**GSN Gene Frameshift Mutations in Alzheimer’s Disease**

Yaling Jiang, Meidan Wan, Xuewen Xiao, Zhuojie Lin, Xixi Liu, Yafang Zhou, Xinxin Liao, Jingyi Lin, Hui Zhou, Lu Zhou, Ling Weng, Junling Wang, Jifeng Guo, Hong Jiang, Zhuohua Zhang, Kun Xia, Jiada Li, Beisha Tang, Bin Jiao*, Lu Shen*

Include:

- Supplementary methods
- Supplementary Figure S1-6
- Supplementary Table S1

**Supplementary methods**

**ATN framework diagnosis classification**

The β amyloid deposition, pathologic tau, and neurodegeneration (ATN) framework[1] was described below: Aβ42 level in CSF < 651 pg/ml, or Aβ42/Aβ40 ratio ≤ 0.1, is defined as positive amyloidosis (corresponding to A+ in the ATN framework); p-tau > 61 pg/ml in CSF is defined as neurofibrillary tangles (corresponding to T+ in the ATN framework); t-tau ≥ 290 pg/ml in CSF is defined as nerve cell death (corresponding to N+ in the ATN framework).

**Cognitive Impairment Multicenter Database and Collaborative Network in China (CI-MDCNC)**

The CI-MDCNC cohort recruited 884 AD patients from 8 clinical sites across China, which is one of the largest multicenter, observational, longitudinal, and natural history studies of AD in China.

**Alzheimer’s Disease Sequencing Project (ADSP) public database**

The ADSP performed whole exome sequencing (WES) of DNA specimens obtained from 5,778 AD cases and 5,136 controls, including 5,519 AD cases and 4,917
cognitively normal elderly controls of European ancestry (EA) and 218 AD cases and 177 controls of Caribbean Hispanic (CH) heritage[2-4].

**Targeted gene sequencing**

The targeted gene sequencing panel includes four genes (GSN, APP, PSEN1, and PSEN2). The specific methods were described previously[5]. In brief, the genomic DNA was broken into 200-250 bp length fragments by Covaris LE220 (MA, USA). End-repairing, A-tailing, adaptor ligation, and a 15-cycle pre-capture PCR amplification were performed in fragmented DNA. The products were captured by the targeted gene sequencing panel and sequenced on Illumina HiSeq X Ten Analyzers (Illumina, San Diego, CA, USA) after quality control. Then, the bioinformatics processing and data analysis were performed.

**3D model structure**

The three-dimensional (3D) models of the mutant protein structures were built by Discovery Studio software. We used homology models of gelsolin in the Protein Data Bank (PDB; No. 3ffn) to construct the 3D structure of the mutant proteins by Discovery Studio software. The known 3D model of the GSN D214N mutant protein was from the PDB (No. 6qbf).

**Global cerebral atrophy-frontal sub-scale (GCA-F)**

The GCA-F scale is defined in the frontal cortex and sulcal. It ranks 0 to 3, with grade 0 representing no cortical atrophy; grade 1, mild atrophy (dilatation of sulci); grade 2, moderate atrophy (loss of gyri volume); and grade 3, end-stage “knife blade atrophy”.[6, 7]

**Posterior atrophy (PA)**

The PA scale is defined in the posterior cingulate sulcus, parieto-occipital sulcus, precuneus, and parietal cortex. It ranks atrophy from 0 to 3, with grade 0 showing no cortical atrophy; grade 1, mild parietal cortical atrophy, with mild enlargement of the posterior cingulate and parieto-occipital sulcus; grade 2, substantial parietal atrophy, with substantial enlargement of the posterior cingulate and parieto-occipital sulcus; grade 3, end-stage “knife-blade” atrophy, with extreme enlargement of the posterior cingulate and parieto-occipital sulcus.[8]
Medial temporal lobe atrophy (MTA)

The MTA scale is defined in the hippocampus, parahippocampal gyrus, and surrounding CSF spaces. It ranks the degree of MTA from 0 to 4, with grade 0 denoting no cortical atrophy; grade 1, enlargement of choroid fissure; grade 2, along with enlargement of the temporal horn of the lateral ventricle; grade 3, moderate loss of hippocampal volume; and grade 4, severe loss of hippocampal volume.[9]

Reference:


