Original Manuscript

SCIENCE PROGRESS

Science Progress
2021, Vol. 104(3) 1–17

DOI: 10.1177/00368504211043766
journals.sagepub.com/home/sci
SAGE

Association between *TFEB* gene polymorphism, gene–environment interaction, and fatty liver disease: A case–control study in China

Chunbao Mo[1,#], Tingyu Mai[2,#], Jiansheng Cai[3,#], Haoyu He[4], Huaxiang Lu[5,6], Xu Tang[5], Quanhui Chen[7], Xia Xu[5], Chuntao Nong[8], Shuzhen Liu[5], Dechan Tan[2], Qiumei Liu[5], Min Xu[5], You Li[2], Zhiyong Zhang[2,5] and Jian Qin[5]

[1]School of Medicine, Southern University of Science and Technology, Shenzhen, Guangdong, China
[2]Department of Environmental Health and Occupational Medicine, School of Public Health, Guilin Medical University, Guilin, Guangxi, China
[3]Guangxi Key Laboratory of Tumor Immunology and Microenvironmental Regulation, Guilin Medical University, Nanning, Guilin, China
[4]Department of Quality Management, The Affiliated Hospital of Stomatology, Guangxi Medical University, Nanning, Guangxi, China
[5]Department of Environmental and Occupational Health, School of Public Health, Guangxi Medical University, Nanning 530021, Guangxi, China
[6]Department of Guangxi Science and Technology Major Project, Guangxi Zhuang Autonomous Region Center for Diseases Control and Prevention, Nanning, Guangxi, China
[7]Department of Hospital Infection-Control, Liuzhou Workers' Hospital, Liuzhou, Guangxi, China
[8]Nanning Municipal Center for Disease Control and Prevention, Nanning, Guangxi, China

Corresponding authors:
Jian Qin, Mailing address: Department of Environmental and Occupational Health, School of Public Health, Guangxi Medical University, Shuangyong Road 22, Nanning 530021, Guangxi, China.

Zhiyong Zhang, Mailing address: Department of Environmental Health and Occupational Medicine, School of Public Health, Guilin Medical University, Huan Cheng North 2nd Road 109, Guilin 541004, Guangxi, China. Email: qinjian@gxmu.edu.cn; rpazz@163.com

[#]These authors contributed equally to this work.

Abstract

Background: Fatty liver disease (FLD) is a serious public health problem that is rapidly increasing. Evidence indicates that the transcription factor EB (*TFEB*) gene may be involved in the pathophysiology of FLD; however, whether *TEFB* polymorphism has an association with FLD remains unclear.
Objectives: To explore the association among *TFEB* polymorphism, gene–environment interaction, and FLD and provide epidemiological evidence for clarifying the genetic factors of FLD.
Methods: This study is a case–control study. Sequenom MassARRAY was applied in genotyping. Logical regression was used to analyze the association between *TFEB* polymorphism and FLD, and the gene–environment interaction in FLD was evaluated by multiplication and additive interaction models.
Results: (1) The alleles and genotypes of each single nucleotide polymorphism and haplotypes of *TFEB* in the case and control groups were evenly distributed; no statistically substantial difference was observed. (2) Logistic regression analysis indicated that *TFEB* polymorphism is not remarkably associated with FLD. (3) In the multiplicative interaction model, rs1015149, rs1062966, rs11754668 and rs2273068 had remarkable interaction with the amount of cigarette smoking. Rs1062966 and rs11754668 also had a considerable interaction body mass index and alcohol intake, respectively. However, no remarkable additive interaction was observed.
Conclusion: *TFEB* polymorphism is not directly associated with FLD susceptibility, but the risk can be changed through gene–environment interaction.

Keywords

TFEB, gene polymorphism, gene–environment interaction, fatty liver disease

Introduction

Health problem has gradually aroused people's concern with the advancement of economy and society and the improvement of living standard. Fatty liver disease (FLD) is a chronic disease and a serious public health problem that is rapidly increasing.[1,2] FLD is the pathological process of excess adipose accumulation in liver cells, caused by many factors, such as disease and drugs. Simple hepatic steatosis may transform into steatohepatitis or cirrhosis as FLD progresses. According to etiology, FLD is classified as nonalcoholic fatty liver disease (NAFLD) and alcoholic liver disease (ALD). Studies have indicated that the pathogenesis of FLD is affected by abnormal fat metabolism, immune response, environmental, genetic, and other factors.[3] Currently, specific medicine for FLD is deficient; the effective prevention and control measures for FLD are early detection and intervention, including diet control and exercise; and alcohol abstinence is the chief measure for patients with ALD.[3]

Similar to most diseases, FLD is likely influenced by environmental and genetic factors. Recently, transcription factor EB (TFEB) has attracted extensive attention in the study of autophagy mechanisms. *TFEB* is the main gene involved in lysosome biosynthesis and encodes TFEB, which is an important regulatory factor for autophagy and lysosomal biosynthesis. TFEB is considered the main activator for autophagy–lysosomal gene transcription and refers to inflammation,[4] cell autophagy,[5] lipid metabolism,[6] and other biological processes. Previous researches indicated that TFEB can regulate the expression of many genes related to lipid degradation, such as cluster of differentiation 36 (CD36), fatty acid-binding proteins (FABPs), and carnitine acetyltransferase (Crat),[7,8]

and TFEB can also regulate lipid degradation factors: peroxisome proliferator-activated receptor alpha (PPARα) and proliferator-activated receptor gamma coactivator 1 alpha (PGC-1α).[9,10] Furthermore, pieces of evidence directly manifest that TFEB can protect mice liver from alcohol-induced damage,[11] and promoting TFEB-mediated lysosomal biogenesis using formononetin can ameliorate the fatty disease process in mice liver.[12]

Single nucleotide polymorphisms (SNPs) are the most common form of mutations in the human genome. Studies have found that SNPs are associated with FLD (e.g. *GCKR* rs780094, *pNPLA3* rs738409, and *TM6SF2* rs58542926).[13–15] Nevertheless, few studies have focused on the relationship between *TFEB* and FLD. Therefore, a case–control approach was adopted in this study to explore the association between *TFEB* polymorphism and FLD, and the gene–environment interaction was evaluated to provide epidemiological evidence of the genetic factors of FLD.

