Journal of Genetics (2019) 98:29
https://doi.org/10.1007/s12041-019-1071-8

REVIEW ARTICLE

Genetics of nonalcoholic fatty liver disease in Asian populations

ARUN KUMAR[1], SHALIMAR[2], GAGANDEEP KAUR WALIA[3], VIPIN GUPTA[1]* and M. P. SACHDEVA[1]*

[1] *Department of Anthropology, University of Delhi, New Delhi 110 007, India*
[2] *Department of Gastroenterology, All India Institute of Medical Sciences, New Delhi 110 029, India*
[3] *Public Health Foundation of India, Gurugram 122 002, India*
*For correspondence. E-mail: Vipin Gupta, udaiig@gmail.com; M. P. Sachdeva, mpsachdeva@rediffmail.com.

Received 16 July 2018; revised 4 October 2018; accepted 25 October 2018; published online 9 March 2019

Abstract. Nonalcoholic fatty liver disease (NAFLD) is characterized by the accumulation of fat in the liver without any history of chronic alcohol consumption. It encompasses a wide spectrum of diseases that range from simple steatosis to nonalcoholic steatohepatitis. NAFLD is strongly associated with obesity, insulin resistance / type-2 diabetes mellitus and the metabolic syndrome. NAFLD is a complex disorder; environmental and genetic factors interact with NAFLD manifestation and determine its progression. In this review, an attempt was made to provide current information on the genetic variants of NAFLD in Asian populations. Literature search was performed by using PubMed, Medline and Google Scholar database. Candidate gene, validation and genomewide association studies (GWASs) were included in this review. A total of 41 studies fulfilled inclusion criteria of which 12 candidate gene studies exclusively focussed on the *PNPLA3* gene and 17 other studies on other important candidate genes such as *NCAN-CILP2*, *PPARG*, *AGTR1*, *FABP1*, *APOC3* etc. reported significant association with NAFLD. Eight validation studies identified associations of variants on *PNPLA3*, *LYPLAL1*, *TM6SF2*, *ADIPOR2*, *STAT3*, *GCKR*, *SAMM50* etc. with NAFLD. Thus, so far, four GWASs have been conducted in Asian population that reported *PNPLA3*, *SAMM50*, *PARVB* and *GATAD2A* genes which were significantly associated with *NAFLD*. Findings indicate that *PNPLA3*, *APOC3*, PPARG, *NCAN* and *GCKR* genes emerge out to be the important biological markers associated with NAFLD.

Keywords. genetics; nonalcoholic fatty liver disease; obesity; liver enzymes; Asia.

Introduction

Nonalcoholic fatty liver disease (NAFLD) is defined by hepatic fat infiltration (>5% hepatocytes) assessed by liver biopsy in the absence of excessive alcohol intake, viral infection, autoimmune or drug-induced liver disease (Brunt and Tiniakos 2010; Aly and Kleiner 2011). With alarming rise in the epidemic of obesity and metabolic syndrome, burden of NAFLD is increasing, and is expected to be the most common indication of liver transplantation (Kemmer *et al.* 2013). The estimated prevalence of NAFLD in the general population around the world is about 20–30% and 67–75% in the obese population (Angulo 2002; Browning *et al.* 2004). In Asia, 15–30% prevalence of NAFLD has been reported in the general population and over 50% in patients with diabetes and metabolic syndrome (Wong 2013). In Sri Lanka, its prevalence is quite high, 33% in general population (Dassanayake *et al.* 2009). In India, the prevalence of NAFLD is 9–32% in general population, with a higher prevalence among overweight/obese and diabetic patients (Amarapurkar *et al.* 2007).

