Lactoferrin levels in cerebrospinal fluid exhibits isoform-specific associations with Alzheimer's disease

Feiyang Zhao[a,b], Raquel Puerta[c], Yaxi Wang[d], Eva Beckett[a], Pablo Garcia Gonzalez[c,e], Sergi Valero[c,e], Adelina Orellana[c], Pilar Sanz[c], Tiffany F. Kautz[a], Jose E. Cavazos[a,f], Maria Victoria Fernandez[c], Amanda Cano[c,e], Sudha Seshadri[a,g,h], Merce Boada[c,e], Valentina R. Garbarino[a,i]*, Agustin Ruiz[a,b,c,e]*

[a]Glenn Biggs Institute for Alzheimer's & Neurodegenerative Diseases, The University of Texas at San Antonio, San Antonio, TX, USA

[b]Department of Microbiology, Immunology and Molecular Genetics. Long School of Medicine, The University of Texas at San Antonio, San Antonio, TX, USA

[c]Ace Alzheimer Center Barcelona-Universitat Internacional de Catalunya, Barcelona, Spain

[d]Department of Biochemistry and Structural Biology. Long School of Medicine. University of Texas San Antonio, TX, USA

[e]CIBERNED, Network Center for Biomedical Research in Neurodegenerative Diseases, National Institute of Health Carlos III, Madrid, Spain

[f]South Texas Medical Science Training Program, University of Texas Health San Antonio, San Antonio, TX, USA

[g]School of Public Health, Boston University and the National Heart, Lung, and Blood Institute Framingham Heart Study, Boston, MA, USA

[h]Department of Neurology, Boston University School of Medicine, Boston, MA, USA

[i]Department of Cell Systems & Anatomy, The University of Texas at San Antonio, San Antonio, TX, USA

*Co-corresponding authors: Agustin Ruiz MD, PhD (ruiza5@uthscsa.edu), Valentina R. Garbarino, PhD (garbarino@uthscsa.edu)

ABSTRACT

Background: Pathological changes in Alzheimer's disease (AD) begin years before the onset of clinical symptoms. Developing cost-effective and minimally invasive biomarkers for preclinical diagnosis remains a critical goal in the field. Lactoferrin, an iron-binding glycoprotein, has emerged as a promising candidate due to its multifunctional roles reported in previous studies. However, whether lactoferrin levels in biofluids are associated with established AD biomarkers, and whether it can serve as a reliable diagnostic indicator, remains controversial.

Methods: We analyzed SOMAscan proteomic data from 1,367 paired plasma and cerebrospinal fluid (CSF) samples from the ACE Alzheimer's Center Barcelona cohort to evaluate lactoferrin levels. Associations between two lactoferrin-targeting SOMAmers and classical AD biomarkers, including total tau (t-tau), phosphorylated tau (p-tau181), and amyloid-beta 42 (Aβ42), were assessed. The age, sex, proteomic principal components were considered as covariates for sensitive analysis. The prognostic value of lactoferrin in predicting dementia progression was further evaluated using survival analysis.

Findings: Among the two lactoferrin-targeting SOMAmers, Seq.2780.35 (LTF2) showed a weak and exclusive association with CSF Aβ42 and syndromic status, whereas Seq.14755.4 (LTF1) was weakly associated with CSF p-tau and t-tau AD biomarker levels displayed expression-dependent stratification consistent with a ventricular-volume–related dilution effect rather than disease. Furthermore, lactoferrin levels were not significantly associated with the progression from mild cognitive impairment (MCI) to dementia.

Interpretation: Isoform-specific lactoferrin expression changes in cerebrospinal fluid (CSF), but not plasma, appears to have biological relevance and diagnostic biomarker potential for AD.

Keywords: Alzheimer's disease, Lactoferrin, CSF, Plasma, Biomarkers

Introduction

Alzheimer's disease (AD) is the most common cause of dementia in the elderly, which is characterized by progressive cognitive decline, loss of neurons, amyloid-beta (Aβ) deposition, and tau pathology (1, 2). Combinations of clinical assessment, structural /functional neuroimaging, and evaluation of cerebrospinal fluid (CSF) based pathological markers, are considered well-established AD biomarkers and show a good diagnostic precision to detect AD progression (2, 3). Despite recent advancements in diagnostic tools, these CSF and imaging-based markers are either invasive or expensive, highlighting an urgent need for new diagnostic biomarkers to address these limitations. Furthermore, the underlying cause of neurodegenerative disease remains unclear. Even the widely accepted hypothesis, that Aβ aggregation leads to a pathological cascade, cannot alone fully explain the pathological initiation and progression of AD (3, 4). In the last decades, there has been growing recognition that immune system declines during aging and persistent inflammation induced by microbes contribute to Aβ deposition, progressive synaptic dysfunction, neuronal loss, and ultimately AD (5). In support of this hypothesis, infections of the central nervous system are known to be associated with AD-like pathology (6-10). Furthermore, oral and gut microorganisms are now known to play a role in the cascade of immune dysregulating events preceding AD (11, 12). The involvement of pathogen-targeting agents and viral infection–related markers in the brain such as amyloid aggregation, reinforces the potential association between AD and immunologic response of the central nervous system to infection(s) (12, 13). These pieces of evidence suggest an important role of both local and systemic inflammation towards AD pathogenesis.

As primary effector molecules of innate immunity, antimicrobial proteins and peptides may contribute significantly to AD development (14, 15), specifically Aβ pathology (16-19). Lactoferrin (LTF), an iron-binding glycoprotein involved in innate immunity, acting as part of the body's first line of defense against bacterial, viral, fungal, inflammatory and

85 carcinogenic compounds (20), has attracted attention in recent years for its potential role
86 in neurodegeneration (21-23). Beyond exocrine secretion by epithelial cells in most
87 mucosal tissues, it is also expressed by the cells of the innate immune system throughout
88 the body, including peripheral neutrophil leukocytes and monocytes, and microglial within
89 the brain (24, 25). In neutrophils, LTF is stored in the secondary granules and released
90 into the blood or infected tissues at high concentrations during inflammation. LTFs
91 expression and release occur not only in peripheral tissues but also in the central nervous
92 system, where microglial cells can release LTF through degranulation, particularly under
93 inflammatory conditions (25, 26).

94 In the last decade, several studies have reported altered LTF abundance in biological
95 fluids of AD patients. For instance, through a cross-sectional investigation, Carro, et al.
96 found that salivary LTF expression can discriminate AD patients from healthy controls
97 more accurately than Aβ42 and total-tau (t-tau) in CSF (27). The same research group
98 further revealed that salivary LTF can reliably distinguish prodromal AD and clinical AD
99 groups from frontotemporal dementia (FTD) with high accuracy (28). Through a later study,
100 they also demonstrated negative associations between salivary LTF and Aβ load, middle
101 temporal cortex thickness, increased FDG uptake in posterior cingulate cortex, and
102 poorer memory (29). Studies from this research group have consistently validated their
103 previous findings, that salivary LTF may be a promising non-invasive and cost-effective
104 AD biomarker. However, a subsequent, independent follow-up study in a mixed memory
105 clinic cohort did not find an association between LTF levels in CSF or saliva and cognitive
106 diagnoses or pathological AD markers (t-tau, p-tau, Aβ) (30). These conflicting findings
107 from different research groups highlight the need for standardization across sample
108 collection and assessment methods, and suggest that LTF may have a complex, context-
109 dependent relationship with specific neurodegenerative disease pathology and
110 progression—relationships which are not fully captured or detectable by current studies
111 and evaluation methods. Disagreement between these studies may stem from several
112 complex aspects of neurodegenerative disease pathology and the biology of LTF that is
113 endogenously generated by a variety of cell types including: immune cell, glial cells,
114 epithelial cells (25, 31). LTF not only exists in the neural system but also within the
115 immune, digestive, circulatory and reproductive systems, so it is important to consider

that circulating LTF levels are potentially modulated by multiple tissues, environmental exposures and potentially even some genetic factors.

To explore the potential involvement of LTF in AD and strengthen our understanding of LTFs role in AD, we examined data from a novel and independent cohort. We utilized existing proteomic data from the ACE Alzheimer Center Barcelona (Spain), which includes a large number of participants with paired CSF and plasma proteomic profiles measured by the high-throughput SOMAscan 7K platform. By leveraging this well-characterized dataset along with readily available detailed clinical phenotyping, we aimed to clarify the relationship between LTF and AD across CSF and plasma, and to assess whether LTF holds promise as a disease-associated or stage-specific biomarker in these two biofluids. This study provides an opportunity to re-evaluate the role of LTF in AD pathogenesis with improved statistical robustness and biological resolution.

2. Material and methods

2.1. Study Cohort

We used the ACE Memory Clinic in Barcelona, Spaine cohort of well characterized CSF and matched plasma samples for all analysis in this study. Participants who presented for cognitive evaluation and met criteria for lumbar puncture were recruited. On the same day, both biological samples were collected, along with clinical, cognitive, and imaging data. Follow-up assessments were conducted annually. The cohort consisted of individuals with MCI, dementia, or subjective cognitive concerns. Diagnoses were made by a multidisciplinary team based on a comprehensive neuropsychological evaluation. Progression to dementia was monitored over time to classify participants as either converters or stable. Converters are defined as subjects who converted to dementia, including AD, vascular dementia, mixed dementia (AD with cerebrovascular disease), frontotemporal dementia, or dementia with Lewy bodies over the study period (from the baseline assessment to last available visit; max = 7.57 years, mean = 2.81 years), and were classified as MCI converters according to previous definitions (32). All converter subjects had a Clinical Dementia Rating (CDR) score of 1. In contrast, those subjects who remained stable at follow-ups were classified as non-MCI converters. Sample processing followed standardized procedures (33). Briefly, plasma centrifuged from blood, as well as CSF were stored at −80 °C for later use. Core AD CSF biomarkers

147 (Aβ40, Aβ42, t-tau, and p-tau181) were quantified using validated immunoassay
148 platforms, with predefined thresholds used for AT(N) classification according to the 2018
149 NIA-AA Research Framework (34). For details on sample acquisition and collection of
150 additional information, refer to previous publications (35-38).