Materials and methods

Study design and population

This case–control study included 228 patients with FLD diagnosed by the presence of at least two of three abnormal findings on abdominal ultrasonography: liver anterior echogenicity ("bright liver"), far-field echo attenuation, and unclear intrahepatic duct structure.[16] Individuals with liver diseases, tumors, and autoimmune diseases caused by drugs and viruses were excluded. A total of 342 healthy individuals who were matched by sex and age (with a variation of ±3 years) in a proportion of 1:1.5 were selected as the control group. All the participants were permanent residents in Gongcheng County, Guilin City, Guangxi Zhuang Autonomous Region, People's Republic of China and signed the informed consent voluntarily after fully understanding the research content and importance of this project. Our research protocol was approved by the Ethics Committee of Guilin Medical University.

Data collection

All the participants were required to answer a questionnaire from a trained researcher to collect information on demography, behavior, exercise, disease history, nutritional diet, and other data. Behavioral factors include smoking and alcohol consumption. The amount of smoking is expressed in pack year, that is, the number of packs (20 cigarettes per bag) per day multiplied by the number of years of smoking. In addition to daily alcohol intake, we also assessed the daily intake of 12 types of food, including cereals and their products, potatoes, vegetables, and fruits, via a dietary survey. The participants were also examined by professional physicians to collect anthropometric indicators, such as height, weight, waist circumference, and blood pressure. Venous blood was collected for subsequent tests and experiments.

Biochemical testing

Two venous blood samples were collected from each participant. One set of samples was tested for biochemical indicators, including triglyceride (TG), total cholesterol (TC), high-density lipoprotein cholesterol (HDL-C), low-density lipoprotein cholesterol (LDL-C), alanine aminotransferase (ALT), and uric acid (UA), by using an automatic biochemical analyzer (Hitachi 7600-020, Kyoto, Japan) at a local hospital. The other set of samples was used in the SNP typing experiment.

Selection and genotyping of SNPs

Functional and validated SNP screening strategies were utilized to screen the target SNPs. This strategy focused on important functional loci and was supplemented by susceptible loci. Finally, the rs1015149, rs1062966, rs11754668, rs14063 and rs2273068 of the *TFEB* gene were selected as the target SNPs for genotyping.

Genomic DNA was isolated from venous blood using a commercial DNA extraction kit (Tiangen, Beijing, China). SNP genotyping was performed using the Sequenom MassARRAY matrix-assisted laser desorption ionization time-of-flight mass spectrometry platform (Sequenom, Inc., San Diego, CA, USA). The primers were designed and synthesized by Bio Miao Biological Technology Co., Ltd (Table S1).

Statistical analysis

Descriptive statistics for continuous and categorical variables were conducted using mean ± standard deviation (SD) and frequency (proportion), respectively. Student's t-test and chi-square test were applied to compare the differences among two groups and genotype subgroups. Pearson's chi-square test was utilized to evaluate the Hardy–Weinberg equilibrium (HWE) before analyzing SNPs data. The samples were considered representative when $p > 0.05$. Haploview was used for haplotype analysis on the possible haplotype and mapping according to the method described by Barrett et al.[17] We performed logistic regression to estimate the effects of genotypes and gene–environment multiplicative interactions on FLD, and odds ratio (OR) and 95% confidence interval (95% CI) were calculated. The test level $\alpha = 0.05$. However, logistic regression was limited to estimate additive interactions; hence, relative excess risk due to interaction (RERI), the attributable proportion of interaction (AP), synergy index (SI), and their 95% CIs were calculated. Additive interactions were considered statistically significant when the 95% CI of RERI and AP did not include 0, and the 95% CI of SI did not contain 1.[18] SPSS 25.0 (IBM, Chicago, IL, USA) and PLINK 1.90 software were used to implement general statistical analysis and gene polymorphism analysis. In addition, R software 4.0.2 and "epiR" package were utilized to complete the calculation of RERI, AP, and SI.

Results

Characteristics of the participants

The demographic and behavioral characteristics of the participants are listed in Table 1. The subjects have a total number of 570 and a roughly equal gender proportion. The age range is 30–83 years with an average of 58.15 years. The majority of the subjects (78.42%) belong to the Yao population. No remarkable differences in gender, age, ethnicity, marital status, hypertension, smoking, drinking, and other factors were observed between the control and case groups at baseline ($p > 0.05$). However, the proportion of subjects with a history of hypertension and the average daily sitting time were significantly higher in the case group than in the control group ($p < 0.05$).

The dietary situation is shown in Table 2. Vegetables and cereals and their products were the main daily dietary intake of the participants. No significant difference was observed in the daily food intake of the two groups ($p > 0.05$).

The clinical indicators are exhibited in Table 3. The two groups showed no statistically significant difference in AST ($t = -1.415$, $p = 0.158$). However, HDL-C was significantly lower in the case group than in the control group, and the other indicators were significantly higher in the case group compared with the control group ($p < 0.05$).

Results indicate that the demographic, behavioral, and dietary variables in the two groups were matched preferably.

