

1 **Mendelian randomization confirms the role of Y-chromosome loss in Alzheimer's**

2 **Disease etiopathogenesis in males**

3 Pablo García-González MSc^{1,2,*}, Itziar de Rojas MSc^{1,2}, Sonia Moreno-Grau PhD^{3,1},
4 Laura Montreal¹, Raquel Puerta MSc¹, Emilio Alarcón-Martín MsC¹, Inés Quintela
5 PhD⁴, Adela Orellana PhD¹, Victor Andrade MSc^{5,6}, Pamela Martino Adami PhD⁵,
6 Stefanie Heilmann-Heimbach PhD⁷, Pilar Gomez-Garre PhD^{8,2}, María Teresa Perriñán
7 MsC^{8,2}, Ignacio Alvarez MsC^{9,10}, Monica Diez-Fairen PhD^{9,10}, Raul Nuñez Llaves
8 MsC¹, Claudia Olivé Roig MSc¹, Guillermo Garcia-Ribas MD, PhD¹¹, Manuel
9 Menéndez-González MD, PhD^{12,13,14}, Carmen Martínez MD, PhD^{15,13}, Miquel Aguilar
10 MD^{9,10}, Mariateresa Buongiorno MD, PhD^{9,10}, Emilio Franco-Macías MD, PhD¹⁶,
11 Maria Eugenia Saez PhD¹⁷, Amanda Cano PhD^{1,2}, Maria Bullido PhD^{18,2,19}, Luis Real
12 PhD^{20,21}, Eloy Rodríguez-Rodríguez MD, PhD^{22,2}, Jose Royo PhD²¹, Victoria Álvarez
13 PhD^{23,13}, Pau Pastor MD, PhD^{9,10}, Gerard Piñol-Ripoll PhD^{24,25}, Pablo Mir MD,
14 PhD^{8,2,26}, Miguel Calero Lara PhD^{2,27}, Miguel Medina Padilla PhD^{2,28}, Pascual Sánchez-
15 Juan MD, PhD^{28,22,2}, Angel Carracedo PhD^{4,29}, Sergi Valero PhD^{1,2}, Isabel Hernandez
16 MD, PhD^{1,2}, Lluís Tàrraga^{1,2}, Alfredo Ramirez MD, PhD^{5,6,30,31,32}, Mercé Boada MD,
17 PhD^{1,2}, Agustín Ruiz MD, PhD^{1,2}

18 **CORRESPONDING AUTHOR:** Pablo García González (pgarcia@fundacioace.org)

19 **TLF:** +34 649 737 121 **ADRESS:** Marques de Sentmenat 57, Barcelona, Spain 08029.

20 ¹Research Center and Memory Clinic, Ace Alzheimer Center Barcelona, Universidad
21 Internacional de Catalunya, Spain

22 ²Networking Research Center on Neurodegenerative Diseases (CIBERNED), Instituto
23 de Salud Carlos III, Madrid, Spain

24 ³Galatea Bio Inc., Hialeah, FL, USA

- 25 ⁴Grupo de Medicina Xenómica, Centro Nacional de Genotipado (CEGEN-PRB3-
26 ISCIII). Universidade de Santiago de Compostela, Santiago de Compostela, Spain
- 27 ⁵Division of Neurogenetics and Molecular Psychiatry, Department of Psychiatry and
28 Psychotherapy, University of Cologne, Medical Faculty, 50937 Cologne, Germany
- 29 ⁶Department of Neurodegenerative diseases and Geriatric Psychiatry, University Clinic
30 Bonn, 53127 Bonn, Germany
- 31 ⁷Institute of Human Genetics, University of Bonn, School of Medicine & University
32 Hospital Bonn, 53127 Bonn, Germany
- 33 ⁸Unidad de Trastornos del Movimiento, Servicio de Neurología y Neurofisiología,
34 Instituto de Biomedicina de Sevilla (IBiS), Hospital Universitario Virgen del
35 Rocío/CSIC/Universidad de Sevilla, Seville, Spain
- 36 ⁹Fundació Docència i Recerca MútuaTerrassa, Terrassa, Barcelona, Spain
- 37 ¹⁰Memory Disorders Unit, Department of Neurology, Hospital Universitari Mutua de
38 Terrassa, Terrassa, Barcelona, Spain
- 39 ¹¹Hospital Universitario Ramon y Cajal, IRYCIS, Madrid
- 40 ¹²Servicio de Neurología, Hospital Universitario Central de Asturias, Oviedo, Spain
- 41 ¹³Instituto de Investigación Sanitaria del Principado de Asturias (ISPA), Oviedo, Spain
- 42 ¹⁴Departamento de Medicina, Universidad de Oviedo, Oviedo, Spain
- 43 ¹⁵Servicio de Neurología, Hospital Universitario de Cabueñes, Gijón, Spain
- 44 ¹⁶Unidad de Demencias, Servicio de Neurología y Neurofisiología. Instituto de
45 Biomedicina de Sevilla (IBiS), Hospital Universitario Virgen del
46 Rocío/CSIC/Universidad de Sevilla, Seville, Spain
- 47 ¹⁷CAEBI, Centro Andaluz de Estudios Bioinformáticos, Sevilla, Spain
- 48 ¹⁸Centro de Biología Molecular Severo Ochoa, UAM-CSIC, Madrid, Spain
- 49 ¹⁹Instituto de Investigación Sanitaria ‘Hospital la Paz’ (IdIPaz), Madrid, Spain

50 ²⁰Unidad Clínica de Enfermedades Infecciosas y Microbiología. Hospital Universitario
51 de Valme, Sevilla, Spain

52 ²¹Departamento de Especialidades Quirúrgicas, Bioquímica e Inmunología. Facultad de
53 Medicina. Universidad de Málaga. Málaga, Spain

54 ²²Neurology Service, Marqués de Valdecilla University Hospital (University of
55 Cantabria and IDIVAL), Santander, Spain

56 ²³Laboratorio de Genética, Hospital Universitario Central de Asturias, Oviedo, Spain

57 ²⁴Unitat Trastorns Cognitius, Hospital Universitari Santa Maria de Lleida, Lleida, Spain

58 ²⁵Institut de Recerca Biomedica de Lleida (IRBLLeida), Lleida, Spain

59 ²⁶Departamento de Medicina, Facultad de Medicina, Universidad de Sevilla, Seville,
60 Spain

61 ²⁷Instituto de Salud Carlos III, Majadahonda, Spain

62 ²⁸Queen Sofía Foundation Alzheimer Center, CIEN Foundation, Madrid, Spain

63 ²⁹Fundación Pública Galega de Medicina Xenómica- CIBERER-IDIS, Santiago de
64 Compostela, Spain

65 ³⁰German Center for Neurodegenerative Diseases (DZNE), 53127 Bonn, Germany

66 ³¹Department of Psychiatry and Glenn Biggs Institute for Alzheimer's and
67 Neurodegenerative Diseases, San Antonio, TX, USA

68 ³²Excellence Cluster on Cellular Stress Responses in Aging-Associated Diseases
69 (CECAD), University of Cologne, Cologne, Germany

70

71 **ABSTRACT:**

72 Mosaic loss of chromosome Y (mLOY) is a common ageing-related somatic event
73 occurring exclusively in men and has been previously associated with Alzheimer's
74 disease (AD). However, mLOY estimation from genotype microarray data only reflects
75 the mLOY degree of subjects at the moment of DNA sampling. Therefore, mLOY

76 phenotype associations with AD can be severely age-confounded in the context of
77 genome-wide association studies. Here, we applied Mendelian randomization to
78 construct an age-independent polygenic risk score of mLOY (mloy-PRS) using 114
79 autosomal variants. The mloy-PRS instrument was associated with an 80% increase in
80 mLOY risk per SD unit ($p=4.22 \cdot 10^{-20}$) and was orthogonal with age. We found that a
81 higher genetic risk for mLOY was associated with faster progression to AD in males
82 with mild cognitive impairment (HR=1.23; $p=0.01$). Importantly, mloy-PRS had no
83 effect on AD conversion or risk in the female group. The male-specificity of the
84 observed effects suggests that these associations of mLOY with AD are caused by the
85 inherent loss of the Y chromosome, and not by the increased genomic instability
86 underlying mLOY risk. Additionally, we found that blood mLOY phenotype was
87 associated with increased CSF levels of total tau and phosphorylated tau181 in subjects
88 with mild cognitive impairment and dementia. Our results strongly suggest that mLOY
89 is involved in AD pathogenesis. Furthermore, we encourage researchers to use this
90 mloy-PRS instrument to find unbiased associations between mLOY and ageing-related
91 diseases.

92

93 INTRODUCTION

94

95 Alzheimer's disease (AD) is the leading cause of dementia worldwide, accounting for
96 60–80% of total cases [1]. While Mendelian inheritance is suspected to cause early
97 onset AD (<65 years) [2], late onset AD (LOAD, >65 years) is a complex, multifactorial
98 disease influenced by both genetic factors and life exposures. The genetic contribution
99 to LOAD is estimated to be 60–80% [3], being *APOE* the most prominent locus
100 discovered to date [4]. However, demographic features also play a predominant role in
101 AD. Notably, age is considered the most important risk factor for LOAD [5]. On the
102 other hand, women represent nearly two thirds of the global population with AD [1] and
103 show higher rates of cognitive decline [6, 7] than men. However, whether sex should be
104 considered a risk factor for AD or instead a source of disease heterogeneity is a matter
105 of intense debate [8]. Recent reviews highlighted the importance of reporting results for
106 sex interactions and sex-stratified AD data instead of the more widely used approach of
107 adjusting data by sex [9]. These approaches may help elucidate differences in sex-

108 specific AD risk profiles, which will be of great value in the incoming age of precision
109 medicine.