NAFLD is strongly associated with obesity, insulin resistance (IR)/type-2 diabetes mellitus (T2DM) and metabolic syndrome (Fan *et al.* 2005; Wong *et al.* 2010). In obesity, particularly central obesity is highly predictive of hepatic steatosis and disease progression (Fan *et al.* 2005; Wong *et al.* 2010). In overweight subjects, the prevalence of steatosis is at least two times more frequent than in lean subjects (Bellentani and Tiribelli 2001). In morbid obesity, almost all patients present steatosis and more than one-third have nonalcoholic steatohepatitis (NASH) (Machado *et al.* 2006). The association with T2DM is also very strong, being five to nine times more frequent in patients with NAFLD as compared to the general population (Anstee *et al.* 2013). There is a considerable number of evidence from twins and family based studies that the NAFLD is a heritable disease, with a heritability ranging

from 22 to 38%. In individuals of Hispanic ancestry, the heritability of population-based NAFLD has been found to be 31–38% (Schwimmer *et al.* 2009; Wagenknecht *et al.* 2009). European family based studies (Family Heart Study, The Old Order Amish and The Framingham Heart Study) have reported heritability ranging from 26 to 27% (Speliotes *et al.* 2011). In African ancestry, from the Insulin Resistance and Atherosclerosis Study, Family Heart Study and Genetic Epidemiology of Arteriopathy Study, the heritability was found to be 22–34% (Palmer *et al.* 2013).

NAFLD is a complex phenotypic outcome of interaction between environmental and genetic factors (Romeo *et al.* 2008). Genetic-based studies with respect to NAFLD are limited in Asian populations; only four genome-wide association studies (GWASs) have been conducted so far. Thus, in the present review, an attempt has been made to discuss the current information related to genetic associations of NAFLD in Asian population.

Data source and search strategy

Our search for literature was limited to Asian populations through PubMed, Medline and Google Scholar databases (search terms in PubMed: genetics and nonalcoholic fatty liver disease and Asia). Different search terms were used to select studies; for GWASs (GWAS and NAFLD and Asia), for candidate gene studies (candidate gene studies and NAFLD and Asia), for single-nucleotide polymorphism studies (SNPs and NAFLD and Asia) and for validation studies (validation studies and NAFLD and Asia). Search was restricted to human population and included studies which were published up to 30th September 2018.

Search details: ((‘genetics’[Subheading] OR ‘genetics’ [All Fields] OR ‘genetics’[MeSH Terms]) AND (‘non-alcoholic fatty liver disease’[MeSH Terms] OR (‘non-alcoholic’[All Fields] AND ‘fatty’[All Fields] AND ‘liver’[All Fields] AND ‘disease’[All Fields]) OR ‘non-alcoholic fatty liver disease’[All Fields] OR ‘nafld’[All Fields]) AND (‘asia’[MeSH Terms] OR ‘asia’[All Fields])) AND ‘humans’[MeSH Terms] for genetics and non-alcoholic fatty liver disease and Asia.

Selection criteria

Studies which meet the following criteria were included in the review: full-text articles published in English, underlying NAFLD as the outcome of the study, genetic association studies, human-population studies, studies published on Asian population. Studies were excluded: if not published in English language, studies on other metabolic and infectious disease interaction with NAFLD, insufficient sample size, duplicate and only review studies.

A total of 2940 studies were found after search term ‘genetics and NAFLD’ and these studies were filtered for humans, hence only 1866 studies were left. These 1866 studies were further filtered for Asia using search term ‘Genetics and NAFLD and Asia’ leaving only 108 studies. These 108 studies were checked for review studies, unrelated title, insufficient sample size and unavailability of full text. Finally, only 41 studies were found to be suitable for inclusion in this review.