Basic information of SNPs

Rs1015149, rs1062966, rs11754668, rs14063, and rs2273068 of *TFEB* are located in chromosome 6 and at the crucial regions that may impact genetic transcription and regulation, such as exon synonymous, promoter, and 3′ untranslated region (3′ UTR). The minimum allele frequency (MAF) of each locus was larger than 0.05, and all the loci of SNPs were consistent with HWE ($p_{HWE} > 0.05$). Therefore, the study subjects are representative (Table 4).

Genotypic frequency and the associations between genotypes and FLD

The alleles and genotypes of each SNP of *TFEB* in the two groups were evenly distributed. No statistically significant difference was observed ($p > 0.05$, Table 5).

FLD was regarded as the dependent variable. Co-dominant, dominant, and recessive models were used for logistic regression analysis using gender and age as adjustment factors. The result indicated that no significant correlations exist between genotypes and FLD in the co-dominant, dominant, and recessive models ($p > 0.05$, Table 6).

Haplotype analysis

Referencing the criteria of $D' \geq 0.8$ and $R^2 \geq 0.2$, strong linkage disequilibrium (LD) was observed among the four SNPs: rs14063, rs1062966, rs2273068, and rs1015149 in one

Table 1. Demographic and behavioral characteristics of the study population.

Variables		Control group (*n* %)	Case group (*n* %)	χ^2/t	*p value*
Gender	Male	192 (56.14)	129 (56.58)	0.011	0.918
	Female	150 (43.86)	99 (43.42)		
Age (years)[a]		58.50 ± 12.75	57.64 ± 11.93	0.806	0.42
Nation	Han	58 (16.96)	40 (17.54)	0.194	0.907
	Yao	268 (78.36)	179 (78.51)		
	Zhuang and others	16 (4.68)	9 (3.95)		
Marital status	No partner	49 (14.33)	28 (12.28)	0.491	0.484
	Have a partner	293 (85.67)	200 (87.72)		
Education	Primary school and below	223 (65.20)	131 (57.46)	3.49	0.062
	Junior high school and above	119 (34.80)	97 (42.54)		
Occupation	Farmer	309 (90.35)	210 (92.11)	0.517	0.472
	Others	33 (9.65)	18 (7.89)		
Household income (Yuan)	<5000	88 (25.73)	66 (28.95)	0.718	0.397
	≥5000	254 (74.27)	162 (71.05)		
Diabetes	No	333 (97.37)	213 (93.42)	5.285	0.022*
	Yes	9 (2.63)	15 (6.58)		
Hypertension	No	284 (83.04)	180 (78.95)	1.514	0.218
	Yes	58 (16.96)	48 (21.05)		
Smoking	No	273 (79.82)	192 (84.21)	1.751	0.186
	Yes	69 (20.18)	36 (15.79)		
Drinking	No	221 (64.62)	150 (65.79)	0.082	0.774
	Yes	121 (35.38)	78 (34.21)		
Smoking amount (pack year)[a]		6.37 ± 17.69	5.03 ± 15.31	0.934	0.351
Alcohol intake (g/day)[a]		17.84 ± 43.75	16.39 ± 34.87	0.419	0.675
Strenuous physical activity (h/day)[a]		1.19 ± 6.04	1.65 ± 8.59	−0.743	0.458
Moderate physical activity (h/day)[a]		4.54 ± 11.43	4.49 ± 11.18	0.049	0.961
Daily walking time (h)[a]		3.38 ± 2.24	3.09 ± 2.14	1.526	0.128
Daily sitting time (h)[a]		3.50 ± 1.76	3.84 ± 1.83	−2.212	0.027

SD: standard deviation.

[a]Mean ± SD; *$p < 0.05$ was considered statistically significant.

haploblock (Fig. S1). Therefore, five haplotypes (G-C-C-C, A-C-C-T, G-T-C-C, A-C-T-T, and G-C-C-T) were constructed in accordance with the LD analysis result. The result in Table 7 shows that all haplotype frequencies did not show statistically significant differences between the control and case groups ($p > 0.05$). Similarly, the regression analysis indicated that five haplotypes were also insignificantly associated with FLD ($p > 0.05$, Table 7).

Table 2. Daily food intake of the participants.

Food category	Control group	Case group	*t*	*p value*
Cereals and their products	262.63 ± 168.13	265.55 ± 152.47	−0.211	0.833
Potatoes	39.72 ± 91.08	36.82 ± 78.86	0.392	0.695
Vegetables	325.79 ± 336.87	295.53 ± 279.52	1.123	0.262
Fruits	256.25 ± 282.77	251.94 ± 249.43	0.187	0.852
Beans and their products	36.04 ± 51.40	36.04 ± 48.59	−0.001	0.999
Nuts	12.69 ± 23.49	14.47 ± 29.42	−0.803	0.422
Meat and poultry	87.23 ± 94.64	77.58 ± 79.25	1.271	0.204
Fish and aquatic products	17.28 ± 29.14	18.77 ± 30.39	−0.585	0.559
Milk and their products	34.85 ± 35.59	37.18 ± 37.40	−0.752	0.452
Eggs and their products	34.74 ± 76.36	30.96 ± 64.04	0.617	0.538
Cooking oil	35.71 ± 27.58	33.94 ± 25.23	0.776	0.438
Salt	9.93 ± 8.24	9.11 ± 6.24	1.276	0.202

Note: Data are expressed as mean ± SD. The unit for each food category is gram. SD: standard deviation.

Table 3. Clinical indicators of the study population.