151 2.2 Protein analysis in CSF and plasma using the SOMAscan panel

152 Paired CSF and plasma proteomic data from 1,369 ACE cohort participants were
153 generated using the SOMAscan 7K Assay Kit in the context of the HARPONE project and
154 the Global Neurodegeneration Proteomics Consortium (36-38). Briefly, this platform
155 measures approximately 7,000 proteins per sample. This assay uses 50 µL of either CSF
156 or plasma and employs modified DNA aptamers (SOMAmers) to quantify specific protein
157 levels. Initially, proteins captured by immobilized aptamers, facilitated by streptavidin
158 beads, and tagged with fluorescent markers. After removal of unbound proteins, the
159 streptavidin beads are released using ultraviolet light. The protein-aptamer complexes
160 were re-captured by monomeric avidin. The detail of the assay is described elsewhere
161 (39). Data was reported in normalized relative fluorescence units (RFU) values using the
162 adaptive normalization by maximum likelihood (ANML) approach, following previously
163 described procedures (39). For the LTF gene (UniProt ID: P02788), two SOMAmers -
164 Seq.14755.4 and Seq.2780.35 - were designed on the SOMAscan platform based on the
165 target coding information provided by SomaLogic. These two SOMAmers have different
166 dilution factors: 20% for Seq.14755.4 and 0.005% for Seq.2780.35 (**Supplement Table**
167 **S1**). For clarity, we henceforth refer to Seq.14755.4 as LTF1 and Seq.2780.35 as LTF2
168 throughout this manuscript.

169 2.2. Statistical methods

170 a. Identification of the association between LTF expression and dementia diagnosis

171 Prior to statistical analysis, SOMAmer measurements were log-transformed and
172 standardized using z-score normalization due to the long tail of the SOMAmer signal
173 among participants. The normality of LTF1/2 distribution was evaluated with Shapiro.test
174 function in R. To examine the association between LTF levels and AD-related biomarkers,
175 the Spearman test was applied to compare the distributions of LTF SOMAmers, given
176 their non-normal distribution. For LTF1, participants were further categorized into high

177 (LTF1-high) and low (LTF1-low) groups to evaluate group-wise differences. Spearman
178 correlation analysis was used to assess the relationships between LTF levels (in CSF and
179 plasma), AD CSF biomarkers expression, and AD risk loci identified in a previous
180 genome-wide association study (40). To address potential inaccuracies in p-values arising
181 from tied ranks in Spearman correlation tests, genotype dosages were transformed into
182 categorical values (0, 1, or 2), followed by Fisher's exact test using the fisher.test function
183 with simulate.p.value = TRUE and B = 10,000. All statistical analyses were performed
184 using R (v4.3). To evaluated whether the confounding factor affects the association
185 between LTF1 and AD biomarkers (Aβ42, p-tau181), linear models were fit with sex, age,
186 age2, principal components and reference genes GAGE2A (seq.18268.5) or OPCML
187 (seq.15622.13) signal as covariates. To perform principal components analysis, we filter
188 out samples with more than 10% protein that have outlier abundance (absolute value ≥
189 mean ± 3 SD), the outlier value left are imputed using function imputPCA in R package
190 missMDA (v: 1.20). PCA was calculated using function prcomp in R package stats (4.5.1).

191 b. Assessment of associations between clinical traits and LTF protein expression

192 Patients were categorized into case and control groups based on AT(N) biomarker
193 classification (status). Individuals with an A-T-(N)- profile were assigned to the control
194 group, while those with A+ (regardless of T or N status) were classified as cases. For
195 further comparison, we also compared A+ v.s. A-, A+T+ v.s. A-T-, A+T+ v.s. A-T- (include
196 A-T-(N)-). Participants were classified into three diagnostic categories: cognitively normal
197 controls (including subjective cognitive decline (SCD)), mild cognitive impairment (MCI),
198 and dementia. Due to the non-linear relation between LTF and CSF biomarkers, we
199 categorize the samples into high and low according to the signals of LTF1 (High: > mean
200 + SD; Low: < mean + SD) for correlation analysis. For LTF2, the outliers were Winsorized
201 (mean ± 3 standard deviation).

202 c. Identification of the LTF central nervous system and systemic co-regulatory gene
203 networks

204 To compare the LTF-associated genes with known co-regulated expression profiles,
205 gene clusters containing 237 brain-expressed genes and 154 bone marrow–expressed
206 genes were retrieved from the Human Protein Atlas database (41). LTF-associated
207 proteins were identified using Spearman correlation analysis via the cor.test function in R.

Genes with an absolute Spearman correlation coefficient greater than 0.1 and ranked among the top 250 proteins by correlation were selected for the positively (+) and negatively (-) correlated groups, respectively. Multiple testing correction was applied using the false discovery rate (FDR) method. Fisher's exact test was used to assess enrichment by comparing the identified genes from the proteomic data with gene sets from the Human Protein Atlas database.

d. Cox proportional hazards models

We applied a series of Cox proportional hazards regression models to assess the association between LTF1/LTF2 and the risk of conversion to dementia. Three hierarchical models were constructed to evaluate the robustness of the associations under different levels of covariate adjustment: Model 1 (unadjusted model): LTF1 or LTF2 signal (LTF) was entered independently to estimate its association with dementia conversion. Model 2 (demographic-adjusted model): Each variable was included in a Cox model adjusted for age at lumbar puncture (Age_LP), sex, years of education (Education), and baseline Mini-Mental State Examination (MMSE) score. Model 3 (Fully adjusted model): In addition to the covariates in Model 2, we further adjusted for polygenic risk score (PRS), CSF phosphorylated tau (p-tau181) levels, and APOE genotype. Statistical tests were performed between the LTF1-high and -low groups. For LTF2, statistical significance was assessed by comparing samples with abundance above versus below the median. Hazard ratios (HRs), 95% confidence intervals (CIs), and p-values were reported for each model. Survival time was defined as the interval from MCI to dementia conversion or last follow-up assessment.

3. Results

3.1 LTF abundance in CSF and Plasma

To investigate whether LTF changes contribute to AD progression and the possibility to serve as a potential biofluid biomarker for AD, we accessed the plasma and CSF profiles of 1,369 individuals from the ACE CSF cohort. In this cohort, two samples lacked detectable signals for both LTF-targeting SOMAmers. Analysis included a total of 792

237 females and 577 males. The subjects consisted of 113 (8.27%) healthy controls; 833
238 (60.9%) diagnosed with MCI at baseline–of whom 44.1% converted to dementia during
239 the follow-up visit, with CSF samples collected within 3 months of first visit–and 394
240 (28.2%) subjects diagnosed with dementia at the time of the lumbar puncture
241 (**Supplement Table S2**). The mean age at the time of lumbar puncture was 72.7 ± 8.8
242 years. As expected, the frequency of the *APOE* ε4 allele was higher in the dementia group
243 (21.9%) compared to MCI group (19.9%) and control group (13.1%). Detailed
244 demographic and clinical characteristics are summarized in **Supplement Table S2**. We
245 first investigated the distribution of all SOMAmers and noticed that distribution patterns of
246 most SOMAmers did not follow a normal distribution including LTF1/2 in both CSF and
247 plasma (**Fig. 1A**). For each individual, two SOMAmers that target LTF were measured in
248 both CSF and plasma. For LTF1 (seq.14755.4) most individuals exhibited low expression
249 in both CSF (132 ± 79) and plasma (370 ± 469) , while a small subset showed high
250 expression in either fluid (**Fig. 1A, B**). The other SOMAmer (LTF2; seq.2780.35)
251 displayed high expression in plasma (1,260 ± 4,703) for a small group of individuals (**Fig.
252 1A, C)** and another group of samples showed CSF-specifically high expression (CSF:
253 19,951 ± 20,804). This result indicates that LTF1 and LTF2 possibly represent two
254 different isoforms or different conformational/binding status of LTF. To further confirm this
255 hypothesis, we compared LTF expression between the two SOMAmers. The correlation
256 analysis reveals the signal between the two SOMAmer shows weak significant correlation
257 only in plasma (Spearman's $\rho = 0.22$; $p < 0.001$), while there was almost no significant
258 correlation in CSF (Spearman's $\rho = 0.01$; $p > 0.05$; **Fig 1D**). Meanwhile, results revealed
259 no significant correlation between CSF and plasma (Spearman's $\rho = 0.02$; $p > 0.05$) for
260 LTF2, with a similarly negligible negative correlation for LTF1 (Spearman's $\rho = -0.09$; $p <$
261 0.01; **Fig 1D**). To examine whether the distribution of LTF1 in CSF is associated with
262 blood-brain barrier integrity, we performed Spearman correlation analysis between LTF1,
263 LTF2 and the albumin quotient (Qalb; a measure of blood–brain barrier (BBB) leakage),
264 revealing that only CSF LTF1 levels (Spearman's $\rho = 0.1$; $p < 0.001$; **Fig. 1E**), not LTF2
265 (Spearman's $\rho = -0.01$; $p > 0.05$), weakly associated with Qalb.

266

3.2 LTF Levels and AD Biomarkers

We sought to determine whether LTF expression is associated with established CSF AT(N) biomarkers in the ACE CSF cohort. To address this, we performed Spearman correlation analysis between LTF1 and LTF2 signals with Aβ42, p-tau181 and t-181 (**Fig 1E**). We found that LTF1 in CSF showed a weak positive correlation with total-tau (Spearman's $\rho = 0.16$; $p < 0.001$), p-tau181 (Spearman's $\rho = 0.16$; $p < 0.001$) and a negligible correlation with Aβ42 (Spearman's $\rho = 0.07$; $p < 0.05$). In contrast, LTF2 showed a weak negative correlation with Aβ42 (Spearman's $\rho = -0.10$; $p < 0.001$) and no significant association with p-tau181 (Spearman's $\rho = 0.01$; $p > 0.05$) or t-tau (Spearman's $\rho = 0.01$; $p > 0.05$). Based on the distribution patterns of LTF1 and LTF2 in CSF and plasma (**Fig. 1A**), we hypothesized that the weak correlation might reflect the heterogeneity in expression across subgroups of participants. LTF1 displayed a non-normal distribution, suggesting it may be expressed in only a subset of individuals. LTF2 appears to follow a normal distribution, with a few distinct outliers. To test this hypothesis, we stratified the samples based on LTF1 signal intensity: high (≥ mean + SD) and low (< mean + SD) and remove the outliers for LTF2. We then repeated the correlation analysis. For LTF2, results remained unchanged (**Fig 1F**; **Fig S1**). In the LTF1-low group, the associations with AD biomarkers were stable (**Fig 1G**; **Fig S1**), however, for LTF1-high groups, the direction of the associations reversed, and correlation coefficient increased, particularly for Aβ (Spearman's $\rho = -0.175$), and p-tau181 (Spearman's $\rho = -0.21$, **Fig 1F**; **Fig S1**). Notably, these coefficients were slightly stronger than those observed between AD polygenic risk score and AD biomarkers (variance explained increase by ~2.0% for p-tau181, 0.2% for t-tau, and 2.2% for Aβ42). Our recent work showed that variability in CSF proteomes is dominated by a ventricular-volume–related dilution axis, implying that CSF biomarker signals can be confounded by physiological variability (42). To investigate its impact on the LTF–AD-biomarker relationship, we first examined correlations of LTF1/LTF2 with the first two proteomic principal components (PC1–PC2) and with reference genes (GAGE2A, OPCML). The result reveals LTF1 correlated with PC1 in the low-expression group (Spearman's $\rho = -0.373$, $p < 0.001$; **Fig. S2**) but not in the high-expression group ($\rho = 0.009$, $p > 0.05$; Fig. S2). LTF2 showed no association with PC1 (Spearman's $\rho = -0.004$, $p > 0.05$; **Fig. S2**)