110 The male-specific region of chromosome Y is one of the most unexplored regions of the
111 human genome and has long been considered a genetic wasteland. Mosaic loss of
112 chromosome Y (mLOY) in blood cells is the most common known form of somatic
113 mosaicism in humans [10–12]. Genetic factors together with age, smoking and other
114 environmental stressors are well known risk factors for mLOY [13]. Genetic variants
115 associated with mLOY risk are mainly related to mitotic processes, cell cycle
116 regulation, DNA damage sensing and response, and apoptotic processes [14]. mLOY
117 was initially considered a phenotypically innocuous, age-related trait [15–18]. However,
118 there is increasing evidence that mLOY in blood cells has a direct effect in the
119 etiopathogenesis of several diseases affecting different tissues. Specifically, blood cell
120 mLOY has been associated with susceptibility to multiple ageing-related diseases, i.e.,
121 AD [19], non-hematological cancer [10, 20], cardiovascular diseases [21, 22] and all-
122 cause mortality risk [10]. The main proposed mechanism to explain blood mLOY
123 pathogenesis is by impairment of immune functions caused by the loss of the Y
124 chromosome in leukocytes [23–25]. However, it has been described that autosomal
125 genetic predisposition for mLOY is associated with breast cancer in women, indicating
126 that the underlying genomic instability can also explain the associations between mLOY
127 and disease risk [14].

128 Here, we aimed to study mLOY's impact on AD risk in the GR@ACE and Dementia
129 Genetics Spanish Consortium (DEGESCO) cohorts [26, 27]. First, we checked for
130 blood mLOY associations with AD in both a case-control setting and in the
131 phenoconversion process from mild cognitive impairment (MCI) to all-cause dementia
132 and AD. Subsequently, to remove age confounding effects, we generated an autosomal,
133 age-independent polygenic risk score (PRS) of mLOY and analysed its effect on AD
134 status and progression in both sexes. Finally, we analysed mLOY's impact in different
135 AD-related biomarkers in CSF.

136

137

138

139 METHODS

140

141 - **The GR@ACE-DEGESCO cohort**

142 The GR@ACE-DEGESCO cohort comprises AD patients and controls from the
143 Spanish population. AD patients were collected from ACE Alzheimer Center Barcelona
144 and 12 other cohorts included in the Dementia Genetics Spanish Consortium
145 (DEGESCO) (Supplementary Table 1). Control individuals were provided by ACE
146 Alzheimer Center (Barcelona, Spain), Valme University Hospital, the Spanish National
147 DNA Bank Carlos III (Salamanca, Spain) and other DEGESCO members. DNA
148 extracted from peripheral blood or saliva (Supplementary Table 1) was genotyped in the
149 Spanish National Center for Genotyping (CeGen, Santiago de Compostela, Spain),
150 using the Axiom 815K Spanish Biobank Array (Thermo Fisher), as previously
151 described [26, 27].

152 - **The ACE MCI-EADB cohort.**

153 The EADB cohort is a prospective cohort conformed by MCI individuals recruited
154 between 2006 and 2013 at ACE Alzheimer Center Barcelona. Briefly, individuals with
155 a Clinical Dementia Rating Scale (CDR) of 0.5 and older than 60 years were selected
156 and underwent at least one follow-up, consisting of neurological, neuropsychological
157 and social work evaluations. Detailed definition of the ascertainment of this cohort has
158 been described [28, 29]. DNA genotyping was performed as described elsewhere [30].
159 Briefly, DNA extracted from peripheral blood was genotyped with the Illumina
160 Infinium Global Screening Array (GSA, GSASharedCUSTOM_24+v1.0) in LIFE &
161 BRAIN CENTER, (EADB node, Bonn, Germany) and SNP genotype calls were
162 obtained from raw probe intensity data in the same center.

163 - **Criteria for AD diagnosis, case-control setup**

164 AD diagnoses were established in all cases by a multidisciplinary working group
165 conformed by neurologists, neuropsychiatrists and social workers, following DSM-IV
166 criteria for dementia and the National Institute on Aging and Alzheimer's Association's
167 (NIA-AA) 2011 guidelines for AD definition. In the present study, individuals were

168 labeled as AD when possible or probable AD was endorsed by neurologists at any point
169 of their clinical history. Written informed consent was obtained from all participants.
170 The Ethics and Scientific Committees have approved this research protocol (Acta
171 25/2016, Ethics Committee. H., Clinic I Provincial, Barcelona, Spain).

172 **Assessment of MCI to dementia/AD conversion**

173 MCI to dementia conversion was determined by integrating CDR, Global Deterioration
174 Scale (GDS), and diagnostic assessments at ACE Alzheimer Center Barcelona, assigned
175 at a consensus conference including neurologists, neuropsychologists and social
176 workers [31]. Conversion to dementia was defined as the first clinical evaluation
177 reporting a diagnosis of AD [32, 33], vascular dementia [34], mixed dementia (AD with
178 cerebrovascular disease), frontotemporal dementia [35, 36] or dementia with Lewy
179 bodies [37], combined with a CDR change from 0.5 to ≥ 1 , and GDS ≥ 4 . AD
180 converters were defined as the fraction of converters to dementia that were diagnosed
181 with AD. Baseline criteria varied depending on whether the exposure was mLOY
182 phenotype or its associated PRS: In the first case, baseline was defined as the moment
183 of blood sampling used afterwards for germline DNA extraction, genome-wide
184 genotyping and mLOY estimation. We selected only those individuals who met
185 Petersen's criteria [38, 39], for amnesic and non-amnesic MCI, at the closest clinical
186 evaluation to DNA sampling. Because genotypes used for PRS estimates are invariable,
187 baseline was defined as the patient's first clinical record meeting Petersen's criteria for
188 PRS analysis. Follow-up time was defined as the time window between baseline and:
189 (a) the date of conversion to dementia (converters) and (b) the date of last clinical
190 evaluation (non-converters). In order to have a prospective cohort, disease progression
191 models only included individuals who were either originally selected as controls/MCIs
192 in the GR@ACE-DEGESCO case-control cohort, or present in the MCI cohort (ACE
193 MCI-EADB).

194 **LOY determination**

195 PennCNV [40–42] was used to process CEL files, following the recommended
196 workflow for Affymetrix arrays [43] to obtain Log R Ratio (LRR) and B allele
197 frequency (BAF) values for each array probe in our dataset. Determination of mLOY
198 was performed using MADloy package for R [44]. Briefly, this method estimates

199 mLOY by normalizing the median LRR of probes found at the male specific region of
200 chromosome Y (mLRR-Y), against the 5% trimmed mean LRR of autosomal
201 chromosomes. Thus, only probes located between pseudoautosomal regions 1 (PAR1)
202 and 2 (PAR2) in chromosome Y, excluding the X transposed region, (chrY:6611498-
203 24510581; hg19/GRCh37) are used to compute mLRR-Y. To call mLOY status, we
204 used the mLRR- Y_{thres} method of the MADloy package. Briefly, a threshold is
205 determined by extrapolating the 99% confidence interval of the positive side of the
206 cohort mLRR-Y distribution [10]. Then, samples with mLRR-Y values below the
207 empirically calculated threshold are assigned mLOY status (or calls). To overcome
208 computational power limitations, mLOY calls were obtained in two randomized
209 batches. B-deviation (Bdev), defined as the mean deviation from the expected BAF
210 (0.5) for heterozygous SNPs, located in PAR1 regions (PAR1-Bdev) was used as a
211 complementary indicator of mLOY (Supplementary Figure 1).

212 **Sample processing and QC**

213 LRR and BAF values for all biallelic markers were obtained from 20 068 CEL files
214 (call rate >0.97 per sample and $>0,985$ per plate). Then, reported male samples were
215 retrieved, and samples with mean LRR-X and LRR-Y corresponding to female (XX) or
216 sexual chromosome aneuploidies were discarded. Additionally, samples with high
217 heterozygosity rate, high chromosome X heterozygosity and population outliers were
218 removed from our dataset. Samples with LRR SD > 0.46 , a standard QC parameter for
219 Affymetrix LRR data, were removed. GENESIS R package [46] was used to examine
220 relatedness within our dataset. Second degree relatives were detected by using a kinship
221 threshold of 0.046875 and were filtered out of the dataset. Finally, outliers in the
222 mLRR-y and b-deviation distribution were removed. Specific QC procedures and
223 sample filtering steps for each analysis are summarized in Supplementary Figure 2.

224 **mLOY autosomal Polygenic Risk Score**

225 Processing, QC and imputation of the genome-wide SNP data were performed as
226 described elsewhere [27, 30]. An mLOY autosomal polygenic risk score (mloy-PRS)
227 was calculated based on independent genome-wide significant variants previously
228 described [14]. Out of the 156 reported SNPs, we excluded those unavailable in our
229 dataset, considered rare variants ($\text{MAF}<0.01$), with low imputation quality ($R^2<0.3$), or

230 located within the sexual chromosomes, leaving us with a final number of 114
231 autosomal SNPs (Supplementary Table 2). mloy-PRS was calculated for all individuals
232 in the GR@ACE-DEGESCO and ACE MCI-EADB cohorts by adding the dosage of
233 risk alleles weighted by their reported effect sizes (beta coefficients). To ease
234 interpretation of results, mloy-PRS units were standardized (SD=1).

235 **Core AD biomarkers & targeted proteomics**

236 Levels of Abeta-42, tau phosphorylated at position 181 (p-tau) and total tau were
237 measured the same day in cerebrospinal fluid (CSF) samples obtained via lumbar
238 punctures using commercially available ELISA immunoassays (INNOTEST® β-
239 AMYLOID (1-42), INNOTEST® hTAU, and INNOTEST® PHOSPHO-TAU(181P)
240 (Fujirebio, Spain).