Clinical indicators	Control group	Case group	*t*	*p value*
WC	76.78 ± 9.14	88.30 ± 8.40	−15.23	<0.001**
BMI	21.99 ± 3.44	28.40 ± 27.07	−3.558	<0.001**
SBP	134.39 ± 23.15	138.47 ± 25.63	−1.972	0.049*
DBP	80.72 ± 13.59	85.22 ± 16.88	−3.512	<0.001**
HbA1C	5.82 ± 0.90	6.17 ± 1.25	−3.595	<0.001**
LDL-C	3.30 ± 0.94	3.63 ± 0.99	−4.026	<0.001**
HDL-C	1.76 ± 0.39	1.58 ± 0.37	5.559	<0.001**
TC	5.40 ± 1.03	5.80 ± 1.03	−4.533	<0.001**
TG	1.21 ± 1.28	2.18 ± 1.86	−6.875	<0.001**
GLU	4.87 ± 1.24	5.37 ± 1.85	−3.543	<0.001**
ALB	43.77 ± 2.34	44.50 ± 2.29	−3.67	<0.001**
ALT	19.29 ± 12.29	25.81 ± 13.06	−5.976	<0.001**
AST	23.55 ± 10.91	24.91 ± 11.77	−1.415	0.158
UA	306.72 ± 91.26	373.28 ± 100.87	−8.175	<0.001**

Note: Data are expressed as mean ± SD; *$p < 0.05$ and **$p < 0.001$ were considered statistically significant. WC: waist circumference (cm); BMI: body mass index; SBP/DBP: systolic/diastolic blood pressure (mmHg); HbA1C: glycosylated hemoglobin (%); LDL-C: low-density lipoprotein cholesterol (mmol/L); HDL-C: high-density lipoprotein cholesterol (mmol/L); TG: triglyceride (mmol/L; TC: total cholesterol (mmol/L); GLU: fasting plasma glucose (mmol/L); ALB: albumin (g/L); ALT: alanine aminotransferase (U/L); AST: aspartate transaminase (U/L); UA: uric acid (μmol/L); SD: standard deviation.

Interactions between environmental factors and SNP in FLD

Multiplicative and additive models were used to evaluate the interactions of each SNP locus with environmental factors, including diabetes, cigarette smoking, alcohol intake, daily sitting time, waist circumference, and body mass index (BMI). Compared with single genes, some gene–environment interactions were remarkably associated with

Table 4. Basic information of SNPs.

Chr	SNPs	Alleles	Region	MAF	Control group p_{HWE}	Case group p_{HWE}
6	rs1015149	C/T	Exon-synonymous	0.312	0.095	0.441
6	rs1062966	C/T	Exon-synonymous	0.188	0.214	0.275
6	rs11754668	C/G	promoter	0.122	1.000	0.288
6	rs14063	G/A	3′ UTR	0.423	0.731	1.000
6	rs2273068	C/T	Exon-synonymous	0.093	1.000	0.739

Note: Chr: chromosome; 3′ UTR: 3′ untranslated region; MAF: minimum allele frequency; p_{HWE}: p value of Hardy–Weinberg equilibrium; SNPs: single nucleotide polymorphisms.

Table 5. Descriptive statistics of *TEFB* genotypes.

SNPs	Alleles/Genotypes	Control group (*n* %)	Case group (*n* %)	χ^2	*p value*
rs1015149	C	390 (0.57)	267 (0.59)	0.210	0.647
	T	292 (0.43)	189 (0.41)		
	CC	114 (33.33)	78 (34.21)	1.018	0.797
	CT	162 (47.37)	111 (48.68)		
	TT	65 (19.01)	39 (17.11)		
rs1062966	C	550 (0.81)	369 (0.81)	0.004	0.947
	T	128 (0.19)	85 (0.19)		
	CC	221 (64.62)	147 (64.47)	0.727	0.867
	CT	108 (31.58)	75 (32.89)		
	TT	10 (2.92)	5 (2.19)		
rs11754668	C	624 (0.92)	406 (0.89)	2.404	0.121
	G	56 (0.08)	50 (0.11)		
	CC	286 (83.63)	181 (79.39)	3.828	0.281
	GC	52 (15.20)	44 (19.30)		
	GG	2 (0.58)	3 (1.32)		
rs14063	G	463 (0.68)	312 (0.69)	0.036	0.849
	A	213 (0.32)	140 (0.31)		
	AA	28 (8.19)	19 (8.33)	0.208	0.976
	AG	157 (45.91)	102 (44.74)		
	GG	153 (44.74)	105 (46.05)		
rs2273068	C	590 (0.87)	406 (0.89)	1.806	0.179
	T	90 (0.13)	48 (0.11)		
	CC	254 (74.27)	183 (80.26)	3.675	0.299
	CT	82 (23.98)	40 (17.54)		
	TT	4 (1.17)	4 (1.75)		

SNPs: single nucleotide polymorphisms.

FLD susceptibility. In the multiplicative interaction model, rs1015149, rs1062966, rs11754668 and rs2273068 had substantial interaction with the smoking amount. Among them, rs1062966 and rs11754668 also had remarkable interaction with BMI and alcohol intake, respectively.

Table 6. Logistic regression analysis between *TEFB* polymorphism and FLD.