298 To assess the influence of potential confounders, we re-estimated the associations
299 between CSF LTF1/LTF2 and AD biomarkers using linear models that adjusted for sex,
300 age, age², proteomic principal components (PC1–PC2), and reference-gene abundance
301 (**Supplementary Tables S3–S6**). After adjusting for PC1, LTF1 was not associated with
302 Aβ42 ($\beta = -0.005$; $p = 0.089$; **Supplementary Table S3**), but it remained negatively
303 associated with p-tau ($\beta = -0.167$; $p < 0.001$; **Supplementary Table S3**). In stratified
304 analyses, LTF1 continued to associate with p-tau in the high-expression group after
305 PC1 adjustment ($\beta = -0.144$; $p = 0.003$), whereas the apparent positive association in
306 the low-expression stratum ($\beta = 0.383$; $p < 0.001$) disappeared after additional
307 adjustment for reference genes ($\beta = 0.010$; $p = 0.837$) and show negative association
308 after adjustment for PC1 ($\beta = -0.167$; $p = 0.001$). For LTF2, the association with Aβ42 (β
309 $= -0.112$; $p < 0.001$) was partially attenuated by reference-gene adjustment ($\beta = -0.087$;
310 $P < 0.001$) but was unchanged by PC1 ($\beta = -0.113$; $p < 0.001$) (**Supplementary Table**
311 **S6**). Collectively, these results suggest that the LTF1 signal in the low-expression group
312 largely biased by physiological (dilution) variation before adjustment, whereas variation
313 in high-expression LTF1 and in LTF2 is more likely to be pathologically relevant.

314 We then compared the LTF1 and LTF2 signal levels across participants with different
315 pathological features. The results showed that LTF2 levels in CSF were significantly
316 associated with AD status defined by AT(N) biomarker ($p = 0.0051$; **Fig S3**) and show
317 trend toward an association with clinical symptom syndromic state (Kruskal-Wallis, $p =$
318 0.050; **Fig S3**), while LTF1 levels were not. To determine the biological relevance of these
319 SOMAmer signals, we performed pathway enrichment analysis on the top co-regulated
320 proteins for LTF1 and LTF2. For LTF1 (in both plasma and CSF), enriched terms could
321 not be found before stratification, however, in the LTF1-high group, we observed
322 enrichment for neural-related pathways in the plasma analysis, including “neurogenesis”,
323 “generation of neurons”, “neuron projection development”, and “neuron differentiation”
324 (**Fig 2A**). Proteins correlated with LTF1 in CSF were enriched in immune-related
325 categories such as “blood microparticle” and “complement and coagulation cascades”
326 (**Fig 2C**). To further explore the differences of high and low populations, we analyzed the
327 LTF-associated protein networks using SomaScan data. Correlation analyses between
328 LTF1/LTF2 and all other quantified proteins were performed, and significantly coregulated
329 proteins were compared with LTF interactors identified in single-cell RNA-seq data from

330 brain and bone marrow. Before stratification, no enrichment overlap was found for LTF1
331 in CSF or plasma. However, in the positive group, significant overlap was observed with
332 bone marrow–derived co-regulated genes (**Fig 3**).

333 LTF2 showed modest enrichment for terms like “secretory granule lumen” and
334 “vesicle lumen” (**Fig. 2B**). Overlap with bone marrow-derived networks was also
335 confirmed. Together, these findings suggest that LTF1 and LTF2 are associated with
336 distinct proteomic profiles and biological processes, possibly reflecting different aspects
337 of the LTF locus biology and AD pathology. Importantly, the diagnostic relevance of LTF1
338 appears to be limited to specific subgroups of samples that have higher expression for
339 LTF1 signal.

340

341 3.3 CSF and Plasma LTF Levels Are Not Associated with AD Progression or 342 Genetic Risk

343 We next examined whether LTF expression is associated with the progression from
344 mild cognitive impairment (MCI) to dementia. Using Cox regression models, we found
345 that neither LTF1 nor LTF2 expression in cerebrospinal fluid (CSF) or plasma was
346 significantly associated with disease progression. To account for potential confounding
347 variables, we implemented nested Cox regression models. The second model included
348 adjustments for age, sex, years of education, and MMSE score, while the third model
349 further adjusted for polygenic risk score (PRS), p-tau181 levels, and APOE genotype.
350 Across all models, LTF1 and LTF2 levels in both CSF and plasma were not significantly
351 correlated with conversion from MCI to dementia (**Fig. 4**).

352 Additionally, we assessed whether variation in LTF expression could be driven by
353 genetic risk for AD. However, no significant associations were found between AD-
354 associated genetic variants and LTF expression levels in either CSF or plasma. These
355 findings suggest that, despite the known roles of LTF in immune response and
356 neuroinflammation, LTF expression levels in CSF and plasma do not appear to reflect AD
357 progression or genetic susceptibility. This may indicate that LTF biomarker potential in
358 CSF and plasma may be limited to identifying the presence of disease pathology, rather

359 than providing a biomarker read for disease risk phenoconversion or genetic risk
360 stratification in AD.

361 4. Discussion

362 Salivary LTF levels remain a controversial candidate biomarker for AD. To further
363 investigate the potential role of LTF as a biomarker for AD, we examined LTF expression
364 in both systemic and central compartments in the context of the AD diagnostic continuum
365 in the largest dataset reported to date (N= 1,367). Our rationale was that if salivary LTF
366 levels are directly associated with AD, similar associations might be detectable in other
367 biological fluids – particularly in CSF, which is in closer contact with the neuropathological
368 changes characteristic of the disease.

369 Our study was inspired by previous work suggesting a potential link between LTF and
370 AD. Specifically, a 2017 study conducted a cross-sectional analysis using saliva samples
371 and proposed salivary LTF as a promising biomarker to differentiate mild cognitive
372 impairment (MCI) and AD from healthy controls (27). Two subsequent studies by the
373 same group further supported this finding: one demonstrating that salivary LFT could
374 distinguish AD from other forms of dementia (28), and the other showed that LTF
375 abundance is closely related to AD pathology (29). Additional support came from a pilot
376 study that reported significant associations between salivary LTF levels and cognitive
377 performance, as measured by the Digit Span Memory Test and Mental Rotation Test
378 scores (43). Collectively, these studies suggest that reduced LTF levels in saliva may
379 serve as a non-invasive biomarker for AD diagnosis. However, a study from an
380 independent laboratory failed to replicate these findings in a consecutive, mixed memory
381 clinic cohort (30), casting doubt on the reliability of salivary LTF as an AD biomarker. Each
382 of these investigations were based on relatively small cohorts (N= 17-274)(27-30, 43),
383 which may reduce statistical power and contribute to inconsistent findings (44). Thus, it
384 remains to be determined whether changes in LTF levels are part of the causal chain
385 driving neurodegeneration, or merely epiphenomena reflecting systemic or local
386 inflammation. Additionally, the function of LTF is complex – it can exhibit both pro-
387 inflammatory and anti-inflammatory properties depending on the context of response (31).

388 Despite growing interest in LTF as a potential biomarker or mediator in AD, several
389 limitations complicate its study across different biological compartments. Beyond its role
390 in inflammation, LTF can act as a mediator of both innate and adaptive responses by
391 promoting the maturation, differentiation and activation of T- B-lymphocytes (45). Its
392 molecular complexity further challenges interpretation: LTF exists in multiple isoforms α,
393 β, γ, and δ-LTF. The canonical isoform is primarily expressed in neutrophils and epithelial
394 cell, secreted into exocrine inflamed tissue and fluid, while δ-LTF localizes to cytoplasm
395 and act as a transcription factor (46, 47). Additionally, Importantly, iron-binding status
396 further complicates LTF biology. LTF exists in different iron-binding states including the
397 iron-free (apo-LTF), native and iron-saturated forms (holo-LTF), which may have distinct
398 biological functions and activities (48, 49). Apo-LTF mainly sequesters iron, while holo-
399 LTF can activate immune signaling pathways via receptors such as LRP1 and CXCR4,
400 influencing MAPK, AKT, and NF-κB pathways (50, 51). Proteolytic cleavage of LTF can
401 also generate biologically active antimicrobial peptides known as lactoferricins, which
402 exhibit different functional properties compared to native LTF (52). Despite this molecular
403 complexity, most current studies-including those evaluating its potential as an AD
404 biomarker - use ELISA kits based on a single LTF antibody, which typically quantifies total
405 or a single protein isoform, without distinguishing among various functional forms.
406 Moreover, these studies often rely on correlation analyses between LTF levels and
407 disease biomarkers, without addressing the biological significance or mechanistic
408 relevance of inter-individual variation in LTF expression.

409 In our study, we aimed to explore the relationship between LTF levels in CSF and
410 plasma with AD status, disease progression and AT(N) biomarkers. Unlike previous work
411 focusing on salivary LTF, we used data derived from high-throughput SOMAscan
412 proteomics on larger, well-characterized cognitive and biomarker cohorts. Our findings
413 suggest that the relationship between LTF abundance and AD biomarkers is complex and
414 compartment specific.