241 CSF and paired plasma samples collected the same day, as described elsewhere [47, 48],
242 underwent targeted proteomics using ProSeek® multiplex immunoassay by Olink
243 Proteomics (Uppsala, Sweden). Protein concentration was measured for 184 proteins
244 included in the commercially available ProSeek® Multiplex panels (Inflammation &
245 Neurology) in both fluids. Quality control details and further description of this data are
246 provided elsewhere [49].

247 **Statistical Analysis**

248 Data processing and analysis was performed with R Software [50]. To harmonize effect
249 directions, the mLRR-Y variable was multiplied by -1 due to lower values of mLRR-Y
250 representing a higher degree of mLOY. Logistic regressions adjusted by age, *APOE*
251 genotype and population structure were fitted for case-control analysis. Survival R
252 package [51] was used to fit Cox proportional-hazards models to assess MCI conversion
253 to all-cause dementia or AD. Due to the age-dependent nature of mLOY [11, 52], only
254 individuals with available age at DNA sampling information were used for analyses
255 involving mLOY or mLRR-Y. To correct for population structure, only the principal
256 components (PCs) that were associated with the dependent variable were included in the
257 models. Population structure was not corrected in the Cox models as PCs showed no
258 effect on disease progression (Supplementary Table 3, Supplementary Figure 3), likely
259 because all MCI samples came from the same center (ACE Alzheimer Center

260 Barcelona). *APOE* genotypes were modelled as a continuous variable ranging from -2 to
261 2, where each *APOE*- ϵ 2 allele contributed with -1 and each *APOE*- ϵ 4 allele added +1,
262 as previously described [29]. In order to control ascertainment and genotyping bias
263 between MCI cohorts (GR@ACE-DEGESCO & ACE MCI-EADB), a dichotomic
264 variable was introduced in models testing association of mLOY-PRS with disease
265 progression. For analysis of Olink proteomic data, linear regressions were adjusted by
266 age, the time window between DNA sampling and lumbar puncture, and *APOE*
267 genotype. Due to the high correlation between many CSF proteins, total tau and p-tau
268 levels (Supplementary Figure 4), we also included models adjusted by total tau levels.
269 Fixed-effect inverse variance weighted meta-analysis was performed with the `rma.uni`
270 function included in the `metafor` R package [53].

271

272 RESULTS

273

274 We computed mean LRR-Y and LRR-X values for 7 954 clinically reported male
275 samples and plotted these values, generating clusters that allowed us to check the sexual
276 chromosome dosages of these individuals (Supplementary Figure 5). We detected one
277 individual with a gain of chromosome Y (GOY, XYY) compatible with a supermale
278 syndrome and three individuals with Klinefelter syndrome (XXY). Women (XX),
279 GOY, Klinefelter individuals, and outliers were discarded prior to mLOY computation.
280 For the 7 843 remaining XY individuals, we removed second-degree or lower relatives,
281 and samples with low genotype call rate (≤ 0.97) or excess heterocigosity ($>3SD$ over
282 cohort mean heterocigosity). We ran principal component analysis to identify the
283 population structure and removed 72 individuals from non-European population ($>6SD$
284 from 1000 Genomes European population mean). Subjects with detectable autosomal
285 chromosomopathies, i.e., Down's Syndrome, were also excluded. After applying these
286 exclusion criteria, we split the remaining 6 955 male samples in two randomized
287 batches and performed mLRR-Y computation and mLOY calling. We did not detect
288 batch effect due to cohort splitting (Supplementary Figure 6). We excluded 12
289 additional samples with LRR SD > 0.46 . Finally, we plotted mLRR-Y and PAR1-
290 Bdev values in order to identify and remove individuals with detectable anomalies in
291 chromosome Y (i.e. partial loss of chromosome Y) or loss of heterocigosity in the

292 PAR1 region (Supplementary Figure 5). QC and filtering steps for analysis of mLOY
293 phenotypes and mLOY-PRS of our data are summarized in Supplementary Figure 2.

294

295 To have a first glimpse at our data, mLRR-Y values of all AD cases and controls were
296 plotted with respect to age (Figure 1). The first thing that became apparent was that our
297 control population is significantly younger than the AD population.
298 Additionally, mLOY occurrence before 65 years was a very rare event in our
299 cohort, indicating that our control population below this age threshold may not be
300 representative for assessing mLOY's effect on AD. Moreover, our control population
301 mostly lacked individuals older than 85 years. Consequently, we decided to establish a
302 65–85 age window for analyzing mLOY's effect on AD. This matches the usual age at
303 onset range for preclinical, prodromal, and mild dementia stages for LOAD in our
304 population [31] and helped reduce the age gap between our case and control groups
305 (Supplementary Table 4). Concordant with previous reports, we observed a clear age-
306 related increase of mLOY events in the older individuals (Figure 1). Age was associated
307 with mLOY occurrence in males aged 65–85, with an estimated 1% increase in the
308 chance of developing LOY every year ($p=3.50 \cdot 10^{-11}$).

309

310 Then, we assessed if continuous mLRR-Y values were differentially distributed
311 amongst cases and controls. Both unadjusted Kolmogorov-Smirnoff test ($D=0.18124$; p
312 $< 2.2 \cdot 10^{-16}$) and ANCOVA adjusted by age at DNA sampling and *APOE* ($F=68.0$;
313 $p=2.54 \cdot 10^{-16}$, Supplementary Table 5) yielded highly significant results in models
314 including all available males in our cohort.

315 Next, we fitted logistic regressions for AD status using mLOY calls and mLRR-Y,
316 defining 3 experimental setups: a) all men with available age at DNA sampling, b) 65–
317 85-year-old men and c) dividing the data in age groups (65–70, 70–75, 75–80 & 80–85
318 years old). Logistic regressions were adjusted by age at blood sampling, *APOE*
319 genotype and relevant PCs (Supplementary Table 6).

320 We found that the continuous mLRR-Y variable was associated with AD in the group
321 including all men ($N=2\ 697$; $OR=2.74$; $p=0.01$), indicating that AD cases had an
322 increased degree of LOY mosaicism compared to controls. We also observed increased
323 mLOY levels in AD males in the 65–85 ($N=1\ 944$; $OR=2.19$; $p=0.09$) and stratified age
324 groups (Table 1) with respect to controls, but these differences were not statistically
325 significant. Importantly, we noticed that the significant effect observed in the model

326 including all available males could be, at least partially, driven by the dramatic age
327 differences between AD cases and controls in our cohort (Figure 1, Supplementary
328 Table 4), even after adjusting our data by age. mLOY calls were not significantly
329 associated to AD in the group including all men (N=2 697; OR=1.14; p=0.35), the 65–
330 85-year-old group (N=1 944; OR=1.04; p=0.81), nor in the age-stratified groups (Table
331 1).

332

333 In order to check for an effect of mLOY in risk of conversion to all-cause dementia and
334 AD, we fitted Cox proportional-hazards models adjusted by age at sampling and *APOE*
335 genotype in our prospective cohort of MCI males (N=400). The continuous mLRR-Y
336 variable had a non-significant risk effect in MCI conversion to all-cause dementia
337 (HR=1.93; p=0.10). The effect size increased when we calculated the model exclusively
338 using conversion to AD but did not reach statistical significance (HR=2.05; p=0.19).
339 mLOY calls also showed similar but smaller non-significant positive effects for
340 conversion to dementia (HR=1.17; p=0.40) and AD (HR=1.38; p=0.20, Figure 2). Cox
341 model results are summarized in Table 2.

342

343 Because the impact of age on mLOY and AD might obscure genuine associations
344 between both phenotypes (mLOY and AD), we decided to construct a mLOY
345 Polygenic Risk Score (mloy-PRS) to evaluate the impact of the genetic variance
346 associated with mLOY phenotype in AD risk. Our rationale was to implement a
347 Mendelian randomization strategy reasoning that, if blood cell mLOY is genuinely
348 associated with AD, the genetic factors linked to blood mLOY risk should be also
349 associated with AD and its related endophenotypes. To this end, mloy-PRS was
350 generated based on a list of autosomal genome-wide significant SNPs associated with
351 the mLOY phenotype identified in a recent genome-wide association study
352 (Supplementary Table 2) [14].

353 For benchmarking purposes of the constructed PRS, we initially validated the effect of
354 mloy-PRS in the mLOY cell phenotype in our cohort (65–85 years old). To this end, we
355 fitted a logistic regression for mLOY calls with PRS, age at DNA sampling and *APOE*
356 genotype as predictors. The PRS (OR=1.80; p=4.22·10⁻²⁰) and age at sampling
357 (OR=1.08; p=5.07·10⁻¹¹), but not *APOE* (OR=0.88; p=0.15), were significantly
358 associated with mLOY in our population (Supplementary Table 7). Importantly, the
359 PRS was orthogonal with age and evenly distributed across the age

360 spectrum (Supplementary Figure 7). Our results reassure the validity of mLOY-PRS as a
361 mendelian randomization instrument for investigating the causal role of mLOY in AD
362 and its endophenotypes, and independently confirm the combined risk effect of
363 previously reported loci in the mLOY phenotype [14].

364

365 In the case-control setup, we checked the effect of mLOY-PRS on AD risk by fitting
366 logistic regressions adjusted by *APOE*, age and relevant PCs (Supplementary Table
367 8). Interestingly, the effect of mLOY-PRS on AD could be measured in the female
368 samples as well. Therefore, 3 analysis groups were established: all (males + females) and
369 stratified by sex. No association between mLOY-PRS and AD was found in the group
370 including both sexes (Table 3). However, after sex stratification, we found a weak non-
371 significant positive effect of mLOY-PRS with respect to AD in the male subgroup (N=2
372 471; OR=1.07; p=0.12), while the effect was mostly neutral in the female subset (N=4
373 978; OR=1.00; p=0.93). Next, we assessed the effect of mLOY-PRS in disease
374 progression. We adjusted Cox models by age, *APOE* genotype and cohort
375 ascertainment. We found a male-specific positive effect of mLOY-PRS in the disease
376 progression models (N=682) (Table 3, Supplementary Figure 8), with a suggestive
377 signal for MCI to dementia progression (HR=1.11; p=0.08) and a statistically significant
378 risk effect for MCI to AD progression (HR=1.23; p=0.01). Of note, no association
379 between mLOY-PRS and conversion to all-cause dementia (HR=0.99, p=0.81) or AD
380 (HR=0.99; p=0.85) was found in the female group (N=1 082).

