

Meta-Analysis of PNPLA3 and TM6SF2 Gene Polymorphisms and Susceptibility to Fatty Liver Disease

INPLASY2026100080

doi: 10.37766/inplasy2026.10.0080

Received: 17 October 2026

Published: 17 October 2026

Liu, YH; Tang, MP; Zhang, YN; Li, WR; Cao, YP; Wang, ZM ;Wei, W; Wu ,LP.

Corresponding author:

Liping Wu

wuliping1985116@163.com

Author Affiliation:

Department of Gastroenterology,
The Third People's Hospital of
Chengdu, the Affiliated Hospital of
Southwest Jiaotong University.

ADMINISTRATIVE INFORMATION

Support - This study was supported by grants from the Third People's Hospital of Chengdu Scientific Research Project (2023PI20) and the Third People's Hospital of Chengdu Clinical Research Program (CSY-YN-03-2024-005).

Review Stage at time of this submission - Completed but not published.

Conflicts of interest - None declared.

INPLASY registration number: INPLASY2026100080

Amendments - This protocol was registered with the International Platform of Registered Systematic Review and Meta-Analysis Protocols (INPLASY) on 17 October 2026 and was last updated on 17 October 2026.

INTRODUCTION

Review question / Objective Among children, adolescents and adults evaluated for fatty liver disease (P), are PNPLA3 rs738409 G or TM6SF2 rs58542926 T carriage (I, interpreted as exposure), compared with the corresponding CC genotype (C), associated with greater odds of fatty liver disease (O) in observational studies (S)? Additional objectives concern robustness and study-level heterogeneity across age, ethnicity, BMI and geographical groups.

Rationale Previous meta-analyses have examined these associations, including a study published by Tong and Wang in 2020. However, differences in disease definitions, variations in study population characteristics, sparse genotype data, sample overlap, and uncertainties in data extraction all complicate the interpretation of the results. This

review employed a consistent statistical framework to evaluate these two variants separately and conducted sensitivity analyses for records where there were uncertainties. However, this parallel evaluation does not establish the practicality of combined genetic effects or clinical screening.

Condition being studied Fatty liver disease reported as nonalcoholic fatty liver disease (NAFLD), metabolic dysfunction-associated fatty liver disease (MAFLD), or metabolic dysfunction-associated steatotic liver disease (MASLD). Original diagnostic definitions are retained because these populations cannot automatically be considered equivalent.

METHODS

Search strategy Searches covered PubMed, Web of Science, Wiley's publisher platform, CNKI, Wanfang and the reported Chinese Medical

Association database, from inception to September 20, 2026, with reference-list screening; We used the following search terms: (TM6SF2) OR (Transmembrane 6 superfamily 2 human gene), (PNPLA3) OR (Patatin-like phospholipase domain-containing protein 3), (NAFLD) OR (MASLD), (case-control) OR (controls), (I148M) OR (rs738409), (E167K) OR (rs58542926), and their corresponding Chinese search terms, and used possible permutations and combinations to narrow the search scope.

Participant or population Children, adolescents and adults from different ethnic and geographical populations, including individuals with fatty liver disease diagnosed using imaging, histopathology or explicit clinical criteria, and corresponding controls. Control definitions follow the original studies; uncertain or potentially misclassified controls are flagged for sensitivity analysis.

Intervention Not applicable as a clinical intervention: no treatment or preventive intervention is evaluated.

Comparator Applicable as a genetic reference. The primary comparisons use PNPLA3 CC and TM6SF2 CC as reference genotypes. Disease-status comparisons involve cases and controls defined by the original studies. Other genetic contrasts are secondary or exploratory.

Study designs to be included Observational genetic association studies, including case-control, cohort, and cross-sectional studies, are eligible. Studies must evaluate PNPLA3 rs738409 and/or TM6SF2 rs58542926 in relation to fatty liver disease and provide genotype counts or sufficient data to estimate odds ratios and 95% confidence intervals for comparisons between cases and controls. For studies originating from clinical trial cohorts, only eligible observational genetic association data are considered; treatment effects are not evaluated.

Eligibility criteria Case-control, cohort and cross-sectional studies providing analysable disease-status and genotype comparisons for either variant. Eligible reports supply genotype counts or sufficient information to estimate odds ratios and confidence intervals. Reviews, editorials, animal studies, laboratory experiments and conference abstracts lacking usable data are excluded.