415 In our hands, lactoferrin expression in CSF and Plasma shows a complex and
416 inconsistent association with AD biomarkers. Three key observations highlight the
417 complexity of interpreting LTF measurements using SOMAmer technology: 1. Lack of
418 consistency between SOMAmers: The two SOMAmers targeting to LTF (LTF1 and LTF2)

419 showed markedly different distributions and negligible correlation in both CSF and plasma
420 (**Fig 1D**). 2. Weak systemic-central correlation: Correlation between LTF SOMAmer
421 signals between CSF and plasma was weak (**Fig 1E**). 3. Inconsistent association with AD
422 biomarkers: In CSF, LTF1 displayed weak and even opposing correlations with AD
423 biomarkers p-tau depending on LTF-high vs. low subgroups (**Fig 1F**, **G**) which may be
424 caused by the variability in CSF protein abundance. Meanwhile, LTF1 is correlated with
425 Qalb, but not for LTF2, suggesting that LTF1 may be more sensitive to
426 barrier/inflammation-related signals, whereas LTF2 likely captures a different
427 conformational or compartment-specific facet of LTF. Notably, CSF LTF2 levels were
428 significantly higher in A+ than A-T-N- individuals (**Fig.S2**, $p = 0.0019$), and this difference
429 was replicated in comparisons between A+T+ and A-T- groups ($p = 0.0068$). This pattern
430 was not observed for LTF1. Together, these findings indicate that LTF measurements
431 obtained with SOMAmer reagents are heterogeneous across targets and fluids.

432 The most prominent studies evaluating LTF as an AD biomarker were conducted
433 using saliva samples. Carro, et al. (27) first reported positive associations between
434 salivary LTF and CSF Aβ42, and negative association with t-tau, suggesting that lower
435 levels of salivary LTF are associated with increased markers of AD pathology. Later
436 studies from the same group supported the discriminatory power of salivary LTF between
437 AD and other dementia types (28), and also linked it to cortical Aβ load and brain integrity
438 in aging adults (29). A separate pilot study (43) found lower salivary LTF in participants
439 with much more severe cognitive impairment. In contrast, our data from CSF and plasma
440 cohorts challenge these findings, but do not necessarily negate them, as we were unable
441 to replicate their work in saliva. We found there is higher CSF LTF2 levels in MCI dementia
442 compared with control (Kruskal-Wallis test; $p = 0.05$). These findings were consistent with
443 those of another independent study (30), which analyzed 222 matched saliva and CSF
444 samples from a consecutive, mixed memory clinic population and found both CSF and
445 saliva LTF expression have higher LTF expression in MCI and AD participants than
446 healthy control although the difference is not statically significant. This discrepancy may
447 arise from difference in sample origin, population heterogeneity, or the biological nature
448 of LTF secretion in different compartments. It is possible that LTF expression in CSF do
449 not reflect the association previously described for salivary LTF in AD. If the salivary LTF
450 finding is genuine, it is most likely capturing a biological signal that is not adequately

measured by SOMAscan assays in CSF or plasma. Moreover, the lack of saliva samples in our study limits direct comparability.

The molecular and multifunctional complexity of LTF likely contributes to these inconsistent results. LTF is involved in both innate and adaptive immune responses, influencing T- and B-cell activation (45). It exists in multiple isoforms (α, β, γ, δ), with distinct localization and functions. For example, δ-lactoferrin acts as a transcription factor and is localized in the cytoplasm (47). Meanwhile, iron-binding status of LTF also potentially impact on it biological functions and activities (48, 49). Another aspect is the technical considerations in SOMAmer-Based measurement. According to SomaLogic documentation, LTF1 targets amino acids 20 – 710, while LTF2 spans the full LTF sequence. These reagents may bind distinct epitopes, requiring different dilutions and potentially detecting different isoforms or modifications of LTF. Consistent with this notion, LTF1 and LTF2 show distinct associations with Qalb, suggesting that structural and conformational differences may underlie compartment-specific behavior and, consequently, their differential relationships with disease biomarkers. This distinction may partly explain the divergent results we observed.

Our study benefits from two notable key methodological strengths, multi-tissue protein profiling, enabling cross-compartment comparison; and large-scale proteomic-level analysis, providing consistently measured insights into disease relevant proteins.

However, several limitations must be acknowledged. First, our sample size, especially for healthy controls, may limit our statistical power. Second, though we had access to CSF and plasma proteomics, we lack this data in a cohort of saliva samples, which restricts our cross-compartment analysis and limits our interpretation of the data compared to previously published reports. We also acknowledge potential confounding introduced by unmeasured variables such as oral health, diet, or use of specific medications that may influence LTF levels or tissue secretion. Unfortunately, no isoform-level validation of SOMAmers exists to provide additional insight into the protein, protein-isoform, or component that is being measured, a critical consideration for proteins with known physical and functional diversity. Finally, we acknowledge that we have pulled our samples from a heterogeneous memory clinic population, which likely dilutes group-level effects that might otherwise be detectable.

Future work should include experimental validation of LTF isoforms and iron-binding forms across compartments, as well as mechanistic studies to clarify their roles in neuroinflammation and AD pathology.

Although lactoferrin shows some association with AD-related markers (particularly Aβ), its value as a biomarker remains uncertain due to its molecular complexity, compartment-specific behavior, and methodological limitations. Integrative multi-omics studies (36) emphasize the disconnection between protein abundance and genetic risk, reinforcing the need to explore regulatory and post-translational mechanisms. While exogenous LTF has shown therapeutic effects in preclinical models (53), such interventions may only benefit individuals with active LTF-related pathways. Identifying LTF-high subgroups in CSF or plasma could help stratify patients and tailor future treatments. However, it requires improved testing systems that could differentiate between LTF isoforms, their fragments, and their iron-binding states for disentangling the systemic and brain-specific roles of LTF in the pathophysiology of AD and related dementias. Ultimately, a deeper understanding of LTF's systemic and brain-specific roles is crucial for developing effective, personalized strategies for AD diagnosis and therapy.

Contributors

Conceptualization: Agustin Ruiz, Valentina R. Garbarino; Formal analysis: Feiyang Zhao; Funding acquisition: Agustin Ruiz, Sudha Seshadri, Merce Boada, Jose E. Cavazos, Sergi Valero, Amanda Cano and Maria Victoria Fernandez; Methodology: Feiyang Zhao, Agustin Ruiz; Project administration: Feiyang Zhao, Agustin Ruiz; Supervision: Merce Boada, Jose E. Cavazos, Agustin Ruiz, Sudha Seshadri; Validation: Feiyang Zhao; Visualization, Feiyang Zhao; Writing – original draft: Feiyang Zhao, Agustin Ruiz, Yaxi Wang; Writing – review & editing: Valentina R. Garbarino, Tiffany Kautz, Agustin Ruiz, Sudha Seshadri, Jose E Cavazos, Yaxi Wang, Eva Beckett, Feiyang Zhao; Proteome database generation and curation, APOE genotyping GWAS analysis and PRS construction: Raquel Puerta, Pablo Gacia Gonzalez, Maria Victoria Fernandez and Amanda Cano; Clinical and molecular data acquisition: Merce Boada, Amanda Cano, Adelina Orellana, Pilar Sanz and Maria Victoria Fernandez. All authors had access to the

data and accept the responsibility to submit for publication. All authors have read and approved the final version of the manuscript.

Declaration of Competing Interest

The authors declare no conflict of interest.

Funding

The study was partially funded by, the Agency for Innovation and Entrepreneurship (VLAIO) grant Nº PR067/21 (Belgium) and Janssen Pharmaceutica (Beerse, Belgium). V.R.G. was/is supported by NIAP25-1146397, 5T-P30-AG066546-03, P30-AG013319, TR004529, and a Clinical & Translational Research Pilot Grant Award from the Institute for Integration of Medicine & Science (IIMS) at UT Health San Antonio. Additionally, A.R. is also supported by the STAR Award. University of Texas System. Tx, United States, The South Texas ADRC. National Institute of Aging. National Institutes of Health. USA. (P30AG066546), the Keith M. Orme and Pat Vigeon Orme Endowed Chair in Alzheimer's and Neurodegenerative Diseases and The Patricia Ruth Frederick Distinguished Chair for Precision Therapeutics in Alzheimer's and Neurodegenerative Diseases. Finally, E.B. is supported by the STX-MSTP T32 training grants at UT San Antonio (NIH T32GM113896, T32GM145432).

The funders had no role in the conceptualization, study design, data collection, analysis, interpretation of data, in writing the paper or in the decision to submit the paper for publication.

Acknowledgments

We would like to thank the patients and controls who participated in this project. We are also grateful to the ACE Alzheimer Center Barcelona Memory Clinic, the clinical trial

and nurse station staff involved in participant evaluations and case management, for their passionate dedication to the patients and families contributing to this research.

References

1. Selkoe DJ. Cell biology of protein misfolding: the examples of Alzheimer's and Parkinson's diseases. Nat Cell Biol. 2004;6(11):1054-61.

2. Simonsen AH, Herukka SK, Andreasen N, Baldeiras I, Bjerke M, Blennow K, et al. Recommendations for CSF AD biomarkers in the diagnostic evaluation of dementia. Alzheimers Dement. 2017;13(3):274-84.

3. Grande G, Valletta M, Rizzuto D, Xia X, Qiu C, Orsini N, et al. Blood-based biomarkers of Alzheimer's disease and incident dementia in the community. Nat Med. 2025;31(6):2027-35.

4. Drachman DA. The amyloid hypothesis, time to move on: Amyloid is the downstream result, not cause, of Alzheimer's disease. Alzheimers Dement. 2014;10(3):372-80.

5. Itzhaki RF, Lathe R, Balin BJ, Ball MJ, Bearer EL, Braak H, et al. Microbes and Alzheimer's Disease. J Alzheimers Dis. 2016;51(4):979-84.

6. Stanley LC, Mrak RE, Woody RC, Perrot LJ, Zhang S, Marshak DR, et al. Glial cytokines as neuropathogenic factors in HIV infection: pathogenic similarities to Alzheimer's disease. J Neuropathol Exp Neurol. 1994;53(3):231-8.

7. Green DA, Masliah E, Vinters HV, Beizai P, Moore DJ, Achim CL. Brain deposition of beta-amyloid is a common pathologic feature in HIV positive patients. AIDS. 2005;19(4):407-11.

8. Smith DB, Simmonds P, Bell JE. Brain viral burden, neuroinflammation and neurodegeneration in HAART-treated HIV positive injecting drug users. J Neurovirol. 2014;20(1):28-38.

9. McQuaid S, Allen IV, McMahon J, Kirk J. Association of measles virus with neurofibrillary tangles in subacute sclerosing panencephalitis: a combined in situ hybridization and immunocytochemical investigation. Neuropathol Appl Neurobiol. 1994;20(2):103-10.