Genotype	β	*S.E.*	Wald χ^2	*p* value	*OR*	95%*CI*
rs1015149						
CC			0.376	0.829	1.000	
CT	−0.009	0.192	0.002	0.962	0.991	0.679–1.445
TT	−0.142	0.251	0.321	0.571	0.868	0.531–1.418
Dominant model TT + CT vs. CC	0.046	0.181	0.063	0.802	1.047	0.733–1.494
Recessive model TT vs. CT + CC	0.137	0.224	0.373	0.541	1.146	0.740–1.777
rs1062966						
CC			0.364	0.834	1.000	
CT	0.049	0.185	0.071	0.789	1.051	0.732–1.508
TT	−0.284	0.559	0.258	0.612	0.753	0.252–2.250
Dominant model TT + CT vs. CC	−0.025	0.180	0.019	0.889	0.975	0.685–1.388
Recessive model TT vs. CT + CC	0.300	0.555	0.293	0.589	1.350	0.455–4.008
rs11754668						
CC			2.507	0.286		
GC	0.293	0.226	1.678	0.195	1.340	0.861–2.087
GG	0.888	0.919	0.933	0.334	2.430	0.401–14.730
Dominant model GG + GC vs. CC	−0.322	0.221	2.120	0.145	0.725	0.470–1.118
Recessive model GG vs. GC + CC	0.322	0.221	2.120	0.145	1.380	0.895–2.129
rs14063						
AA			0.150	0.928		
AG	0.003	0.326	0.000	0.992	1.003	0.529–1.903
GG	−0.066	0.180	0.136	0.712	0.936	0.658–1.331
Dominant model AA + AG vs. GG	0.056	0.173	0.104	0.747	1.057	0.754–1.484
Recessive model AA vs. AG + GG	−0.037	0.314	0.014	0.906	0.964	0.521–1.782
rs2273068						
CC			3.896	0.143		
CT	−0.413	0.218	3.605	0.058	0.662	0.432–1.013
TT	0.297	0.715	0.173	0.678	1.346	0.331–5.469
Dominant model TT + CT vs.CC	0.366	0.211	3.009	0.083	1.442	0.954–2.180
Recessive model TT vs. CT + CC	−0.389	0.713	0.297	0.586	0.678	0.168–2.745

FLD: fatty liver disease; 95% CI: 95% confidence interval; OR: odds ratio.

Table 7. Haplotype frequencies and logistic regression analysis between haplotypes and FLD.

Haplotypes	Control Freq.	Case Freq.	χ^2	p^a	β	p^b	*OR (95%CI)*
G-C-C-C	38.30%	39.70%	0.235	0.628	0.073	0.674	1.076 (0.765–1.513)
A-C-C-T	18.20%	20.50%	0.977	0.323	−0.108	0.654	0.898 (0.561–1.437)
G-T-C-C	18.70%	18.80%	0.001	0.979	−0.042	0.852	0.959 (0.616–1.493)
A-C-T-T	12.90%	10.50%	1.453	0.228	−0.160	0.697	0.852 (0.382–1.903)
G-C-C-T	11.20%	10.30%	0.212	0.645	0.029	0.880	1.029 (0.708–1.496)

Note: Haplotypes consisted of these four SNPs (rs14063, rs1062966, rs2273068, and rs1015149) of *TFE*; adjusting by gender and age; Freq.: frequencies (%); p^a: *p* value of chi-square test; p^b: *p* value of logistic analysis; FLD: fatty liver disease; 95% CI: 95% confidence interval; OR: odds ratio.

Notably, the 95% CIs of RERI and AP contained 0, and the 95% CI of SI contained 1; thus, all gene–environment additive interactions with FLD were not statistically significant (Table 8, Table S3).

Discussion

In this case–control study, we analyzed the association of *TFEB* polymorphisms and FLD and assessed gene–environment interactions to provide epidemiological evidence of the genetic factors related to the occurrence and development of FLD. Results showed that the alleles and genotypes of each SNP and haplotypes of *TFEB* in the case and control groups were evenly distributed; no statistically significant difference was observed. Logistic regression analysis indicated that *TFEB* polymorphism is not substantially associated with FLD. Previous studies have suggested that autophagy plays an important role in maintaining liver homeostasis.[19] *TFEB* knockout in mice may result in the hepatic accumulation of fatty acid-β and impaired oxidation in hepatocytes, which lead to elevated fatty acid and glycerol levels and lipid metabolism disorders in hepatocytes.[7] According to the results, *TFEB* may be involved in the pathophysiological basis of FLD. However, the relationship between *TFEB* polymorphism and FLD was not observed in this study; thus, these SNPs and haplotypes may not affect the normal expression of *TFEB*.

Gene–environment interaction plays an important role in the occurrence and development of complex diseases, such as FLD. Zhu et al. demonstrated that the gene–gene interaction between *AGTR1* and *PPARγ* is associated with the occurrence of NAFLD in the Chinese population.[20] Zhang et al. found that people with 11391G/A(AA) and EC-SOD (CG + GG) genotypes suffer a higher risk of NAFLD, and these genotypes have an interaction with *Helicobacter pylori* infection.[21] Therefore, the analysis of gene–environment interaction might be conducive to understand etiological factors and guide the prevention and treatment of FLD. The result exhibited that some SNP loci, such as rs1062966, and rs11754668, had positive interactions with smoking, which is a risk factor for FLD susceptibility.[22] This finding is consistent with the results of Zhang et al. on the interactions between *GPX-1* polymorphism and smoking in NAFLD and also agrees with the results of Oniki et al.[23] Interestingly, rs1015149 and rs2273068 had a negative interaction with smoking in FLD; thus, they are considered "protective factors" (OR = 0.96 and 0.97, respectively). Smoking is a recognized risk factor that is remarkably associated with the occurrence of many diseases. However, the relationship between smoking and FLD is not clear yet.[24–27] In this study, we found that smoking might reduce the risk of FLD of individuals who carry the CT + CC genotype of rs1015149 or the TT + CT genotype of rs2273068. However, the results do not "advocate" smoking in these populations for FLD prevention. A more rational explanation for the reduced FLD risk is that compared with individuals who carry other genotypes, people with the CT + CC genotype of rs1015149 or the TT + CT genotype of rs2273068 may be more able to offset the risk of FLD caused by smoking. The same explanation can also be utilized to explain the interaction between rs11754668 and alcohol intake in this study.

Table 8. Results of gene–environment multiplication and additive interactions (only the parts with statistical significance are exhibited).