Supplementary Figure S1-6

A

<table>
<thead>
<tr>
<th>Gene</th>
<th>Primer sequence</th>
<th>Annealing temperature</th>
<th>Size</th>
</tr>
</thead>
<tbody>
<tr>
<td>GSN-K346fs</td>
<td>Forward 5'-AGCTGGCAAGCTCTACAG-3'</td>
<td>60 °C</td>
<td>181 bp</td>
</tr>
<tr>
<td></td>
<td>Reverse 5'-TTCCCTCTCCCGTCTGTTG-3'</td>
<td></td>
<td></td>
</tr>
<tr>
<td>GSN-P3fs</td>
<td>Forward 5'-ACTTAAGGTCCGGGACC-3'</td>
<td>56 °C</td>
<td>454 bp (Mut)</td>
</tr>
<tr>
<td></td>
<td>Reverse 5'-CCTCAAGTTACATCCGG-3'</td>
<td></td>
<td></td>
</tr>
<tr>
<td>GAPDH</td>
<td>Forward 5'-AGTAAAGGCGCCCTGGA-3'</td>
<td>60 °C</td>
<td>131 bp</td>
</tr>
<tr>
<td></td>
<td>Reverse 5'-TGGACAGACGCCCATCT-3'</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

B

GSN-K346fs cDNA sequencing

Figure S1 PCR primers and cDNA sequences. (A) The PCR primers for GSN frameshift mutations. (B) The cDNA sequencing of Case 1 with the K346fs mutation was the same as that of Control 1.

Figure S2 Identification of AD model cells. (A-B) Representative western blot (A) and quantification (B) of APP expression in SH-SY5Y and SH-SY5Y-APPswe cells. SH-SY5Y-APPswe cells expressed more APP proteins. *** $P < 0.001$. 

Figure S3 The 3D model of mutant gelsolin proteins. (A) The 3D model of human gelsolin wild type (GSN-WT) from the Protein Data Bank (No. 3ffn). (B) The different 3D configurations between GSN-WT and GSN D214N mutant proteins. The arrows indicate the difference in protein structures between WT and mutant proteins. The different colours of proteins represent different secondary types of proteins. Amino acids of interest are marked in yellow. The orange square indicates the amplification regions. (C-E) Similar 3D configurations between GSN-WT and GSN mutant proteins (D214N (B), R188I (C), Y301C (D), and R525H (E)).
Figure S4 (A) The gelsolin level changes with age in AD patients and NC. The plasma gelsolin level was compensatively increased in the age between 61 to 80 years. In contrast, the gelsolin level was lower than that in the control in the stage more than 80 years. (B) The ROC curve of plasma gelsolin levels between AD patients and NC (age 61-80 years) showed a moderate test power to distinguish AD from controls. (C) The ROC curve of plasma gelsolin levels between AD patients with GSN frameshift mutation carriers and noncarriers showed good test power to distinguish carriers from noncarriers. *P < 0.05 versus the NC group, and ****P < 0.001.

Figure S5 (A) The negative correlation between CSF gelsolin and CSF p-tau181 in AD patients analyzed by Pearson correlation analysis. n = 82. (B) The positive correlation between CSF gelsolin and CSF t-tau in AD patients analyzed by Pearson correlation analysis. n = 82. (C) The negative correlation between plasma gelsolin and plasma Aβ42 in AD patients analyzed by Pearson correlation analysis. n = 124.
Figure S6 The relationship between gelsolin and brain atrophy. (A-B) The gelsolin level (plasma (A) and CSF (B)) changes with GCA-F scores in AD patients. (C-D) The gelsolin level (plasma (C) and CSF (D)) changes with PA scores in AD patients. (E-F) The gelsolin level (plasma (E) and CSF (F)) changes with MTA scores in AD patients.
## Supplementary Table

### Table S1 Frameshift mutations of the GSN gene in the CI-MDCNC cohort

<table>
<thead>
<tr>
<th>AA Change</th>
<th>SNP</th>
<th>Mutation type</th>
<th>gnomAD_e xome_ALL</th>
<th>gnomAD_e xome_EAS</th>
<th>ExAC_ALL</th>
<th>ExAC_EAS</th>
<th>Reve</th>
<th>Functional predictions: pathogenic (total)</th>
<th>Patients</th>
<th>Controls</th>
</tr>
</thead>
<tbody>
<tr>
<td>exon1:c.8_35del:p.P3fs</td>
<td>rs764841269</td>
<td>Frameshift</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>NA</td>
<td>NA</td>
<td>10</td>
<td>0</td>
</tr>
<tr>
<td>exon10:c.1408delG:p.G470fs</td>
<td>rs758660795</td>
<td>Frameshift</td>
<td>4.06E-06</td>
<td>5.79E-05</td>
<td>8.24E-06</td>
<td>1.00E-04</td>
<td>NA</td>
<td>NA</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>exon12:c.1655dupC:p.S552fs</td>
<td>rs769989772</td>
<td>Frameshift</td>
<td>1.25E-05</td>
<td>2.00E-04</td>
<td>1.91E-05</td>
<td>3.00E-04</td>
<td>NA</td>
<td>NA</td>
<td>2</td>
<td>0</td>
</tr>
</tbody>
</table>

Carriers (n)  
13 0

Frequency (%)  
1.47% 0.00%

\( P \) value  
<0.001***