitale^{6,7}, Sara Bandres-Ciga⁷, Mathew J. Koretsky^{6,7}, Andrew B. Singleton^{7,29}, Caroline B. Pantazis⁷, Victor Valcour^{3,4,30}, Agustin Ibañez^{17,31}, Kenneth S. Kosik^{2,34} & Jennifer S. Yokoyama^{3,4,32,34}✉