Gene–environment interaction		Multiplication interactions			Additive interactions		
		B	*p* value	*OR* (95%*CI*)	RERI (95%*CI*)	AP (95%*CI*)	SI (95%*CI*)
rs1015149	RM × Smoking amount	−0.037	0.032*	0.960 (0.930,1.000)	0.003 (−0.025–0.031)	0.002 (−0.017–0.021)	1.007 (0.947–1.072)
rs1062966	DM × Smoking amount	0.027	0.035*	1.030 (1.000,1.050)	0.000 (−0.015–0.015)	0.000 (−0.018–0.018)	1.000 (0.924–1.083)
	DM × BMI	0.270	<0.001**	1.310 (1.140,1.510)	−0.256 (−0.374 to −0.138)	−121.1 (−509.0–266.7)	1.345 (1.145–1.579)
rs11754668	DM × Smoking amount	0.032	0.039*	0.970 (0.940,1.000)	−0.007 (−0.021–0.007)	−0.008 (−0.025–0.010)	1.067 (0.842–1.352)
	DM × Alcohol intake	−0.017	0.037*	0.980 (0.970,1.000)	−0.006 (−0.020–0.007)	−0.007 (−0.022–0.009)	1.098 (0.600–2.010)
	RM × Smoking amount	0.032	0.039*	1.030 (1.000,1.060)	−0.007 (−0.025–0.011)	−0.006 (−0.021–0.008)	0.946 (0.771–1.162)
	RM × Alcohol intake	0.017	0.037*	1.020 (1.000,100.030)	−0.006 (−0.021–0.009)	−0.006 (−0.019–0.008)	0.927 (0.572–1.501)
rs2273068	DM × Smoking amount	−0.028	0.024*	0.970 (0.950,1.000)	0.000 (−0.029–0.029)	0.000 (−0.017–0.017)	1.000 (0.961–1.041)

Note: DM: dominant model; RM: recessive model; 95% CI: 95% confidence interval; OR: odds ratio; SI: synergy index; RERI: relative excess risk due to interaction. *$p < 0.05$ and **$p < 0.001$ were considered as statistically significant.

No significant additive interaction was observed in this study. This result is in agreement with the result of Zhao et al.[28] In other words, additive interaction may not be remarkable even if the factors studied have substantial multiplicative interaction. In fact, the interaction between multiple factors is based on multiplication and synergism, whereas additive interaction is relatively rare. In the field of medicine, the analysis models for gene–environment, gene–gene, and gene–environment–gene interactions, such as cross-generation analysis,[29] multifactor dimensionality reduction (MDR),[30] and generalized MDR,[31] are based on multiplication. Although substantial multiplicative interaction results were not observed in the present study, this study still provided a meaningful attempt to explore the gene–environment interaction in FLD, which might be ignored.

Dietary factors are important influencing factors of FLD. Numerous works of literature have reported the association between different dietary patterns or food intake and the incidence of FLD. For example, the high intake of meat, high-fat dairy products, and refined grains may increase the risk of FLD, whereas a diet based on fruits, vegetables, whole grains, fish, and olive oil can reduce FLD risk.[32,33] A cross-sectional study based on Chinese adolescents illustrated that adolescents who have a traditional Chinese diet have lower risks of FLD compared with those with the Western diet.[34] Therefore, the potential impact of diet on the result needs to be fully considered and controlled to reduce analysis error. In addition, in this study, we ensured data quality and improved the reliability of the results. The advantages are as follows: (1) This study is the first epidemiological study to uncover the associations of *TFEB* polymorphism and gene–environment interaction with FLD. (2) All participants were from the same area with relatively similar genetic backgrounds, living environments, and habits; these similarities were helpful to control potential confounding factors. (3) In terms of grouping, gender and age (±3 years) were adopted in matching to reduce the influence of gender and age on the results to a certain extent.

However, this study also has many deficiencies that need to be further improved. First, the sample size is relatively small, and sampling error is difficult to decrease. Second, the conclusions are based on the population from Gongcheng County. Therefore, the applicability of the conclusions to other populations is limited, and extrapolation is deficient. Third, few studies focused on the loci of SNPs selected in this study. Hence, references on the target SNP selection were insufficient. This insufficiency increases the uncertainty of the research conclusion and enhances the difficulty of interpretation simultaneously. Therefore, an optimized SNP screening strategy is required. Fourth, the degree of FLD was not classified. Thus, the effect of research factors on the process of FLD might have been ignored. Finally, the causal demonstration power is not strong because of the case–control design.

Conclusion

The polymorphisms of the rs1015149, rs1062966, rs11754668, rs14063, and rs2273068 of the *TFEB* gene are not directly associated with FLD susceptibility, but the risk can be changed through the gene–environment interaction.

Acknowledgments

Not applicable.

Authors' contributions

CM, TM, and JC contributed to conception and design: HH, HL, XT, QC, XX, CN, SL, TD, QL, and MX contributed to acquisition of data; CM, TM, JC, and HH contributed to analysis and interpretation of data; CM, TM, JC, LY, and JQ contributed to writing, review, and/or revision of the manuscript; YL, ZZ, and JQ contributed to administrative, technical, or material support; ZZ and JQ contributed to study supervision. All authors approved the final manuscript.

Declaration of conflicting interests

The authors declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The study was supported by the National Natural Science Foundation of China (Grant No. 81760577, 81960583, and 81560523), the Guangxi Science and Technology Development Project (Grant No. AD17129003 and AD18050005), the Guangxi Natural Science Found for Innovation Research Team (2019GXNSFGA245002), the Guangxi Graduate Education Innovation Project (Grant No. YCSW2020230), and the Guangxi Scholarship Fund of Guangxi Education Department of China.

Availability of data and materials

Please contact the authors for reasonable requests.

Ethics approval and consent to participate

Our research protocol was approved by the Ethics Committee of Guilin Medical University. Informed consent was signed by all study participants.