381

382 Following these results, we proceeded to examine the existence
383 of associations between mLRR-Y, the continuous mLOY variable, and the levels
384 of core AD biomarkers in CSF (Abeta-42, p-tau and total tau). Only individuals aged
385 65–85 at the moment of lumbar puncture (LP) were kept, and those with a gap higher
386 than 5 years between DNA sampling and the LP were excluded from the analysis
387 (N=214). Linear regressions were adjusted by *APOE* genotype, age at LP, and
388 time window between blood sampling and LP. To account for the effect of syndromic
389 status on the levels of Abeta-42, p-tau and total tau, we calculated the effect in two
390 groups (MCI N=148; dementia N=66) and then performed inverse-variance weighted
391 fixed effect meta-analysis of the effect of mLRR-Y. We found that both p-tau ($\beta=41.92$;
392 p=0.01) and total tau ($\beta=396.69$; p=0.004) levels were increased in individuals with a
393 higher degree of mLOY (Figure 3).

394

395 Afterwards, we checked for mLRR-Y associations with proteomics data obtained with
396 the Olink ProSeek® multiplex immunoassay for paired plasma and CSF samples in 135
397 MCI males. Because mLOY is known to affect the immune system [10, 24, 54], and
398 inflammation is involved in many processes related to AD pathogenesis [55], we
399 analysed Olink Neurology and Inflammation panels. We detected inflation in our
400 models ($\lambda=1.86$), with most proteins showing increased levels in the CSF of individuals
401 with a higher degree of blood mLOY (Figure 3). A similar pattern was observed when
402 we analysed the effect of *APOE* genotype and total tau levels, with a big fraction of the
403 proteins showing increased CSF levels in individuals carrying *APOE* risk alleles or
404 displaying higher tau levels, respectively (Figure 3). Moreover, after adjusting our
405 models by total tau we lost most CSF associations and the inflation factor was
406 drastically reduced to $\lambda=0.86$ (Figure 3). After covariation with total tau, we found 7
407 nominally significant markers in plasma and 1 nominally significant marker in
408 CSF. Nevertheless, no proteins passed FDR correction suggesting that most mLOY
409 associations can be explained by the previously observed correlation between mLOY
410 and tau levels. Summary statistics for association of mLRR-Y to the CSF and plasma
411 proteins are available (Supplementary Tables 9-12).

412

413 DISCUSSION

414

415 In the present study, we found that MCI males with high genetic risk of developing
416 mLOY have increased chances of progressing to AD over time. The autosomal loci
417 used to construct the mLOY PRS had no effect on AD progression in the female subset
418 of our cohort, strongly suggesting that the observed effect is produced via loss of the Y
419 chromosome among men. Importantly, modelling mLOY through its associated genetic
420 variance allowed us to observe mLOY-induced alterations in AD pathogenesis in an
421 age-independent manner, unparalleled in previous studies. These results add to previous
422 evidence reporting mLOY as a male-specific AD pathogenic factor.

423 mLOY is the most common known form of somatic mosaicism among men [10].
424 Concordantly, we detected mLOY in 18.9% men aged 65–85 in our cohort. Though
425 classically thought of as a harmless age-related trait, recent studies have revealed that

426 mLOY increases risk for all-cause mortality and several diseases [10, 20–22]. With
427 such a high prevalence in the elderly population, interest in determining mLOY’s effect
428 in age-related diseases has raised over the past decade. Previous studies reported that
429 mLOY significantly increases AD risk and progression rate [19]. A more recent
430 publication claims that transcriptomic extreme downregulation of chromosome Y
431 decreases AD resilience in men [56]. However, whether mLOY acts as an AD
432 promoting factor or is just a by-product of ageing is not yet clearly established.

433 Concordant with previous studies [19], we found a higher degree of mLOY mosaicism
434 (mLRR-Y) in our AD vs. control population in unadjusted Kolmogorov-Smirnoff
435 models ($D=0.18124$; $p < 2.2e-16$) and age-adjusted ANCOVA ($F=68.0$; $p=2.54 \cdot 10^{-16}$).
436 We then performed a case-control logistic regression in all available males in our
437 cohort, obtaining statistically significant results ($OR=2.74$; $p=0.01$). However, even
438 though age was adjusted in our data, these results should be taken cautiously due to the
439 dramatic age differences between the AD and control groups (Figure 1), as age is an
440 important risk factor for both phenotypes. Thus, aiming to reduce age confounding in
441 our models, we restricted analysis to men aged 65–85 years old. However, despite not
442 completely correcting the age gap between the groups, this also reduced our sample size
443 (Supplementary Table 4). We found an increased degree of LOY mosaicism (mLRR-Y)
444 in 65–85-year-old cases vs. controls ($OR=2.19$; $p=0.09$), but statistical significance was
445 not reached in the models.

446 Previous reports also found that mLOY increases the rate of AD conversion in MCI
447 males [19]. We selected individuals recruited in ACE Alzheimer Center with an MCI
448 diagnosis at the moment of sampling and available clinical follow ups, and fitted Cox
449 proportional-hazards models. Even though we found risk, or positive, effect directions
450 for mLOY phenotypes towards AD progression, no statistical significance was reached
451 in the models (Figure 3). However, given our small sample size ($N=400$), we may have
452 lacked enough statistical power in this analysis. Remarkably, we noticed that the
453 quantitative mLRR-Y variable performed superiorly in the case-control and disease
454 progression models compared to mLOY calls, implying that, if effects were genuine, the
455 mLOY-induced increase in AD risk and progression may be proportional to the mosaic
456 fraction of LOY cells in blood.

457 Due to the age-dependent nature of both mLOY and AD, controlling age confounding
458 was very challenging in our cohort. For this reason, we checked mLOY causality in AD
459 by creating an instrument variable and conducted a mendelian randomization study. To
460 this end, we generated an age-independent and sex-independent mLOY-associated PRS,
461 using 114 independent autosomal genetic variants (Supplementary Table 2) previously
462 associated with mLOY [14]. Of note, mloy-PRS successfully predicted mLOY events in
463 our data and was not associated with age or *APOE* genotype (Supplementary Figure 7,
464 Supplementary Table 7). A recently published work found similar effect sizes of this
465 PRS for predicting mLOY calls [57]. Therefore, analysis of mloy-PRS instead of
466 mLOY phenotypes allowed us to overcome the main limitations of the study (age
467 differences and sample size) by providing an age-independent mLOY instrument, and
468 allowing us to increase the effective sample size in two ways: (a) removing the need to
469 restrict analysis to samples with available age at DNA sampling information and (b)
470 allowing us to introduce all MCI individuals with subsequent clinical records in disease
471 progression models instead of only those with an MCI diagnose at the closest clinical
472 evaluation to DNA sampling.

473

474 Importantly, we found a male-specific, statistically significant (HR=1.23; p=0.01)
475 association between mloy-PRS and MCI phenoconversion to AD. Case-control models
476 also reported positive, or risk, effects of mloy-PRS in AD in the male subset (OR=1.07;
477 p=0.12), but no statistical significance was reached in the models, even though our
478 sample size was considerably larger in the case-control dataset ($N_{\text{males}}=2\ 471$) than in
479 the longitudinal, prospective MCI dataset ($N_{\text{males}}=682$). These results suggest that
480 mLOY could be more involved in the MCI, early clinical stages of AD etiopathogenesis
481 than in the preclinical stages of the disease, namely AD risk. However, because the PRS
482 only explains a fraction of the variance that causes mLOY, a higher sample size may be
483 needed to reach enough statistical power to obtain more robust associations in the case-
484 control models. Importantly, mloy-PRS effects were neutral in the female groups (Table
485 3), implying that the observed effect of mloy-PRS on AD is unlikely to be driven by the
486 same mechanisms that confer mLOY risk (increased genomic instability and
487 impairment of DNA reparation mechanisms) [14]. Instead, the observed effects were
488 male-specific and therefore more likely produced via loss of the Y chromosome
489 exclusively in men.

490

491 One of the most commonly proposed mechanisms to explain blood cell LOY
492 pathogenesis is the impairment of immune functions [10, 14, 19]. Interestingly,
493 deregulation of the immune system is one of the hallmark features of AD [58], and
494 genome-wide association studies are revealing an increasing number of genes related to
495 immune functions [30]. LOY has been reported to cause deregulation in the expression
496 levels of approximately 500 autosomal transcripts in leukocytes [24]. Furthermore,
497 levels of CD99, a cell surface protein involved in several key immune functions, such as
498 leukocyte migration through the vascular endothelium, cell adhesion and apoptosis [59,
499 60], have been found to be significantly lowered in immune cells with LOY [24, 54].
500 Thus, mLOY-induced alterations in the homeostasis and migration of leukocytes
501 through the brain-blood barrier could explain the observed associations. Further studies
502 are necessary to corroborate our findings and to identify the specific mechanisms within
503 mLOY modifying AD etiopathogenesis. Of note, functional restoration of the lost Y-
504 chromosome loci promoting aberrant clonal expansion or transcriptomic deregulation of
505 LOY leukocytes could be an attractive therapeutic strategy for combating AD
506 progression.