Genes associated with NAFLD

PNPLA3, also called adiponutrin, with a molecular weight of ~53 kDa, is mainly expressed in intracellular membrane fractions of hepatocytes (Sanyal 2011). *PNPLA3* is found on the surface of lipid droplets in liver and acts as a triglyceride hydrolase (He *et al.* 2010). The overexpression of the *PNPLA3* (I148M) mutant allele in human hepatocyte culture results in lipid accumulation (He *et al.* 2010). GWASs have suggested that a nonsynonymous sequence variation (rs738409) that results in the substitution of isoleucine with methionine at residue 148 (I148M) in the *PNPLA3* gene, contributes to the differences in hepatic lipid content and susceptibility to NAFLD (Rotman *et al.* 2010). The variation in the *PNPLA3* gene also contributes to the ethnic and interindividual differences in hepatic fat content and susceptibility to NAFLD (Romeo *et al.* 2008). For instance, a study found that *PNPLA3* (rs738409 GG genotype) increases the risk of NAFLD in Chinese Han and Uygur populations (Zhang *et al.* 2014) (table 1).

Of 41 studies reviewed in the present study, 12 candidate gene studies had focussed exclusively on *PNPLA3* gene and reported association with NAFLD (odds ratio (OR) ranging from 1.732 to 5.22) (table 1). NAFLD is reported to be affected more by the G allele of *PNPLA3* rs738409 in normal weight and in overweight individuals (Nishioji *et al.* 2015). Further, the G allele of rs738409 *PNPLA3* has also found to be associated with the raised levels of liver enzymes (aspartate aminotransferase (AST) and alanine aminotransferase (ALT)), ferritin, higher fasting insulin, homoeostasis model assessment of insulin resistance (HOMA-IR) and the fibrosis stage in the patients with NAFLD, after adjustment for age, gender and body mass index (BMI) (Hotta *et al.* 2010; Bhatt *et al.* 2013). Among normal weight individuals, *PNPLA3* G/G genotype is reported to be associated with risk for NAFLD (Oniki *et al.* 2015) suggesting its role independent of BMI.

Obesity, high haemoglobin A1c, IR, hyperferritinaemia and the *PNPLA3* G allele are known to be independently associated with NAFLD in nonobese individuals (Wei *et al.* 2015). The role of *PNPLA3* (I148M variant) has also been reported to influence the degree of steatosis in patients with elevated plasma ALT levels (Lin *et al.* 2011; Li *et al.* 2012).

Seventeen candidate gene studies other than *PNPLA3* have reported genetic associations with NAFLD, the ORs range from 1.00 to 3.68 (table 2).

Table 1. Association of *PNPLA3* (rs738409) gene with NAFLD among Asian populations.

Reference	Population (study design)	Sample size	Risk alleles/genotype	Results
Hotta *et al.* (2010)	Japanese population (case–control)	253 cases and 578 controls	G-allele	OR = 1.73, 95% CI: 1.25–2.38, $P = 9.4 \times 10^{-10}$
Lin *et al.* (2011)	Taiwanese Children (population-based)	520 obese children	CG genotype GG genotype	CG, OR = 2.96, 95% CI: 1.57–5.59, $P = 0.0008$ GG, OR = 5.84, 95% CI: 2.59–13.16, $P < 0.0001$
Wang *et al.* (2011)	Normoglycaemic (case–control)	156 cases and 723 controls	G allele	OR = 2.03, 95% CI: 1.23–3.375, $P < 0.05$
Li *et al.* (2012)	Chinese population (case–control)	203 cases and 202 controls	M allele	Frequencies 0.32, 0.54 and 0.87 in mild ($n = 105$), moderate ($n = 83$) and severe ($n = 15$) cases, respectively ($P = 7.6 \times 10^{-8}$)
Bhatt *et al.* (2013)	Asian Indians (case–control)	162 cases and 173 controls	G allele	OR = 1.98, 95% CI: 1.43–2.73, $P = 0.04$
Lin *et al.* (2013)	Taipei, Taiwan (population-based)	781 obese children	CG genotype GG genotype	OR = 1.6, 95% CI: 1.07–2.45, $P = 0.023$ OR = 2.6, 95% CI: 1.50–4.68, $P = 0.001$
Zhang *et al.* (2014)	Uygur and Han from Northwestern China (case–control)	396 cases and 399 controls	G allele	Han: OR = 5.22, 95% CI: 1.94–14.04, $P = 0.001$ Uygur: OR = 4.29, 95% CI: 1.60–11.48, $P = 0.004$
Lee *et al.* (2014)	Korean population (case–control)	155 cases and 184 controls	CG+GG genotype	With NAFLD OR = 2.568, 95% CI: 1.109–5.945, $P = 0.028$ and with advanced liver fibrosis OR =18.573, 95% CI: 2.035–169.526, $P = 0.010$
Nishioji *et al.* (2015)	Japanese population (hospital-based)	824 participants	G allele	In normal weight OR = 12.00, 95% CI: 3.71–38.79, $P = 3.3 \times 10^{-5}$ In overweight OR = 13.40, 95% CI: 2.92–61.36, $P = 0.0008$
Oniki *et al.* (2015)	Japanese (cross sectional, retrospective longitudinal)	740 and 393 Participants	G allele	Normal weight subjects OR = 3.06, 95% CI: 1.11–8.43, $P < 0.05$
Wei *et al.* (2015)	Database of the Hong Kong Government (population-based)	911 community subjects	G allele	Common in non-obese than obese NAFLD patients (78.4% versus 59.8%, $P = 0.001$)
Song *et al.* (2016)	Han Chinese (case–control)	384 cases and 384 controls	GG of rs2896019	OR = 2.17, 95% CI: 1.37–3.45, $P = 0.002$