ORCID iDs

Chunbao Mo https://orcid.org/0000-0003-0738-7773
Tingyu Mai https://orcid.org/0000-0001-8150-665X

Supplemental material

Supplemental material for this article is available online. **Fig S1** TFEB pairwise linkage disequilibrium patterns **Table S1** Sequence information of primers **Table S2** Results of gene–environment multiplication and additive interactions.

References

1. Diehl AM, Farpour-Lambert NJ, Zhao L, et al. Why we need to curb the emerging worldwide epidemic of nonalcoholic fatty liver disease. *Nat Metab* 2019; 1: 1027–1029.
2. Perumpail BJ, Khan MA, Yoo ER, et al. Clinical epidemiology and disease burden of nonalcoholic fatty liver disease. *World J Gastroenterol* 2017; 23: 8263–8276.

3. Carr RM, Oranu A and Khungar V. Nonalcoholic fatty liver disease: pathophysiology and management. *Gastroenterol Clin North Am* 2016; 45: 639–652.
4. Brady OA, Martina JA and Puertollano R. Emerging roles for TFEB in the immune response and inflammation. *Autophagy* 2018; 14: 181–189.
5. Settembre C, Di Malta C, Polito VA, et al. TFEB links autophagy to lysosomal biogenesis. *Science* 2011; 332: 1429–1433.
6. Li M, Wang Z, Wang P, et al. TFEB: a emerging regulator in lipid homeostasis for atherosclerosis. *Front Physiol* 2021; 12: 639920.
7. Settembre C, De Cegli R, Mansueto G, et al. TFEB controls cellular lipid metabolism through a starvation-induced autoregulatory loop. *Nat Cell Biol* 2013; 15: 647–658.
8. Settembre C and Ballabio A. Lysosome: regulator of lipid degradation pathways. *Trends Cell Biol* 2014; 24: 743–750.
9. Sardiello M, Palmieri M, di Ronza A, et al. A gene network regulating lysosomal biogenesis and function. *Science* 2009; 325: 473–477.
10. Barbara S, Ryan JS, Shaun GW, et al. The small GTPase Rab7 as a central regulator of hepatocellular lipophagy. *Hepatology* 2015; 61: 1896–1907.
11. Chao X, Wang S, Zhao K, et al. Impaired TFEB-mediated lysosome biogenesis and autophagy promote chronic ethanol-induced liver injury and steatosis in mice. *Gastroenterology* 2018; 155: 865–879.e12.
12. Wang Y, Zhao H, Li X, et al. Formononetin alleviates hepatic steatosis by facilitating TFEB-mediated lysosome biogenesis and lipophagy. *J Nutr Biochem* 2019; 73: 108214.
13. Cai W, Weng D, Yan P, et al. Genetic polymorphisms associated with nonalcoholic fatty liver disease in Uyghur population: a case-control study and meta-analysis. *Lipids Health Dis* 2019; 18: 14–14.
14. Boccuto L and Abenavoli L. Genetic and epigenetic profile of patients with alcoholic liver disease. *Ann Hepatol* 2017; 16: 490–500.
15. Martina G, Sonia C, Ariel EF, et al. Role of TM6SF2 rs58542926 in the pathogenesis of nonalcoholic pediatric fatty liver disease: a multiethnic study. *Hepatology* 2016; 63: 117–125.
16. Jian GF, Lai W and Hui Z. Guidelines of prevention and treatment of nonalcoholic fatty liver disease (2018, China). *J Dig Dis* 2019; 20: 163–173.
17. Barrett JC, Fry B, Maller J, et al. Haploview: analysis and visualization of LD and haplotype maps. *Bioinformatics* 2005; 21: 263–265.
18. Andersson T, Alfredsson L, Källberg H, et al. Calculating measures of biological interaction. *Eur J Epidemiol* 2005; 20: 575–579.
19. Song TT, Cai RS, Hu R, et al. The important role of TFEB in autophagy-lysosomal pathway and autophagy-related diseases: a systematic review. *Eur Rev Med Pharmacol Sci* 2021; 25: 1641–1649.
20. Danford CJ and Lai M. NAFLD: a multisystem disease that requires a multidisciplinary approach. *Frontline Gastroenterol* 2019; 10: 328–329.
21. Zhang C, Guo L, Qin Y, et al. Correlation between *Helicobacter pylori* infection and polymorphism of adiponectin gene promoter-11391G/A, superoxide dismutase gene in nonalcoholic fatty liver disease. *Zhong Nan Da Xue Xue Bao Yi Xue Ban* 2016; 41: 359–366.
22. Zhang C and Guo L. Correlation of polymorphisms of adiponectin receptor 2 gene+33371Gln/Arg, cytochrome P4502E1 gene Rsa I and smoking with nonalcoholic fatty liver disease. *Nan Fang Yi Ke Da Xue Xue Bao* 2014; 34: 1481–1487.
23. Oniki K, Hori M, Saruwatari J, et al. Interactive effects of smoking and glutathione S-transferase polymorphisms on the development of non-alcoholic fatty liver disease. *Toxicol Lett* 2013; 220: 143–149.