507

508 One strength of our study is that we modelled LOY through an age-independent PRS
509 instead of just analyzing the age-dependent mLOY phenotype and adjusting our data by
510 age. In our opinion, this allowed a clearer and more robust approach for inferring
511 causality between mLOY and AD. We also obtained an independent validation of our
512 findings through AD-related biomarkers, with mLOY phenotypes associated with
513 higher levels of total tau and p-tau in the CSF and displaying the proteomic
514 neurodegenerative biochemical signature observed with other AD-related factors
515 (Figure 3). Higher tau levels are associated with faster rates of cognitive decline [61],
516 supporting the hypothesis that mLOY modulates disease progression. This work,
517 however, also faced several limitations: a) the lack of age at sampling information for
518 most controls and b) the control population being significantly younger than the AD
519 population. Both of these limitations ultimately decreased our statistical power to find
520 more robust mLOY-AD associations in the case-control models.

521

522 In summary, the present study did not find such strong associations between the blood
523 mLOY phenotype and AD as those reported previously [19]. Due to the demographic
524 features of the GR@ACE/DEGESCO cohort, with older AD patients and younger

525 population-based controls, adjusting our data by age was challenging. Consequently, we
526 modelled the genetic variance associated with mLOY risk, generating a PRS that was
527 associated with MCI conversion to AD in a male-specific manner. This approach
528 allowed us to efficiently control the effect of ageing and to evaluate the potential
529 causality of the mLOY phenotype. Furthermore, lack of association of mLOY-PRS and
530 AD in women suggests that the observed effect is produced via the inherent loss of the
531 Y chromosome and that mLOY could be a male-specific AD risk factor. Larger studies
532 may benefit from modelling mLOY using Mendelian randomization, as case and control
533 populations do not always represent the same age groups in AD cohorts, and date of
534 DNA sampling of the subjects may not be available.

535

536 ACKNOWLEDGEMENTS

537 We would like to thank patients and controls who participated in this project. The
538 present work has been performed as part of the doctoral thesis of P.G. at the University
539 of Barcelona (Barcelona, Spain). P.G. is supported by CIBERNED employment plan
540 CNV-304-PRF-866. CIBERNED is integrated into ISCIII (Instituto de Salud Carlos
541 III). I. de Rojas is supported by a national grant from the Instituto de Salud Carlos III
542 FI20/00215. A. Cano acknowledges the support of the Spanish Ministry of Science,
543 Innovation and Universities under the grant Juan de la Cierva (FJC2018-036012-I).
544 M.B. and A.R. are also supported by national grants PI13/02434, PI16/01861,
545 PI17/01474, PI19/01240 and PI19/01301. The Genome Research @ Fundacio ACE
546 project (GR@ACE) is supported by Grifols SA, Fundación bancaria La Caixa,
547 Fundació ACE, and CIBERNED. Acción Estratégica en Salud is integrated into the
548 Spanish National Research and Innovation Plan and funded by ISCIII (Instituto de Salud Carlos
549 III)—Subdirección General de Evaluación and the Fondo Europeo de Desarrollo
550 Regional (FEDER—Una manera de hacer Europa). Some control samples and data from
551 patients included in this study were provided in part by the National DNA Bank Carlos
552 III (www.bancoadn.org, University of Salamanca, Spain) and Hospital Universitario
553 Virgen de Valme (Sevilla, Spain); they were processed following standard operating
554 procedures with the appropriate approval of the Ethical and Scientific Committee.
555 Genotyping of the ACE MCI-EADB samples was performed in the context of EADB
556 (European Alzheimer DNA biobank) funded by the JPCo-fuND FP-829-029 (ZonMW
557 project number 733051061). This work was supported by a grant (European Alzheimer

558 DNA BioBank, EADB) from the EU Joint Program—Neurodegenerative Disease
559 Research (JPND).

560

561

562 REFERENCES

563

- 564 1. Alzheimer's Association. 2021 Alzheimer's disease facts and figures. *Alzheimer's*
565 *Dement.* 2021; 17:327–406.
- 566 2. Wingo TS, Lah JJ, Levey AI, Cutler DJ. Autosomal recessive causes likely in
567 early-onset Alzheimer disease. *Arch Neurol.* 2012; 69:59–64.
- 568 3. Gatz M, Reynolds CA, Fratiglioni L, Johansson B, Mortimer JA, Berg S, et al.
569 Role of genes and environments for explaining Alzheimer disease. *Arch Gen*
570 *Psychiatry.* 2006; 63:168–174.
- 571 4. Strittmatter WJ, Saunders AM, Schmechel D, Pericak-Vance M, Enghild J,
572 Salvesen GS, et al. Apolipoprotein E: High-avidity binding to β -amyloid and
573 increased frequency of type 4 allele in late-onset familial Alzheimer disease.
574 *Proc Natl Acad Sci USA.* 1993; 90:1977–1981.
- 575 5. Herrup K. Reimagining Alzheimer's disease - An age-based hypothesis. *J*
576 *Neurosci.* 2010; 30:16755–16762.
- 577 6. Lin KA, Choudhury KR, Rathakrishnan BG, Marks DM, Petrella JR,
578 Doraiswamy PM. Marked gender differences in progression of mild cognitive
579 impairment over 8 years. *Alzheimer's Dement Transl Res Clin Interv.* 2015;
580 1:103–110.
- 581 7. Holland D, Desikan RS, Dale AM, McEvoy LK. Higher rates of decline for
582 women and apolipoprotein e ϵ 4 carriers. *Am J Neuroradiol.* 2013; 34:2287–2293.
- 583 8. Snyder HM, Asthana S, Bain L, Brinton R, Craft S, Dubal DB, et al. Sex biology
584 contributions to vulnerability to Alzheimer's disease: A think tank convened by
585 the Women's Alzheimer's Research Initiative. *Alzheimer's Dement.* 2016;
586 12:1186–1196.

- 587 9. Ferretti MT, Iulita MF, Cavedo E, Chiesa PA, Dimech AS, Chadha AS, et al. Sex
588 differences in Alzheimer disease — The gateway to precision medicine. *Nat Rev*
589 *Neurol.* 2018; 14:457–469.
- 590 10. Forsberg LA, Rasi C, Malmqvist N, Davies H, Pasupulati S, Pakalapati G, et al.
591 Mosaic loss of chromosome y in peripheral blood is associated with shorter
592 survival and higher risk of cancer. *Nat Genet.* 2014; 46:624–628.
- 593 11. Jacobs PA, Brunton M, Brown WMC, Doll R, Goldstein H. Change of human
594 chromosome count distributions with age: Evidence for a sex difference. *Nature.*
595 1963; 197:1080–1081.
- 596 12. Jacobs PA, Court Brown WM, Doll R. Distribution of human chromosome
597 counts in relation to age. *Nature.* 1961; 191:1178–1180 .
- 598 13. Guo X, Dai X, Zhou T, Wang H, Ni J, Xue J, et al. Mosaic loss of human Y
599 chromosome: what, how and why. *Hum Genet.* 2020; 139:421–446
- 600 14. Thompson DJ, Genovese G, Halvardson J, Ulirsch JC, Wright DJ, Terao C, et al.
601 Genetic predisposition to mosaic Y chromosome loss in blood. *Nature.* 2019 575:
602 652–657.
- 603 15. United Kingdom Cancer Cytogenetics Group. Loss of the Y chromosome from
604 normal and neoplastic bone marrows. *Genes, Chromosom Cancer.* 1992; 5:83–
605 88.
- 606 16. Wiktor A, Rybicki BA, Piao ZS, Shurafa M, Barthel B, Maeda K, et al. Clinical
607 significance of Y chromosome loss in hematologic disease. *Genes Chromosom*
608 *Cancer.* 2000; 27:11–16.
- 609 17. Stone JF, Sandberg AA. Sex chromosome aneuploidy and aging. *Mutat Res*
610 *DNAGing.* 1995; 338:107–113
- 611 18. Wong AK, Fang B, Zhang L, Guo X, Lee S, Schreck R. Loss of the Y
612 chromosome: An age-related or clonal phenomenon in acute myelogenous
613 leukemia/myelodysplastic syndrome? *Arch Pathol Lab Med.* 2008 132:1329–
614 1332
- 615 19. Dumanski JP, Lambert JC, Rasi C, Giedraitis V, Davies H, Grenier-Boley B, et
616 al. Mosaic Loss of Chromosome y in Blood Is Associated with Alzheimer

- 617 Disease. *Am J Hum Genet.* 2016; 98:1208–1219.
- 618 20. Noveski P, Madjunkova S, Stefanovska ES, Geshkovska NM, Kuzmanovska M,
619 Dimovski A, et al. Loss of Y chromosome in peripheral blood of colorectal and
620 prostate cancer patients. *PLoS One.* 2016; 11:e0146264 .
- 621 21. Haitjema S, Kofink D, Van Setten J, Van Der Laan SW, Schoneveld AH, Eales J,
622 et al. Loss of y Chromosome in Blood Is Associated with Major Cardiovascular
623 Events during Follow-Up in Men after Carotid Endarterectomy. *Circ Cardiovasc*
624 *Genet.* 2017; 10:e001544
- 625 22. Loftfield E, Zhou W, Graubard BI, Yeager M, Chanock SJ, Freedman ND, et al.
626 Predictors of mosaic chromosome Y loss and associations with mortality in the
627 UK Biobank. *Sci Rep.* 2018; 8:12316
- 628 23. Forsberg LA. Loss of chromosome Y (LOY) in blood cells is associated with
629 increased risk for disease and mortality in aging men. *Hum Genet.* 2017; 136(5):
630 657–663.
- 631 24. Dumanski JP, Halvardson J, Davies H, Rychlicka-Buniowska E, Mattisson J,
632 Moghadam BT, et al. Immune cells lacking Y chromosome show dysregulation
633 of autosomal gene expression. *Cell Mol Life Sci.* 2021; 78:4019–4033.
- 634 25. Guo X. Loss of Y chromosome at the interface between aging and Alzheimer’s
635 disease. *Cell Mol Life Sci.* 2021; 78:7081–7084.
- 636 26. Moreno-Grau S, de Rojas I, Hernández I, Quintela I, Montreal L, Alegret M, et
637 al. Genome-wide association analysis of dementia and its clinical
638 endophenotypes reveal novel loci associated with Alzheimer’s disease and three
639 causality networks: The GR@ACE project. *Alzheimer’s Dement.* 2019; 15:1333–
640 1347.
- 641 27. de Rojas I, Moreno-Grau S, Tesi N, Grenier-Boley B, Andrade V, Jansen IE, et
642 al. Common variants in Alzheimer’s disease and risk stratification by polygenic
643 risk scores. *Nat Commun.* 2021; 12:3417.
- 644 28. Espinosa A, Alegret M, Valero S, Vinyes-Junqué G, Hernández I, Mauleón A, et
645 al. A longitudinal follow-up of 550 mild cognitive impairment patients: Evidence
646 for large conversion to dementia rates and detection of major risk factors