Information sources The manuscript reports searches of PubMed, Web of Science, CNKI, Wanfang and a Chinese Medical Association

database, from inception to September 20, 2026, supplemented by reference-list screening.

Main outcome(s) The primary outcome is fatty liver disease status at the assessment reported in each study. Associations are expressed as unadjusted odds ratios under the dominant model: PNPLA3 CG/GG versus CC and TM6SF2 CT/TT versus CC. Separate REML random-effects estimates, modified Hartung–Knapp 95% confidence intervals and 95% prediction intervals are reported. No common follow-up interval is specified; disease progression and treatment response are not primary outcomes.

Quality assessment / Risk of bias analysis The manuscript reports independent Newcastle–Ottawa Scale assessments by two reviewers, with disagreements adjudicated by a third. Selection, comparability and exposure/outcome ascertainment are evaluated. Scores of 7–9, 4–6 and 0–3 represent study-defined descriptive categories.. Control-group Hardy–Weinberg equilibrium is assessed separately using an exact test; deviation alone does not establish bias.

Strategy of data synthesis Each variant is analysed separately using unadjusted odds ratios. The dominant model is primary; allelic, homozygous, heterozygous and recessive models are supplementary. REML random-effects pooling uses modified Hartung–Knapp confidence intervals and reports Q , I^2 , τ^2 and 95% prediction intervals. Identified dependent comparisons are addressed before pooling. A 0.5 correction applies to all four cells only in tables containing zero cells; uninformative double-zero comparisons are excluded. Peters testing assesses dominant-model small-study effects, supplemented by Egger and standardized Begg tests. Study IDs and DOIs link publications and comparisons. Extraction records include population characteristics, diagnostic criteria, sample sizes and genotype counts. Directly extracted data are classified as Grade A; reconstructed data as Grade B. Overlapping samples, shared controls and uncertain counts are flagged. These grades describe data acquisition, not methodological quality.

Subgroup analysis Exploratory analyses examine age, ethnicity, BMI (<25 , $25\text{--}30$ and ≥ 30 kg/m²), and China versus other regions. Meta-regression uses common residual heterogeneity, modified Hartung–Knapp covariance and F tests. Holm correction covers 16 omnibus tests separately within the main, conservative and Grade A sets; BMI pairwise comparisons receive separate

correction. Study-level differences cannot establish individual interactions.

Sensitivity analysis Analyses exclude questionable records, restrict to Grade A data within the conservative set, and apply HWE and original-NAFLD-label restrictions. Additional checks exclude preprints or theses, retain one comparison per publication, omit publications sequentially, and use DerSimonian–Laird estimation. Sparse TM6SF2 models receive Mantel–Haenszel and conditional exact analyses. Further checks address uncertain sources, allelic variance, single-study elderly strata and jointly reporting publications. Each variant is analysed separately using unadjusted odds ratios. The dominant model is primary; allelic, homozygous, heterozygous and recessive models are supplementary. REML random-effects pooling uses modified Hartung–Knapp confidence intervals and reports Q , I^2 , τ^2 and 95% prediction intervals. Identified dependent comparisons are addressed before pooling. A 0.5 correction applies to all four cells only in tables containing zero cells; uninformative double-zero comparisons are excluded. Peters testing assesses dominant-model small-study effects, supplemented by Egger and standardized Begg tests.

Language restriction Search using English and the corresponding Chinese search terms.

Country(ies) involved All participants are from China.

Keywords PNPLA3; TM6SF2; NAFLD; MAFLD; MASLD; genetic susceptibility; systematic review; meta-analysis.

Contributions of each author

Author 1 - Yuhao Liu.

Email: liuyh8413@163.com

Author 2 - Mingpu Tang.

Email: 2622348350@qq.com

Author 3 - Yanan Zhang.

Email: 982547527@qq.com

Author 4 - Wenrui Li.

Email: 1978803450@qq.com

Author 5 - Yuping Cao.

Email: 2046638044@qq.com

Author 6 - Zhiming Wang.

Email: wangzm42@126.com

Author 7 - Wei Wen.

Email: mirsquito@163.com

Author 8 - Liping Wu.

Email: wuliping1985116@163.com