24. Akhavan Rezayat A, Dadgar Moghadam M, Ghasemi Nour M, et al. Association between smoking and non-alcoholic fatty liver disease: a systematic review and meta-analysis. *Sage Open Med* 2018; 6: 205031211774522.
25. Wijarnpreecha K, Scribani M, Kim D, et al. The interaction of nonalcoholic fatty liver disease and smoking on mortality among adults in the United States. *Liver Int* 2019; 39: 1202–1206.
26. Liu P, Xu Y, Tang Y, et al. Independent and joint effects of moderate alcohol consumption and smoking on the risks of non-alcoholic fatty liver disease in elderly Chinese men. *PLoS One* 2017; 12: e0181497.
27. Chavez-Tapia NC, Lizardi-Cervera J, Perez-Bautista O, et al. Smoking is not associated with nonalcoholic fatty liver disease. *World J Gastroenterol* 2006; 12: 5196–5200.
28. Zhao L, Yu C, Lv J, et al. Fluoride exposure, dopamine relative gene polymorphism and intelligence: a cross-sectional study in China. *Ecotoxicol Environ Saf* 2021; 209: 111826.
29. Chen Z, Yu D, Xu ZW, et al. C-reactive protein gene polymorphisms and gene-environment interactions in ischaemic stroke. *Neurol Res* 2015; 37: 979–984.
30. Manuguerra M, Matullo G, Veglia F, et al. Multi-factor dimensionality reduction applied to a large prospective investigation on gene-gene and gene-environment interactions. *Carcinogenesis* 2007; 28: 414–422.
31. Lou X, Chen G, Yan L, et al. A generalized combinatorial approach for detecting gene-by-gene and gene-by-environment interactions with application to nicotine dependence. *Am J Hum Genet* 2007; 80: 1125–1137.
32. Hassani ZS, Mansoori A and Hosseinzadeh M. Relationship between dietary patterns and Non-alcoholic fatty liver disease: a systematic review and meta-analysis. *J Gastroenterol Hepatol* 2020; 36: 1470–1478.
33. Salehi-Sahlabadi A, Sadat S, Beigrezaei S, et al. Dietary patterns and risk of non-alcoholic fatty liver disease. *BMC Gastroenterol* 2021; 21: 41.
34. Liu X, Peng Y, Chen S, et al. An observational study on the association between major dietary patterns and non-alcoholic fatty liver disease in Chinese adolescents. *Medicine (Baltimore)* 2018; 97: e0576.

Abbreviations

FLD	fatty liver disease
TFEB	transcription factor EB
NAFLD	nonalcoholic fatty liver disease
ALD	alcoholic liver disease
CD36	cluster of differentiation 36
FABPs	fatty acid-binding proteins
Crat	carnitine acetyltransferase
PPARα	peroxisome proliferator-activated receptor alpha
PGC-1α	proliferator-activated receptor gamma coactivator 1 alpha
SNPs	single nucleotide polymorphisms
WC	waist circumference
BMI	body mass index
SBP/DBP	systolic/diastolic blood pressure
HbA1C	glycosylated hemoglobin
LDL-C	low-density lipoprotein cholesterol
HDL-C	high-density lipoprotein cholesterol

TG	triglyceride
TC	total cholesterol
GLU	fasting plasma glucose
ALB	albumin
ALT	alanine aminotransferase
AST	aspartate transaminase
UA	uric acid
HWE	Hardy–Weinberg equilibrium
RERI	relative excess risk due to interaction
AP	attributable proportion of interaction
SI	synergy index
LD	linkage disequilibrium
DM	Dominant model
RM	Recessive model
MDR	multifactor dimensionality reduction

Author biographies

Chunbao Mo received Master degree in pathology and pathophysiology from Guilin Medical University in 2021. Now he is a Doctoral candidate in the School of Medicine, Southern University of Science and Technology, Guangdong, China. His research focuses on environmental and molecule epidemiology.

Tingyu Mai is a Postgraduate in the School of Public Health, Guilin Medical University, Guangxi, China. His research focuses on environmental epidemiology.

Jiansheng Cai is a Doctoral candidate in the School of Public Health, Guangxi Medical University, Guangxi, China and works in Guangxi Key Laboratory of Tumor Immunology and Microenvironmental Regulation, Guilin Medical University, Guilin, China. His research focuses on environmental epidemiology.

Haoyu He, Doctor of Medicine, is employed in the Affiliated Hospital of Stomatology, Guangxi Medical University, Guangxi, China. His research focuses on environmental epidemiology.

Huaxiang Lu is a Doctoral candidate in the School of Public Health, Guangxi Medical University, Guangxi, China and works in the Guangxi Zhuang Autonomous Region Center for Diseases Control and Prevention. His research focuses on environmental and infectious diseases epidemiology.

Xu Tang is a Doctoral candidate in the School of Public Health, Guangxi Medical University, Guangxi, China. His research focuses on environmental epidemiology.

Quanhui Chen received Master degree of medicine from Guangxi Medical University in 2020. Now he is employed in the Liuzhou Workers' Hospital, Guangxi, China. His research focuses on epidemiology and hospital infection management.

Xia Xu is a Doctoral candidate in the School of Public Health, Guangxi Medical University, Guangxi, China. Her research focuses on environmental epidemiology.

Chuntao Nong received Master degree in medicine from Guangxi Medical University in 2021. Now he is employed in Nanning Municipal Center for Disease Control and Prevention. His research focuses on environmental epidemiology.

Shuzhen Liu is a Postgraduate in the School of Public Health, Guangxi Medical University, Guangxi, China. Her research focuses on environmental epidemiology.

Dechan Tan is a Postgraduate in the School of Public Health, Guilin Medical University, Guangxi, China. Her research focuses on environmental epidemiology.

Qiumei Liu is a Postgraduate in the School of Public Health, Guangxi Medical University, Guangxi, China. Her research focuses on environmental epidemiology.

Min Xu is a Postgraduate in the School of Public Health, Guangxi Medical University, Guangxi, China. Her research focuses on environmental epidemiology.

You Li is an Associate Professor in the School of Public Health, Guilin Medical University, Guangxi, China. Her research focuses on environmental epidemiology.

Zhiyong Zhang is a Professor in the School of Public Health, Guilin Medical University, Guangxi, China. His research focuses on environmental epidemiology.

Jian Qin is a Professor in the School of Public Health, Guangxi Medical University, Guangxi, China. His research focuses on environmental epidemiology.