- 647 involved. *J Alzheimer's Dis.* 2013; 34:769–780.
- 648 29. Lacour A, Espinosa A, Louwersheimer E, Heilmann S, Hernández I, Wolfsgruber
649 S, et al. Genome-wide significant risk factors for Alzheimer's disease: Role in
650 progression to dementia due to Alzheimer's disease among subjects with mild
651 cognitive impairment. *Mol Psychiatry.* 2017; 22:153–160.
- 652 30. Bellenguez C, Küçükali F, Jansen IE, Kleineidam L, Moreno-Grau S, Amin N, et
653 al. New insights into the genetic etiology of Alzheimer's disease and related
654 dementias. *Nat Genet.* 2022; 54:412–436.
- 655 31. Boada M, Tárraga L, Hernández I, Valero S, Alegret M, Ruiz A, et al. Design of
656 a comprehensive Alzheimer's disease clinic and research center in Spain to meet
657 critical patient and family needs. *Alzheimer's Dement.* 2014; 10:409-415.
- 658 32. McKhann G, Drachman D, Folstein M, Katzman R, Price D, Stadlan EM.
659 Clinical diagnosis of alzheimer's disease: Report of the NINCDS-ADRDA work
660 group* under the auspices of department of health and human services task force
661 on alzheimer's disease. *Neurology.* 1984; 34:939–944.
- 662 33. McKhann GM, Knopman DS, Chertkow H, Hyman BT, Jack CR, Kawas CH, et
663 al. The diagnosis of dementia due to Alzheimer's disease: Recommendations
664 from the National Institute on Aging-Alzheimer's Association workgroups on
665 diagnostic guidelines for Alzheimer's disease. *Alzheimer's Dement.* 2011; 7:263–
666 269
- 667 34. Román GC, Tatemichi TK, Erkinjuntti T, Cummings JL, Masdeu JC, Garcia JH,
668 et al. Vascular dementia: Diagnostic criteria for research studies: Report of the
669 ninds-airen international workshop. *Neurology.* 1993; 43:250–260.
- 670 35. Neary D, Snowden JS, Gustafson L, Passant U, Stuss D, Black S, et al.
671 Frontotemporal lobar degeneration: A consensus on clinical diagnostic criteria.
672 *Neurology.* 1998; 51:1546–1554.
- 673 36. Rascovsky K, Hodges JR, Knopman D, Mendez MF, Kramer JH, Neuhaus J, et
674 al. Sensitivity of revised diagnostic criteria for the behavioural variant of
675 frontotemporal dementia. *Brain.* 2011; 134:2456–2477.
- 676 37. McKeith IG, Dickson DW, Lowe J, Emre M, O'Brien JT, Feldman H, et al.

- 677 Diagnosis and management of dementia with Lewy bodies: Third report of the
678 DLB consortium. *Neurology*. 2005; 65:1863–1872.
- 679 38. Petersen RC, Smith GE, Waring SC, Ivnik RJ, Tangalos EG, Kokmen E. Mild
680 cognitive impairment: Clinical characterization and outcome. *Arch Neurol*. 1999;
681 56:303–308.
- 682 39. Petersen RC. Mild cognitive impairment as a diagnostic entity. *J. Intern. Med*.
683 2004; 256:183–194.
- 684 40. Diskin SJ, Li M, Hou C, Yang S, Glessner J, Hakonarson H, et al. Adjustment of
685 genomic waves in signal intensities from whole-genome SNP genotyping
686 platforms. *Nucleic Acids Res*. 2008; 36:e126.
- 687 41. Wang K, Li M, Hadley D, Liu R, Glessner J, Grant SFA, et al. PennCNV: An
688 integrated hidden Markov model designed for high-resolution copy number
689 variation detection in whole-genome SNP genotyping data. *Genome Res*. 2007;
690 17:1665–1674.
- 691 42. Wang K, Chen Z, Tadesse MG, Glessner J, Grant SFA, Hakonarson H, et al.
692 Modeling genetic inheritance of copy number variations. *Nucleic Acids Res*.
693 2008. 21:e138.
- 694 43. Kendall KM, Rees E, Escott-Price V, Einon M, Thomas R, Hewitt J, et al.
695 Cognitive Performance Among Carriers of Pathogenic Copy Number Variants:
696 Analysis of 152,000 UK Biobank Subjects. *Biol Psychiatry*. 2017; 82:103–110.
- 697 44. González JR, López-Sánchez M, Cáceres A, Puig P, Esko T, Pérez-Jurado LA.
698 MADloy: robust detection of mosaic loss of chromosome Y from genotype-
699 array-intensity data. *BMC Bioinformatics*. 2020; 21:533.
- 700 45. Rodríguez-Santiago B, Malats N, Rothman N, Armengol L, Garcia-Closas M,
701 Kogevinas M, et al. Mosaic uniparental disomies and aneuploidies as large
702 structural variants of the human genome. *Am J Hum Genet*. 2010; 87:129–138.
- 703 46. Gogarten SM, Sofer T, Chen H, Yu C, Brody JA, Thornton TA, et al. Genetic
704 association testing using the GENESIS R/Bioconductor package. *Bioinformatics*.
705 2019; 35:5346–5348.
- 706 47. Rodriguez-Gomez O, Sanabria A, Perez-Cordon A, Sanchez-Ruiz D, Abdelnour

- 707 C, Valero S, et al. FACEHBI: A Prospective Study of Risk Factors, Biomarkers
708 and Cognition in a Cohort of Individuals with Subjective Cognitive Decline.
709 Study Rationale and Research Protocols. *J Prev Alzheimer's Dis.* 2017; 4:100–
710 108.
- 711 48. Pesini P, Pérez-Grijalba V, Monleón I, Boada M, Tárraga L, Martínez-Lage P, et
712 al. Reliable measurements of the β -amyloid pool in blood could help in the early
713 diagnosis of AD. *Int J Alzheimers Dis.* 2012. 2012:604141.
- 714 49. Adami PVM, Orellana A, García P, Kleineidam L, Alarcón-Martín E, Montreal
715 L, et al. Matrix metalloproteinase 10 is linked to the risk of progression to
716 dementia of the Alzheimer's type. *Brain.* 2022. e-pub ahead of print 28 January
717 2022; doi: 10.1093/brain/awac024.
- 718 50. R Core Team (2020). R: A language and environment for statistical computing. R
719 Foundation for Statistical Computing, Vienna, Austria. URL [https://www.R-](https://www.R-project.org/)
720 [project.org/](https://www.R-project.org/).
- 721 51. Arthur Allignol, Aurelien Latouche (2022). CRAN Task View: Survival
722 Analysis. Version 2022-03-03. URL <https://CRAN.R-project.org/view=Survival>.
- 723 52. Pierre R V., Hoagland HC. Age-associated aneuploidy: Loss of Y chromosome
724 from human bone marrow cells with aging. *Cancer.* 1972; 30:889–894.
- 725 53. Viechtbauer W. Conducting Meta-analysis in R with the metafor package. *J Stat*
726 *Softw.* 2010; 16:1–48.
- 727 54. Mattisson J, Danielsson M, Hammond M, Davies H, Gallant CJ, Nordlund J, et
728 al. Leukocytes with chromosome Y loss have reduced abundance of the cell
729 surface immunoprotein CD99. *Sci Rep.* 2021; 11:15160.
- 730 55. Kinney JW, Bemiller SM, Murtishaw AS, Leisgang AM, Salazar AM, Lamb BT.
731 Inflammation as a central mechanism in Alzheimer's disease. *Alzheimer's*
732 *Dement Transl Res Clin Interv.* 2018; 4:575–590.
- 733 56. Caceres A, Jene A, Esko T, Perez-Jurado LA, Gonzalez JR. Extreme
734 downregulation of chromosome Y and Alzheimer's disease in men. *Neurobiol*
735 *Aging.* 2020; 90: 150.e1–150.e4.
- 736 57. Riaz M, Mattisson J, Polekhina G, Bakshi A, Halvardson J, Danielsson M, et al.