10. Eimer WA, Vijaya Kumar DK, Navalpur Shanmugam NK, Rodriguez AS, Mitchell T, Washicosky KJ, et al. Alzheimer's Disease-Associated beta-Amyloid Is Rapidly Seeded by Herpesviridae to Protect against Brain Infection. Neuron. 2018;99(1):56-63 e3.

11. Miklossy J. Alzheimer's disease - a neurospirochetosis. Analysis of the evidence following Koch's and Hill's criteria. J Neuroinflammation. 2011;8:90.

12. Hill JM, Clement C, Pogue AI, Bhattacharjee S, Zhao Y, Lukiw WJ. Pathogenic microbes, the microbiome, and Alzheimer's disease (AD). Front Aging Neurosci. 2014;6:127.

13. Welling MM, Nabuurs RJ, van der Weerd L. Potential role of antimicrobial peptides in the early onset of Alzheimer's disease. Alzheimers Dement. 2015;11(1):51-7.

14. Haage V, De Jager PL. Neuroimmune contributions to Alzheimer's disease: a focus on human data. Mol Psychiatry. 2022;27(8):3164-81.

15. Chen L, Shen Q, Liu Y, Zhang Y, Sun L, Ma X, et al. Homeostasis and metabolism of iron and other metal ions in neurodegenerative diseases. Signal Transduct Target Ther. 2025;10(1):31.

573 16. Soscia SJ, Kirby JE, Washicosky KJ, Tucker SM, Ingelsson M, Hyman B, et al. The
574 Alzheimer's disease-associated amyloid beta-protein is an antimicrobial peptide. PLoS One.
575 2010;5(3):e9505.

576 17. White MR, Kandel R, Tripathi S, Condon D, Qi L, Taubenberger J, et al. Alzheimer's
577 associated beta-amyloid protein inhibits influenza A virus and modulates viral interactions with
578 phagocytes. PLoS One. 2014;9(7):e101364.

579 18. Bourgade K, Garneau H, Giroux G, Le Page AY, Bocti C, Dupuis G, et al. beta-Amyloid
580 peptides display protective activity against the human Alzheimer's disease-associated herpes
581 simplex virus-1. Biogerontology. 2015;16(1):85-98.

582 19. Bourgade K, Le Page A, Bocti C, Witkowski JM, Dupuis G, Frost EH, et al. Protective
583 Effect of Amyloid-beta Peptides Against Herpes Simplex Virus-1 Infection in a Neuronal Cell
584 Culture Model. J Alzheimers Dis. 2016;50(4):1227-41.

585 20. Wang B, Timilsena YP, Blanch E, Adhikari B. Lactoferrin: Structure, function, denaturation
586 and digestion. Crit Rev Food Sci Nutr. 2019;59(4):580-96.

587 21. Yong SJ, Veerakumarasivam A, Lim WL, Chew J. Neuroprotective Effects of Lactoferrin in
588 Alzheimer's and Parkinson's Diseases: A Narrative Review. ACS Chem Neurosci. 2023.

589 22. Qian ZM, Li W, Guo Q. Lactoferrin/lactoferrin receptor: Neurodegenerative or
590 neuroprotective in Parkinson's disease? Ageing Res Rev. 2024;101:102474.

591 23. Tsatsanis A, McCorkindale AN, Wong BX, Patrick E, Ryan TM, Evans RW, et al. The acute
592 phase protein lactoferrin is a key feature of Alzheimer's disease and predictor of Abeta burden
593 through induction of APP amyloidogenic processing. Mol Psychiatry. 2021;26(10):5516-31.

594 24. Bennett RM, Kokocinski T. Lactoferrin content of peripheral blood cells. Br J Haematol.
595 1978;39(4):509-21.

596 25. Fillebeen C, Ruchoux MM, Mitchell V, Vincent S, Benaissa M, Pierce A. Lactoferrin is
597 synthesized by activated microglia in the human substantia nigra and its synthesis by the human
598 microglial CHME cell line is upregulated by tumor necrosis factor alpha or 1-methyl-4-
599 phenylpyridinium treatment. Brain Res Mol Brain Res. 2001;96(1-2):103-13.

600 26. Wang W, An Q, Zou Y, Dai Y, Meng Q, Zhang Y. Lactoferrin alleviates the adverse effects
601 of early-life inflammation on depression in adults by regulating the activation of microglia. Mol
602 Med. 2025;31(1):50.

603 27. Carro E, Bartolome F, Bermejo-Pareja F, Villarejo-Galende A, Molina JA, Ortiz P, et al.
604 Early diagnosis of mild cognitive impairment and Alzheimer's disease based on salivary lactoferrin.
605 Alzheimers Dement (Amst). 2017;8:131-8.

606 28. Gonzalez-Sanchez M, Bartolome F, Antequera D, Puertas-Martin V, Gonzalez P, Gomez-
607 Grande A, et al. Decreased salivary lactoferrin levels are specific to Alzheimer's disease.
608 EBioMedicine. 2020;57:102834.

609 29. Reseco L, Atienza M, Fernandez-Alvarez M, Carro E, Cantero JL. Salivary lactoferrin is
610 associated with cortical amyloid-beta load, cortical integrity, and memory in aging. Alzheimers
611 Res Ther. 2021;13(1):150.

612 30. Gleerup HS, Jensen CS, Hogh P, Hasselbalch SG, Simonsen AH. Lactoferrin in
613 cerebrospinal fluid and saliva is not a diagnostic biomarker for Alzheimer's disease in a mixed
614 memory clinic population. EBioMedicine. 2021;67:103361.

615 31. Legrand D. Lactoferrin, a key molecule in immune and inflammatory processes. Biochem
616 Cell Biol. 2012;90(3):252-68.

32. Espinosa A, Alegret M, Valero S, Vinyes-Junque G, Hernandez I, Mauleon A, et al. A longitudinal follow-up of 550 mild cognitive impairment patients: evidence for large conversion to dementia rates and detection of major risk factors involved. J Alzheimers Dis. 2013;34(3):769-80.

33. Hansson O, Batrla R, Brix B, Carrillo MC, Corradini V, Edelmayer RM, et al. The Alzheimer's Association international guidelines for handling of cerebrospinal fluid for routine clinical measurements of amyloid beta and tau. Alzheimers Dement. 2021;17(9):1575-82.

34. Jack CR, Jr., Bennett DA, Blennow K, Carrillo MC, Dunn B, Haeberlein SB, et al. NIA-AA Research Framework: Toward a biological definition of Alzheimer's disease. Alzheimers Dement. 2018;14(4):535-62.

35. Morato X, Puerta R, Cano A, Orellana A, de Rojas I, Capdevila M, et al. Associations of plasma SMOC1 and soluble IL6RA levels with the progression from mild cognitive impairment to dementia. Brain Behav Immun Health. 2024;42:100899.

36. Puerta R, de Rojas I, Garcia-Gonzalez P, Olive C, Sotolongo-Grau O, Garcia-Sanchez A, et al. Linking genomic and proteomic signatures to brain amyloid burden: insights from GR@ACE/DEGESCO. Funct Integr Genomics. 2025;25(1):73.

37. Puerta R, Cano A, Garcia-Gonzalez P, Garcia-Gutierrez F, Capdevila M, de Rojas I, et al. Head-to-Head Comparison of Aptamer- and Antibody-Based Proteomic Platforms in Human Cerebrospinal Fluid Samples from a Real-World Memory Clinic Cohort. Int J Mol Sci. 2024;26(1).

38. Imam F, Saloner R, Vogel JW, Krish V, Abdel-Azim G, Ali M, et al. The Global Neurodegeneration Proteomics Consortium: biomarker and drug target discovery for common neurodegenerative diseases and aging. Nat Med. 2025.

39. Candia J, Cheung F, Kotliarov Y, Fantoni G, Sellers B, Griesman T, et al. Assessment of Variability in the SOMAscan Assay. Sci Rep. 2017;7(1):14248.

40. Bellenguez C, Kucukali F, Jansen IE, Kleineidam L, Moreno-Grau S, Amin N, et al. New insights into the genetic etiology of Alzheimer's disease and related dementias. Nat Genet. 2022;54(4):412-36.

41. Uhlen M, Oksvold P, Fagerberg L, Lundberg E, Jonasson K, Forsberg M, et al. Towards a knowledge-based Human Protein Atlas. Nat Biotechnol. 2010;28(12):1248-50.

42. García-González P, Puerta R, de Rojas I, Marquié M, Emon A, Smets B, et al. Integrative multi-omics analysis reveals physiological and genetic drivers of CSF biomarker variability: implications in neurodegeneration studies. Alzheimer's Association International Conference; 2025 Jul 30; Toronto, Canada2025.

43. Hammerschlag BL, Butts B, Likos K, Verble DD, Nimmagadda N, Virani R, et al. Pilot: Salivary Lactoferrin as a Biomarker of Alzheimer's Disease. medRxiv. 2024.

44. Button KS, Ioannidis JP, Mokrysz C, Nosek BA, Flint J, Robinson ES, et al. Power failure: why small sample size undermines the reliability of neuroscience. Nat Rev Neurosci. 2013;14(5):365-76.

45. Legrand D. Overview of Lactoferrin as a Natural Immune Modulator. J Pediatr. 2016;173 Suppl:S10-5.

46. Siebert PD, Huang BC. Identification of an alternative form of human lactoferrin mRNA that is expressed differentially in normal tissues and tumor-derived cell lines. Proc Natl Acad Sci U S A. 1997;94(6):2198-203.

47. Mariller C, Hardiville S, Hoedt E, Huvent I, Pina-Canseco S, Pierce A. Delta-lactoferrin, an intracellular lactoferrin isoform that acts as a transcription factor. Biochem Cell Biol. 2012;90(3):307-19.

662 48. Jiang R, Lopez V, Kelleher SL, Lonnerdal B. Apo- and holo-lactoferrin are both internalized
663 by lactoferrin receptor via clathrin-mediated endocytosis but differentially affect ERK-signaling
664 and cell proliferation in Caco-2 cells. J Cell Physiol. 2011;226(11):3022-31.

665 49. Lopez V, Kelleher SL, Lonnerdal B. Lactoferrin receptor mediates apo- but not holo-
666 lactoferrin internalization via clathrin-mediated endocytosis in trophoblasts. Biochem J.
667 2008;411(2):271-8.

668 50. Tian M, Han YB, Yang GY, Li JL, Shi CS, Tian D. The role of lactoferrin in bone remodeling:
669 evaluation of its potential in targeted delivery and treatment of metabolic bone diseases and
670 orthopedic conditions. Front Endocrinol (Lausanne). 2023;14:1218148.