- 737 A polygenic risk score predicts mosaic loss of chromosome Y in circulating
738 blood cells. *Cell Biosci.* 2021; 11:205.
- 739 58. Bettcher BM, Tansey MG, Dorothée G, Heneka MT. Peripheral and central
740 immune system crosstalk in Alzheimer disease — a research prospectus. *Nat Rev*
741 *Neurol.* 2021; 17:689–701.
- 742 59. Schenkel AR, Mamdouh Z, Chen X, Liebman RM, Muller WA. CD99 plays a
743 major role in the migration of monocytes through endothelial junctions. *Nat*
744 *Immunol.* 2002 3:143–150.
- 745 60. Vestweber D. How leukocytes cross the vascular endothelium. *Nat Rev Immunol.*
746 2015; 15:692–704.
- 747 61. Sämgård K, Zetterberg H, Blennow K, Hansson O, Minthon L, Londos E.
748 Cerebrospinal fluid total tau as a marker of Alzheimer’s disease intensity. *Int J*
749 *Geriatr Psychiatry.* 2010; 25:403–410.
- 750
- 751
- 752
- 753
- 754
- 755
- 756
- 757
- 758
- 759
- 760
- 761

762

763

764

765 **Table 1. Logistic regression results using mLRR-Y and mLOY calls as predictors for AD status.** 3 experimental setups were defined: a) all men with
766 available age at DNA sampling, b) 65–85-year-old men, and c) a stratification of 65–85-year-old men in age groups (65–70, 70–75, 75–80, & 80–85 years
767 old). Models were adjusted by age at blood sampling, *APOE* genotype, and PCs. In the age-stratified models, only the effect of mLRR-Y or mLOY are
768 displayed.

		All available males					65-85 year-old males					Stratified 65-85 year-old males					
		OR	SE	P	CI2.5	CI97.5	OR	SE	P	CI2.5	CI97.5	Age group	OR	SE	P	CI2.5	CI97.5
mLRR-Y	mLRR-Y	2.74	0.4	1.13·10⁻²	1.29	6.16	2.19	0.47	9.28·10⁻²	0.18	1.1	65-70	1.47	1.43	0.79	0.07	21.96
	Age	1.16	0.01	2.74·10 ⁻¹⁰⁴	1.15	1.18	1.27	0.01	6.61·10 ⁻⁷⁰	1.24	1.31	70-75	1.74	0.77	0.48	0.39	8.18
	APOE	2.52	0.08	1.61·10 ⁻³²	2.17	2.94	2.86	0.10	5.40·10 ⁻²⁸	2.38	3.46	75-80	2.37	0.77	0.27	0.57	12.16
	PC1	2.01·10 ³	2.6	3.39·10 ⁻³	12.48	3.29·10 ⁵	0.01	3.33	0.11	0.00	3.48	80-85	44.44	2.51	0.13	0.99	2.63·10 ⁴
	PC2	0.02	2.58	0.15	0.00	3.8	0.00	3.45	0.12	0.00	4.00	META	2.20	0.50	0.115	0.83	5.88
mLOY calls	mLOY	1.14	0.14	0.35	0.87	1.49	1.04	0.16	0.81	0.76	1.43	65-70	0.55	0.45	0.19	0.22	1.30
	Age	1.16	0.01	1.11·10 ⁻¹⁰⁶	1.15	1.18	1.28	0.01	1.73·10 ⁻⁷¹	1.24	1.31	70-75	1.17	0.3	0.6	0.65	2.11
	APOE	2.52	0.08	2.17·10 ⁻³²	2.17	2.94	2.86	0.10	6.93·10 ⁻²⁸	2.37	3.46	75-80	1.00	0.25	0.99	0.62	1.66
	PC1	2.29·10 ³	2.59	2.78·10 ⁻³	14.5	3.69·10 ⁻⁵	0.00	3.32	0.10	0.00	2.96	80-85	3.70	0.77	0.09	1.00	24.08
	PC2	0.02	2.58	0.14	0.00	3.44	0.00	3.44	0.11	0.00	3.34	META	1.03	0.17	0.86	0.74	1.45

769 **Table 2. Cox proportional-hazards model results for conversion of MCI males to all-cause-dementia or AD-dementia.** Models were
 770 adjusted by age at sampling and *APOE* genotype.

771

	Conversion to all-cause-dementia				Conversion to AD-dementia			
	HR	SE	P-value	CI95 %	HR	SE	P-value	CI95 %
<i>Association results for mLRR-Y (continuous mLOY variable)</i>								
mLRR-Y	1.93	0.40	0.10	0.87-4.27	2.05	0.55	0.19	0.70-5.97
Age	1.10	0.01	$3.62 \cdot 10^{-15}$	1.07-1.12	1.13	0.02	$2.38 \cdot 10^{-12}$	1.09-1.16
APOE	1.29	0.13	$4.10 \cdot 10^{-02}$	1.01-1.66	1.56	0.17	$8.89 \cdot 10^{-3}$	1.12-2.17
<i>Association results for mLOY calls</i>								
mLOY	1.17	0.19	0.40	0.81-1.70	1.38	0.25	0.20	0.85-2.24
Age	1.10	0.01	$1.33 \cdot 10^{-15}$	1.07-1.13	1.13	0.02	$1.62 \cdot 10^{-12}$	1.09-1.17
APOE	1.28	0.13	$5.09 \cdot 10^{-2}$	1.00-1.63	1.54	0.17	$1.05E \cdot 10^{-2}$	1.11-2.13

772

773

774

775 **Table 3. Association results for mLOY PRS.** (a) Results of logistic regressions for case-control AD. Only individuals aged 65–85 were included in the
 776 models. Models were adjusted by *APOE*, age, and principal components. (b) Results of joint analysis of prospective MCIs in the GR@ACE-DEGESCO and
 777 EADB-DEGESCO cohorts using Cox proportional-hazards models for progression from MCI to all-cause dementia or AD. Models were adjusted by *APOE*,
 778 age, and cohort ascertainment.

a. Logistic regression results:

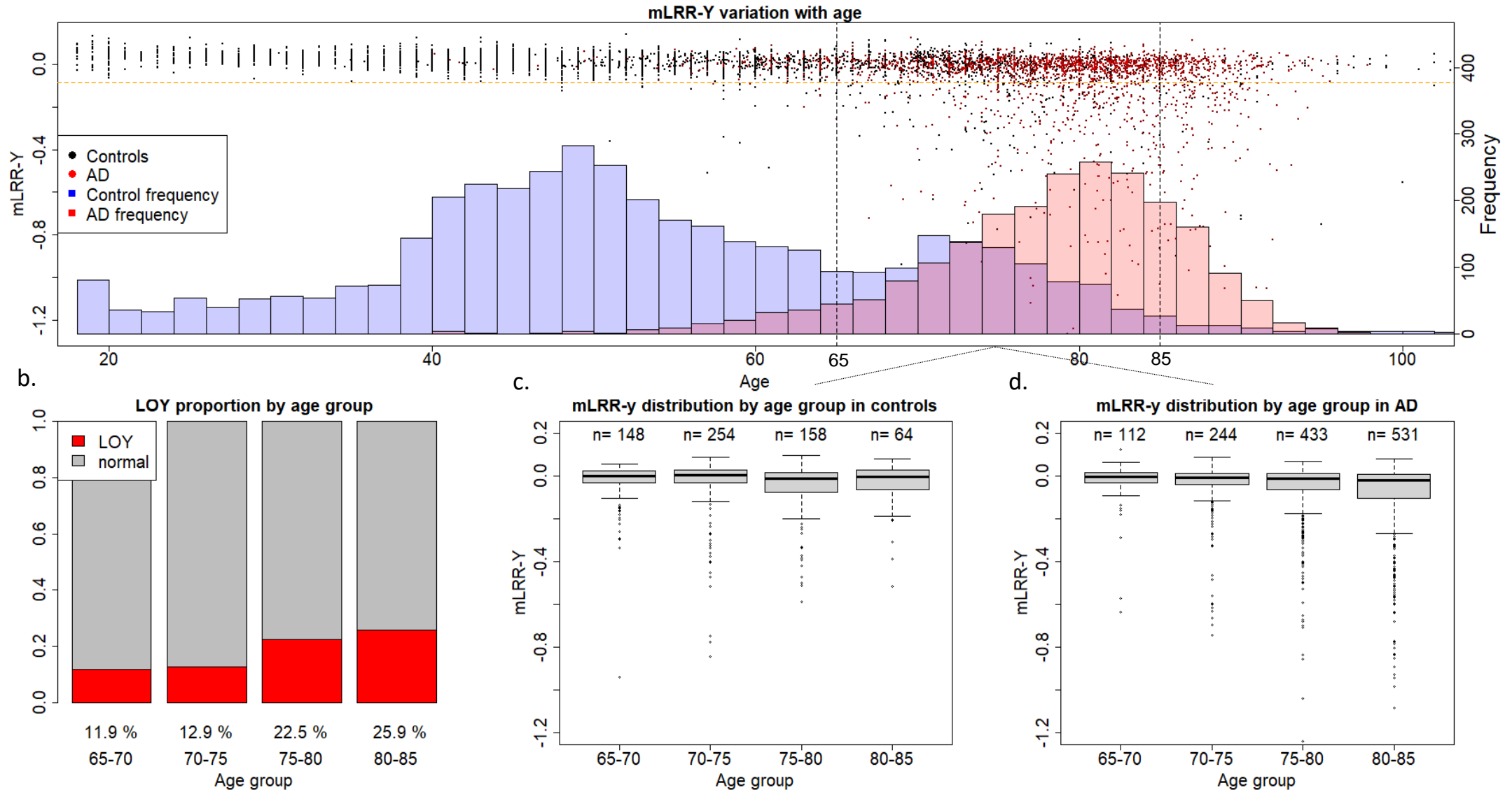
	OR	SE	Z	P-value	2.5 % CI	97.5 % CI
CASE-CONTROL, ALL	1.03	0.03	1.09	0.28	0.98	1.09
CASE-CONTROL, MALES	1.07	0.05	1.55	0.12	0.98	1.18
CASE-CONTROL, FEMALES	1.00	0.03	0.10	0.93	0.90	1.10

779

b. Cox model results:

	HR	SE	Z	P-value	2.5 % CI	97.5 % CI
ALL, MCI to dementia	1.04	0.04	0.97	0.33	0.96	1.12
ALL, MCI to AD	1.07	0.05	1.45	0.15	0.98	1.16
MALES, MCI to dementia	1.11	0.06	1.77	7.68·10⁻²	0.99	1.26
MALES, MCI to AD	1.23	0.08	2.53	1.14·10⁻²	1.05	1.43
FEMALES, MCI to dementia	0.99	0.05	-0.23	0.81	0.90	1.08
FEMALES, MCI to AD	0.99	0.06	-0.19	0.85	0.89	1.11

Figure 1. mLRR-y variation with age in the GR@ACE-DEGESCO cohort. a. Age and mLRR-Y distribution in the case and control groups. The dots represent mLRR-Y values for individual samples and the histogram represents the age distribution across the case and control groups **b.** Proportion of individuals with mLOY in the different age groups based on age at blood sampling. **c–d** mLRR-Y distribution for the different age groups based on age at blood sampling in control and AD individuals, respectively.



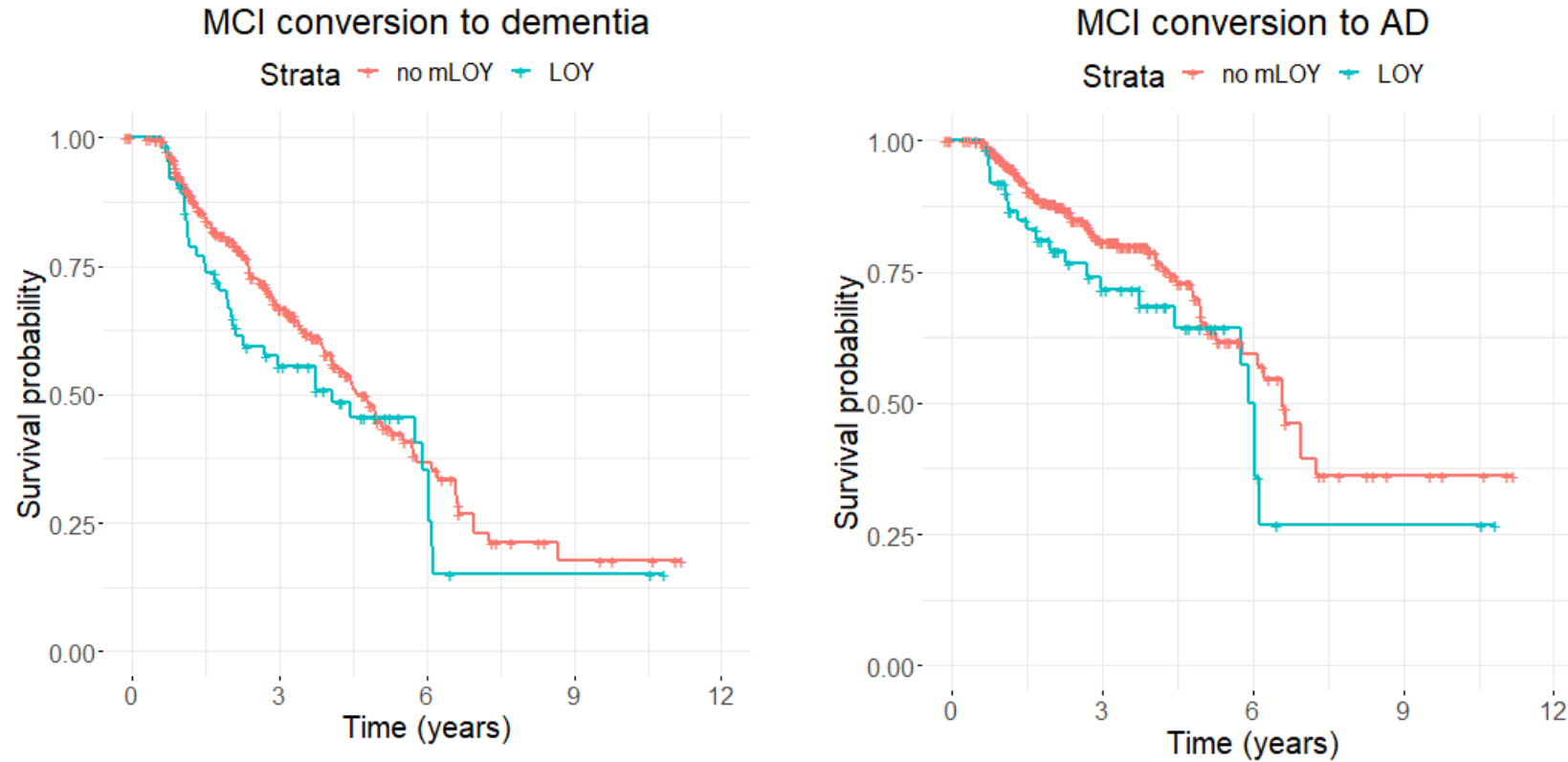


Figure 2. Association of mLOY phenotypes with risk of conversion to dementia and AD-dementia over time for MCI males in the GR@ACE-DEGESCO cohort. Kaplan-Meier plots showing survival time in years for conversion to dementia (a) or AD (b) for individuals with LOY (blue) or without LOY (red) for prospective MCI males in the GR@ACE-DEGESCO cohort.

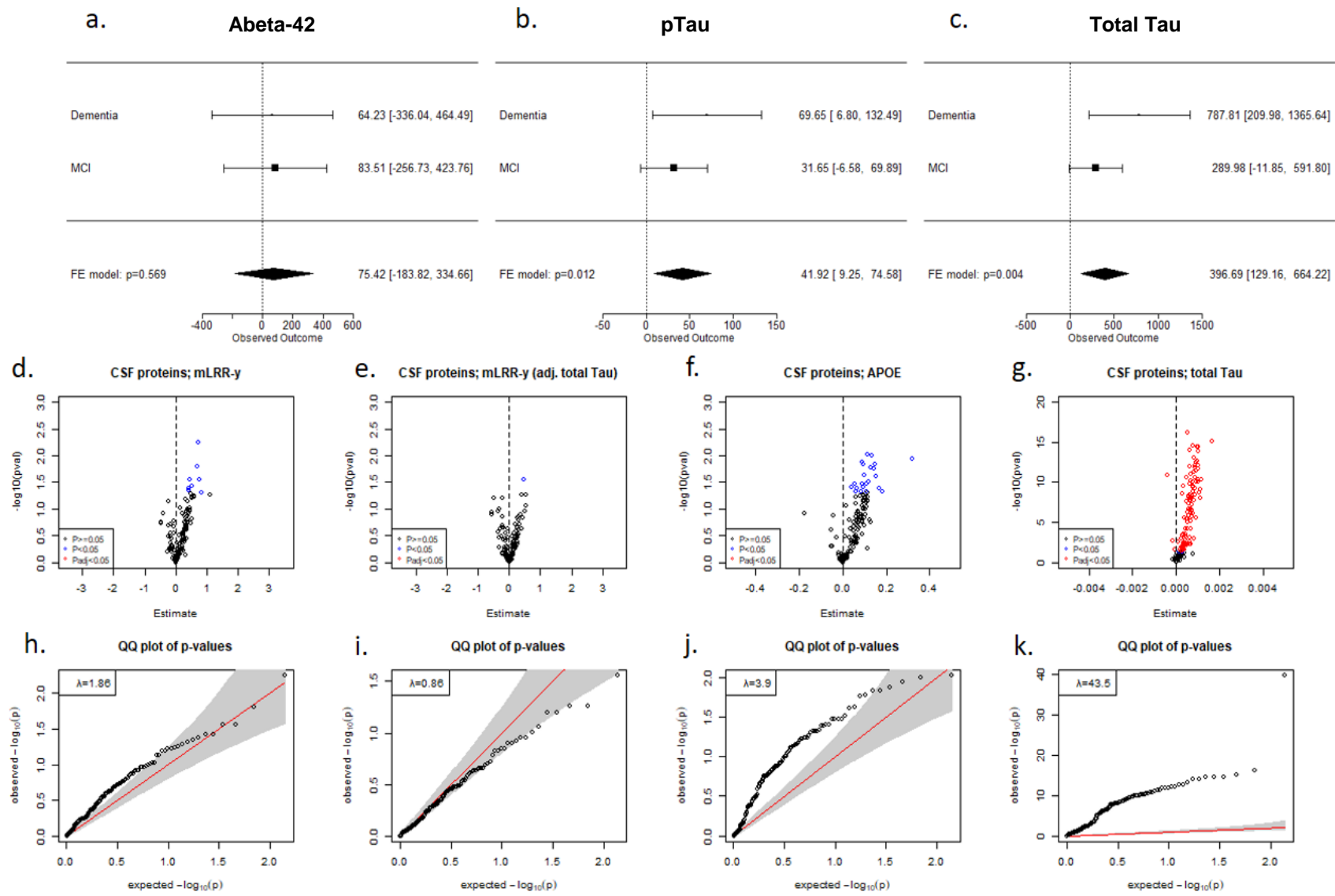


Figure 3. mLRR-Y associations with CSF protein levels. a-c. Forest plots showing the effect size obtained in linear regression models for mLRR-Y on Abeta-42 (a), phospho-Tau (b), and total Tau (c) in males with MCI or dementia, along with meta-analysis results. **d-g.** Volcano plots showing association of CSF proteins in the Olink Inflammation and Neurology panels with mLRR-Y (d), mLRR-Y adjusted by total Tau (e), APOE genotype (f), and total Tau (g). **h-k.** QQ plots obtained in the models for mLRR-Y (h), mLRR-Y adjusted by total Tau (i), APOE genotype (j), and total Tau (k). Models were adjusted by age, the time window between CSF and DNA sampling, and APOE genotype.