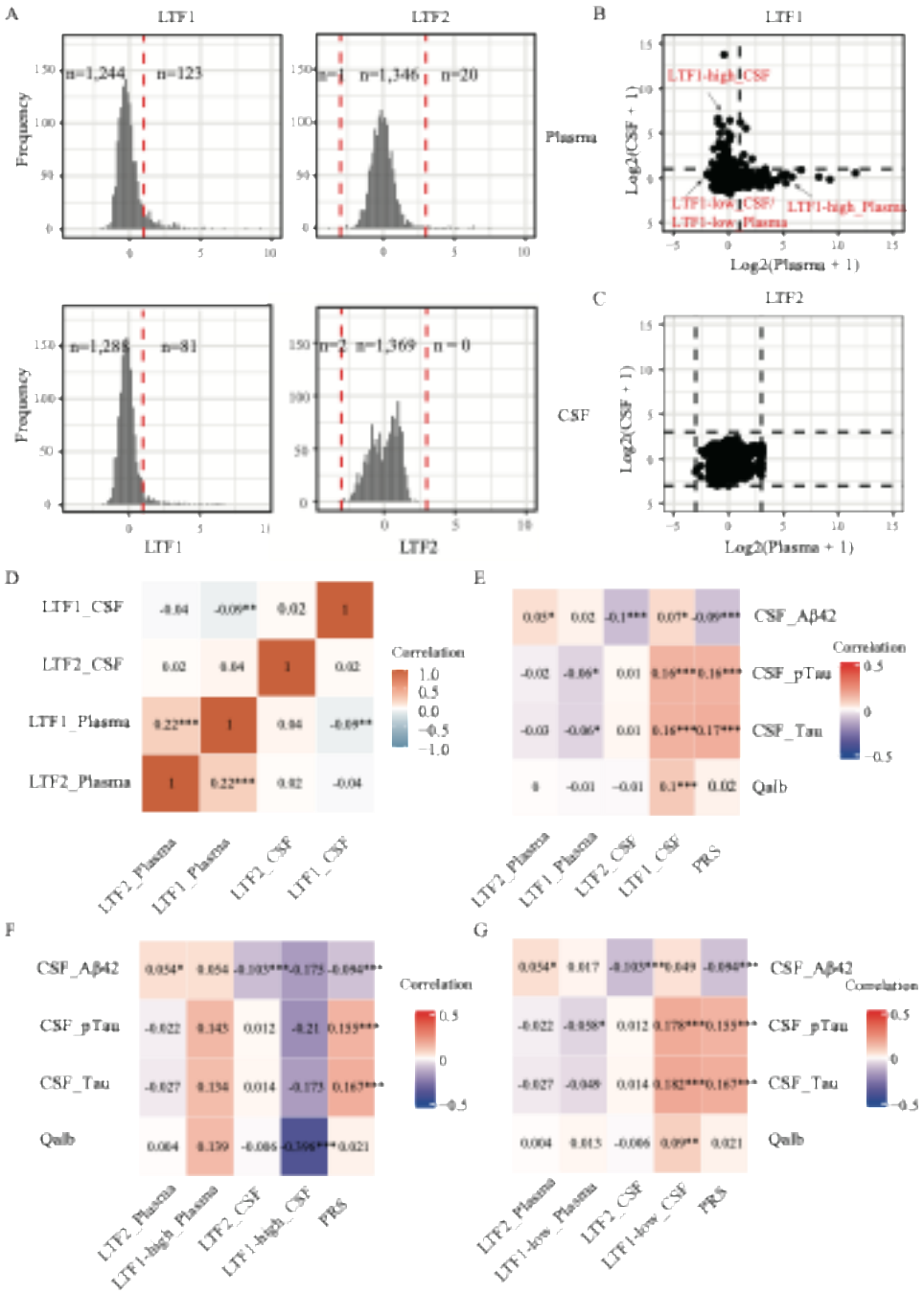
671 51. Takayama Y, Aoki R, Uchida R, Tajima A, Aoki-Yoshida A. Role of CXC chemokine receptor
672 type 4 as a lactoferrin receptor. Biochem Cell Biol. 2017;95(1):57-63.

673 52. Gifford JL, Hunter HN, Vogel HJ. Lactoferricin: a lactoferrin-derived peptide with
674 antimicrobial, antiviral, antitumor and immunological properties. Cell Mol Life Sci.
675 2005;62(22):2588-98.

676 53. Zheng J, Xie Y, Li F, Zhou Y, Qi L, Liu L, et al. Lactoferrin improves cognitive function and
677 attenuates brain senescence in aged mice. Journal of Functional Foods. 2020;65:103736.

678

679



680

681

682

683

684

Fig. 1. Distributions of SOMAmer signals targeting lactoferrin (LTF) and their associations with Alzheimer's disease (AD) biomarkers.

A: Distributions of two SOMAmers targeting LTF (LTF1 on the left, LTF2 on the right) in plasma (top) and CSF (bottom). Red dotted lines indicate the cutoffs defined as mean + standard deviations for LTF1 and mean ± 3 standard deviations for LTF2. All p-value were calculated by Shapiro.test for normality test.

B–C: Comparisons of LTF1 and LTF2 signals between CSF and plasma; dotted lines represent the cutoffs shown in panel A.

D: Heatmap showing the Spearman correlation between LTF1 and LTF2.

E–G: Heatmaps showing Spearman correlations between the two SOMAmers, polygenic risk scores (PRS), and AD biomarkers. Panel E shows results for all samples; panels F and G show correlations after stratifying by LTF1 level into high and low for LTF1, respectively. *, $p < 0.05$; **, $p < 0.01$; ***, $p < 0.001$.

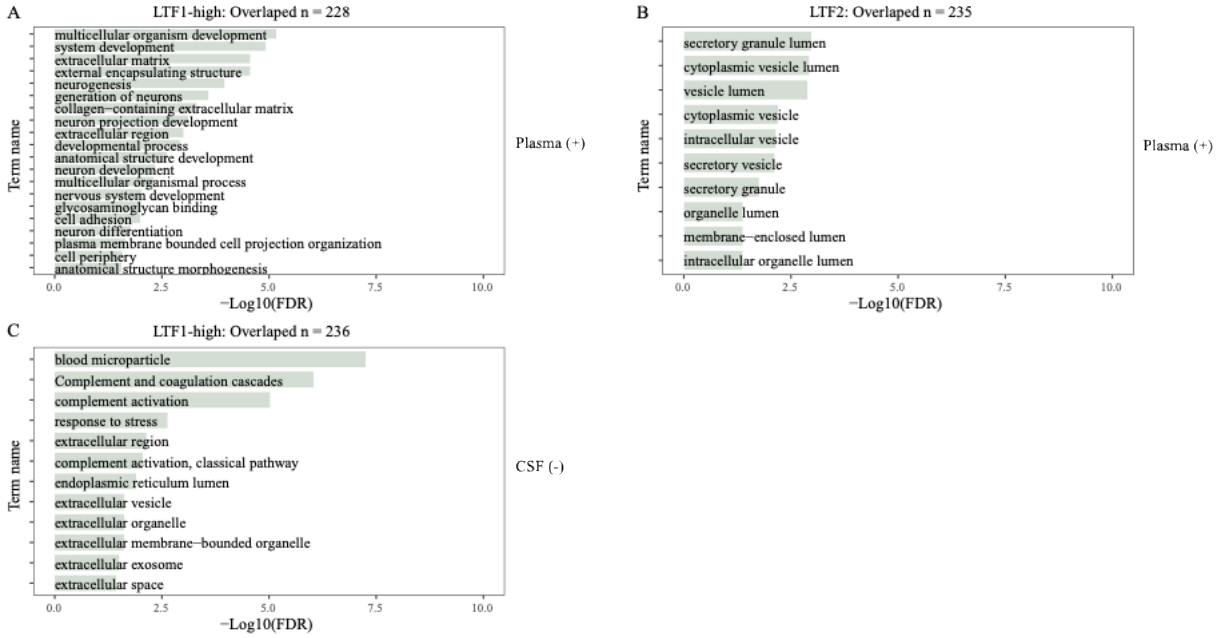
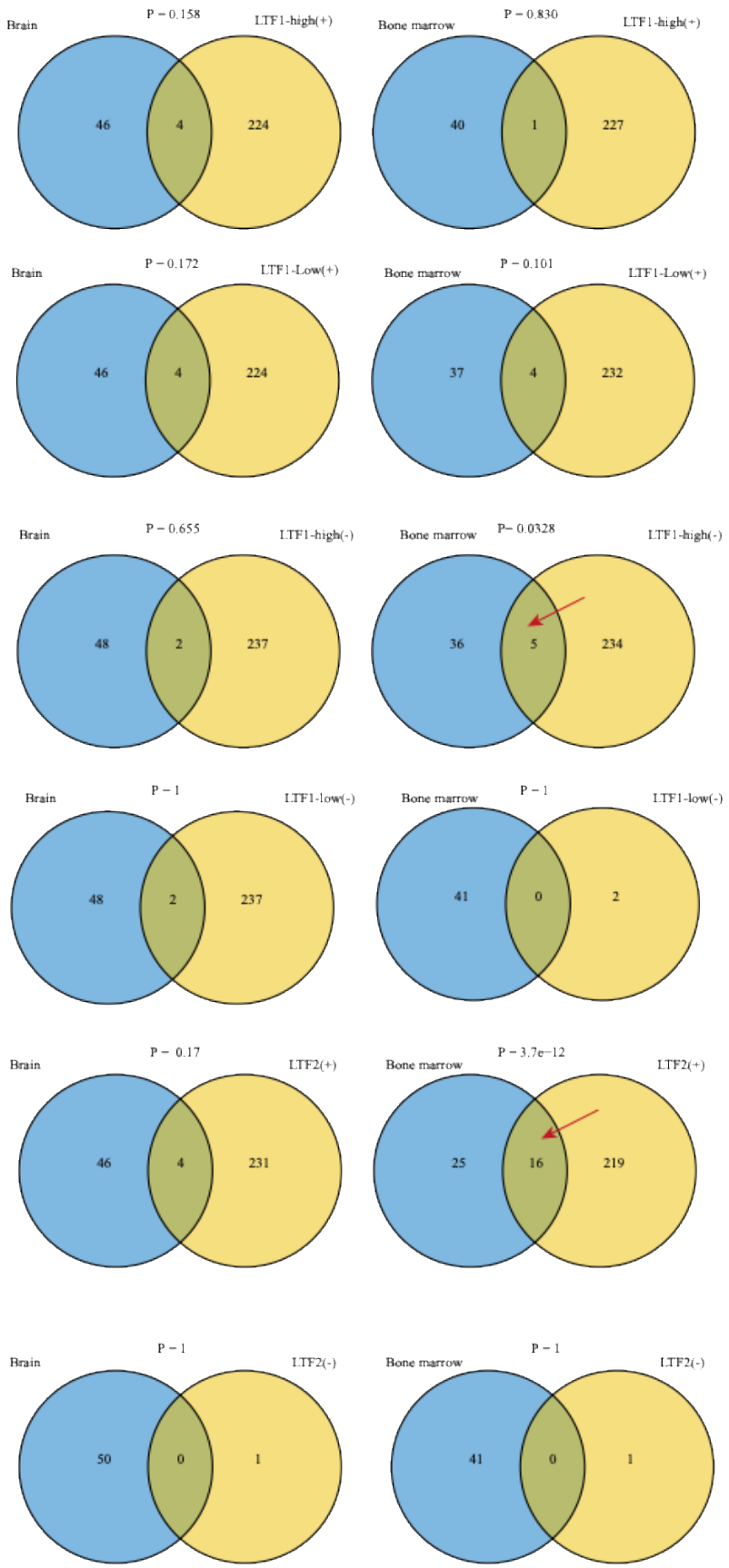


Fig. 2. GO enrichment analysis of lactoferrin-associated proteins identified from SOMAscan plasma and CSF proteomic data.

A: Significantly enriched Gene Ontology (GO) biological process terms among proteins positively correlated with the plasma-derived LTF1 signal in individuals classified as + (n = 226).

B: GO cellular component terms enriched among proteins positively correlated with the plasma-derived LTF2 signal (n = 226).

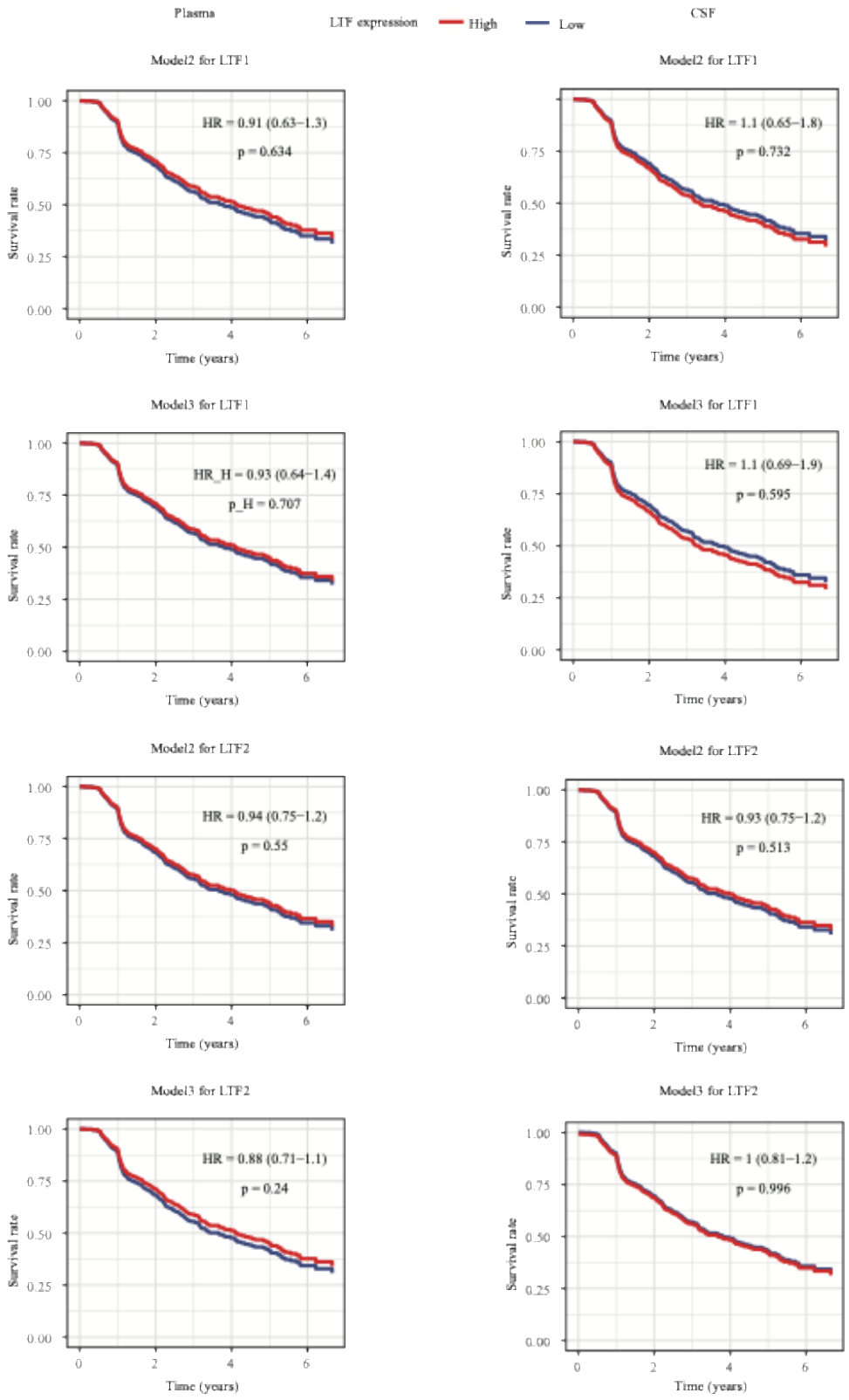
C: GO biological process and cellular component terms enriched among proteins negatively correlated with the CSF-derived LTF1 signal in the LTF1-positive group (n = 236). Enrichment significance is represented by -log10(FDR)-adjusted p-values.


Brain
P = 0.158
LTF1-high(+)
46
4
224
Bone marrow
P = 0.830
LTF1-high(+)
40
1
227
Brain
P = 0.172
LTF1-Low(+)
46
4
224
Bone marrow
P = 0.101
LTF1-Low(+)
37
4
232
Brain
P = 0.655
LTF1-high(-)
48
2
237
Bone marrow
P= 0.0328
LTF1-high(-)
36
5
234
Brain
P = 1
LTF1-low(-)
48
2
237
Bone marrow
P = 1
LTF1-low(-)
41
0
2
Brain
P = 0.17
LTF2(+)
46
4
231
Bone marrow
P = 3.7e−12
LTF2(+)
25
16
219
Brain
P = 1
LTF2(-)
50
0
1
Bone marrow
P = 1
LTF2(-)
41
0
1

709

710 **Fig. 3. Overlap between LTF-derived correlation signatures in plasma and tissue-**
711 **specific co-expression modules.**
712 Columns show tissue modules from the Human Protein Atlas database (Brain, #237 and
713 Bone marrow, #154 blue circles). Rows stratify our signatures by (i) LTF1 expression
714 stratum—LTF1-high vs LTF1-low and (ii) correlation sign with the LTF signal: “+” = top
715 positively correlated genes, “−” = top negatively correlated genes. For LTF2,
716 correlations were computed across all samples (no expression stratification). In each
717 panel, the right circle is the LTF-derived signature (yellow) and the left circle is the tissue
718 co-expression set (blue). Statistical significance was assessed with Fisher’s exact test
719 (background = all tested genes after QC, n = 5,419).

720


Plasma
LTF expression
High
Low
CSF
Model2 for LTF1
HR = 0.91 (0.63−1.3)
p = 0.634
Model2 for LTF1
HR = 1.1 (0.65−1.8)
p = 0.732
Model3 for LTF1
HR_H = 0.93 (0.64−1.4)
p_H = 0.707
Model3 for LTF1
HR = 1.1 (0.69−1.9)
p = 0.595
Model2 for LTF2
HR = 0.94 (0.75−1.2)
p = 0.55
Model2 for LTF2
HR = 0.93 (0.75−1.2)
p = 0.513
Model3 for LTF2
HR = 0.88 (0.71−1.1)
p = 0.24
Model3 for LTF2
HR = 1 (0.81−1.2)
p = 0.996
Survival rate
Time (years)

722 **Fig. 4. Survival analysis for lactoferrin (LTF) using nested models.**
723 The left column shows results based on plasma data; the right column shows results
724 based on CSF data. For LTF1, survival rates were compared between LTF1 high and
725 LTF1 low. For LTF2, samples were split by the median (High: ≥ median; Low: < median).
726 Model1: Surv(Time(years), status) ~ LTF + Age_LP + sex + education + mmse_csf;
727 Model2: Surv(Time(years), status) ~ LTF + csf_ p_tau + PRS + APOE + Age_LP + sex +
728 education + mmse_csf. HR represents the hazard ratio of the positive group, 95%
729 confidence intervals is included, and p is the corresponding p-value.

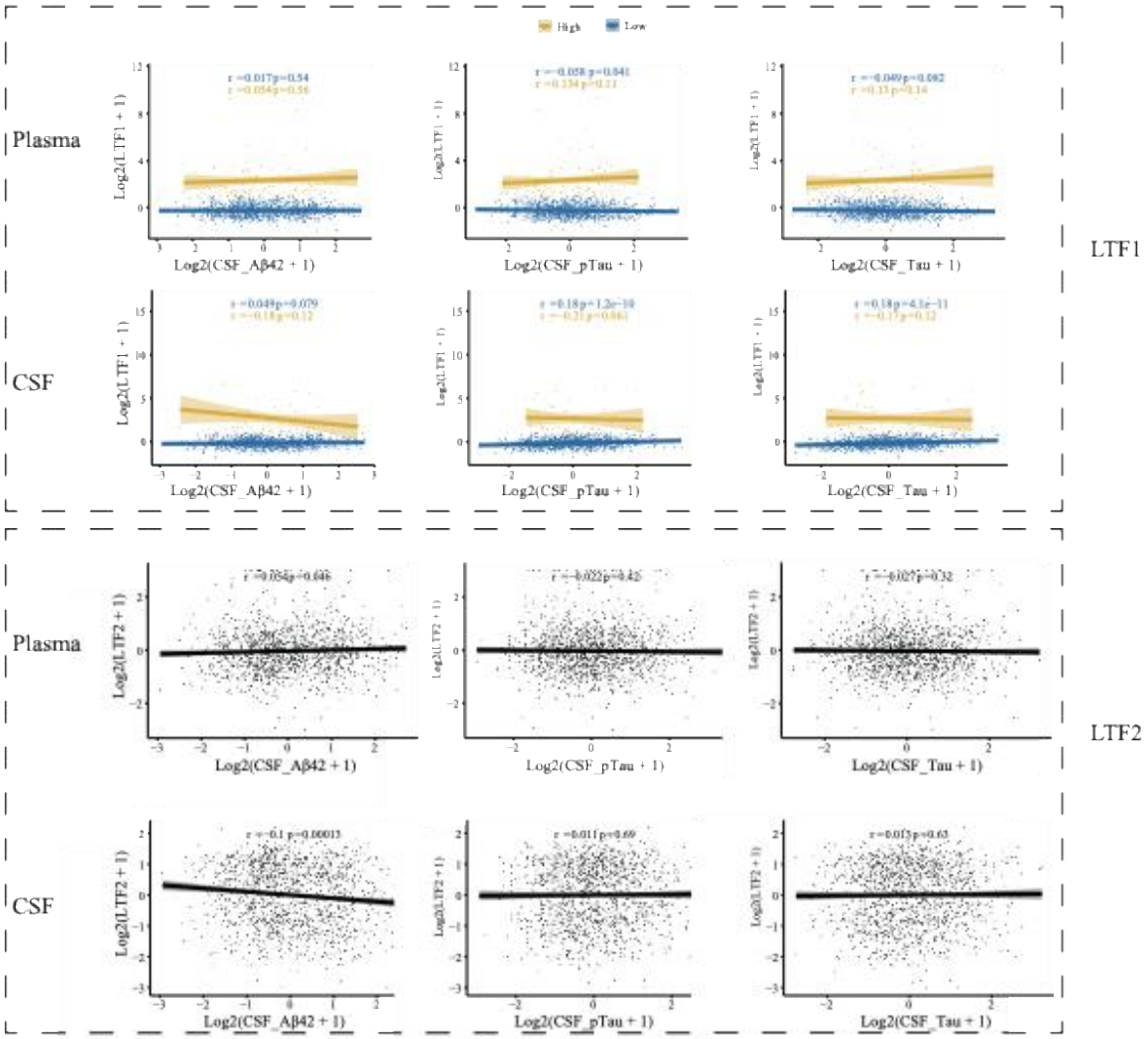


Fig. S1. Scatter plots showing the correlations between LTF1/LTF2 and AD biomarkers (Aβ42, p-tau, and t-tau).

The top row shows correlations with LTF1, and the bottom row with LTF2. LTF1 is stratified by LTF1 expression—LTF1-high (yellow) vs LTF1-low (blue). For LTF2, correlations were computed across all samples. Spearman correlation coefficients (r) and corresponding *p* values (*p*) are indicated in each panel.

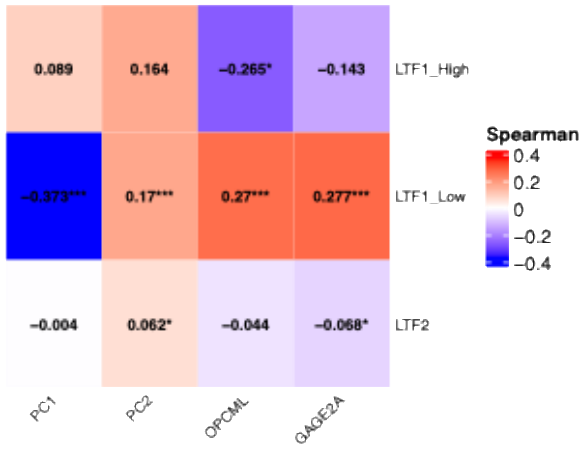
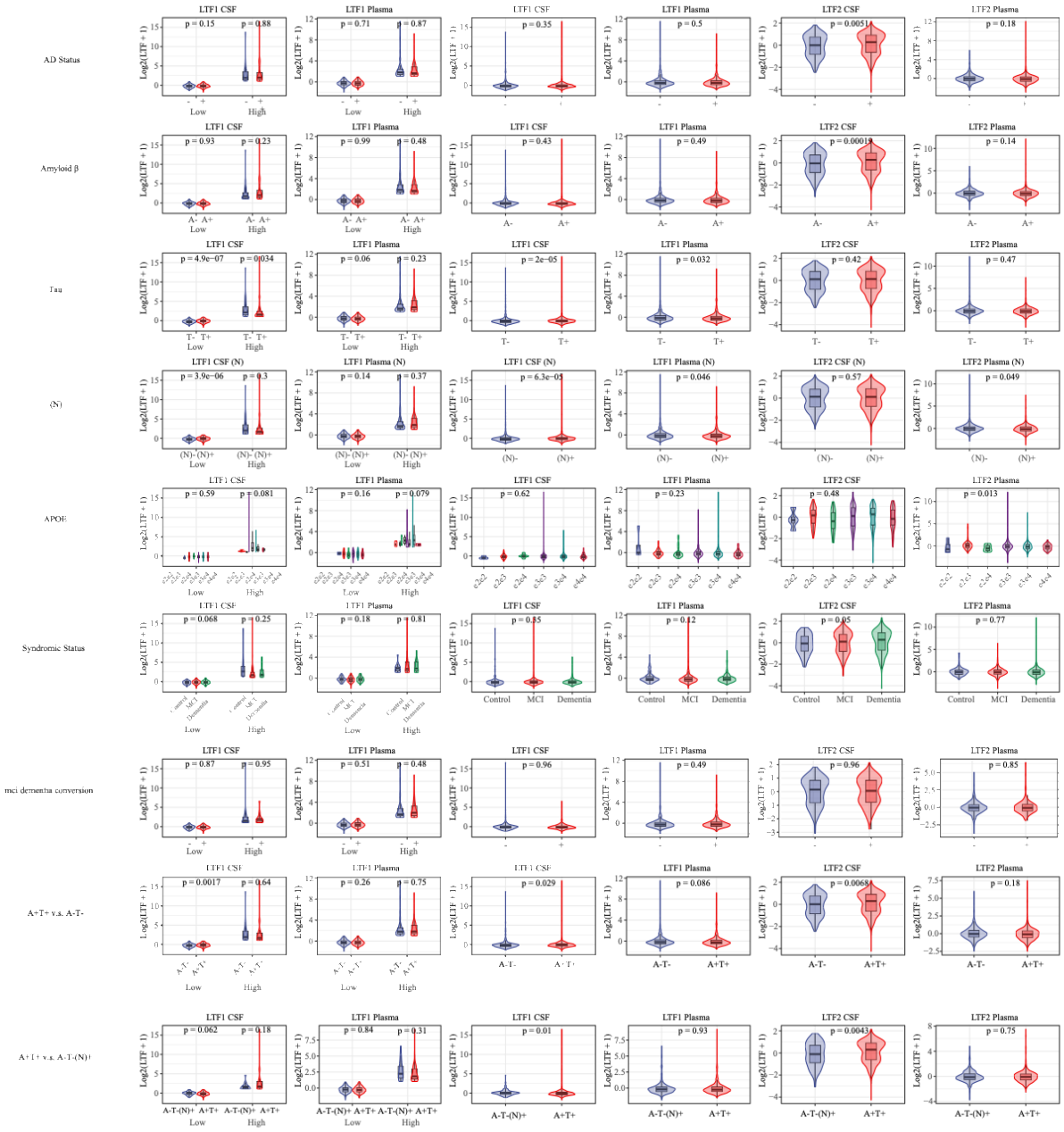


Fig. S2. **Spearman correlation heatmap between CSF LTF1-/LTF2 and the first two proteomic principal components (PC1–PC2) and reference genes (OPCML, GAGE2A).** Cells display Spearman's ρ; color encodes effect size. Significance: *, $p < 0.05$; **, $p < 0.01$; **, $p < 0.001$.


AD Status
Amyloid β
Tau
(N)
APOE
Syndromic Status
mci dementia conversion
A+T+ v.s. A-T-
A+T+ v.s. A-T-(N)+
LTF1 CSF
LTF1 Plasma
LTF2 CSF
LTF2 Plasma
LTF1 CSF (N)
LTF1 Plasma (N)
LTF2 CSF (N)
LTF2 Plasma (N)
Log2(LTF + 1)
Low
High
Control
MCI
Dementia

753 **Fig. S3. Comparison of LTF1 and LTF2 levels across AD biomarker profiles,**
754 **clinical diagnosis, APOE genotypes, and longitudinal conversion status.**
755 Each row represents a comparison between groups stratified by APOE genotype,
756 dementia diagnosis, A/T/N biomarker positivity, or longitudinal conversion status.
757 Columns 1 – 2 show CSF and plasma LTF1 levels stratified by LTF1 abundance (high,
758 low) respectively; columns 3 – 4 show the CSF and plasma LTF1 levels before
759 stratification. Columns 5 – 6 show CSF and plasma LTF2 levels. In each plot, different
760 color were used to discriminate the status of diagnosis, AD biomarkers, APOE
761 genotypes or conversion status respectively. Statistical comparisons were performed
762 using the Kruskal-Wallis test.

Supplement Table

Table S1. Annotation of the two SOMAmers that target to LTF**.**

Table S2. Demographic and clinical data of the participants in the ACE cohort.

Abbreviations: LP: lumbar puncture; MCI, mild cognitive impairment; AD, Alzheimer's disease dementia; MMSE, Mini Mental State Examination; LTF, Lactoferrin; NA, not applicable; SD, standard deviation.

Table S3. This table show the results of the linear regression between csf Aβ42, p-tau and LTF1 in CSF using sex, age, age^2, reference gene, PC1, PC2. Results are shown for models adjusting for each covariate individually and for a fully adjusted model including all covariates simultaneously.

Table S4. This table show the results of the linear regression between csf Aβ42, p-tau and LTF1 in CSF using sex, age, age^2, reference gene, PC1, PC2 as covariates in LTF1 high express population. Results are shown for models adjusting for each covariate individually and for a fully adjusted model including all covariates simultaneously.

Table S5. This table show the results of the linear regression between csf Aβ42, p-tau and LTF1 in CSF using sex, age, age^2, reference gene, PC1, PC2 as covariates in LTF1 low express population. Results are shown for models adjusting for each covariate individually and for a fully adjusted model including all covariates simultaneously.

Table S6. This table show the results of the linear regression between csfA β42, p-tau181 and LTF2 in CSF using sex, age, age^2, reference gene, PC1, PC2 as covariates. Results are shown for models adjusting for each covariate individually and for a fully adjusted model including all covariates simultaneously.