OR, odds ratio; CI, confidence interval; *P*, probability value; NAFLD, nonalcoholic fatty liver disease.

Candidate gene studies in the Asian region revealed that other loci apart from *PNPLA3* may also have potential associations with NAFLD, such as genetic variations in *ATGR1* (rs3772622, rs3772633, rs2276736, rs3772630 and rs3772627) (Yoneda *et al.* 2009), G/C genotype in *SREBP-2* 1784 (Bhatt *et al.* 2011) and alleles Ala and T in *PPARγ* (Yang *et al.* 2011; Bhatt *et al.* 2013). GG genotype and G carrier (GG+CG) of rs2228314, G>C polymorphism in the *SREBP-2* gene are known to be significantly associated with NAFLD after adjusting for age, sex, BMI and diabetes (Wang *et al.* 2014). After controlling the effects of age, sex, sex-adjusted BMI and *PNPLA3* (rs738409) polymorphism, the *PPARGC1A* (rs8192678) was reported as an independent risk factor for developing NAFLD and

Table 2. Association of other genetic loci with NAFLD in Asian populations.

Reference	Population (study design)	Study sample	Targeted gene/SNPs	Results
Yoneda *et al.* (2009)	Japanese (case–control)	167 Cases and 435 controls	*ATGR1* (12 SNPs)	rs3772622 had strongest association amongst the 12 SNPs (OR = 1.95, 95% CI: 1.49–2.55, $P = 0.0000012$)
Yang *et al.* (2012)	Chinese (case–control)	436 Cases, and 467 controls	*PPARG* Pro12Ala variant	OR = 1.87, 95% CI: 1.13–2.85, $P = 0.009$
Bhatt *et al.* (2011)	North Indians (case–control)	162 Cases, and 173 controls	*SREBP-2* 1784 G/C	Frequency of G allele 0.79 and C allele was 0.21 (P =0.01)
Peng *et al.* (2014)	Chinese (population case–control)	553 Cases and 553 controls	*FABP1* (rs2241883 and rs1545224)	OR =1.32, 95% CI: 1.01–1.71, $P < 0.05$ OR = 1.52, 95% CI: 1.14–2.02, $P < 0.05$
Bhatt *et al.* (2013)	Asian Indians (case–control)	162 Cases and 173 controls	Ala and T of *PPARγ* (Pro12Ala, C161T)	OR =1.64, 95% CI: 1.09–2.45, P =0.05
Zain *et al.* (2013)	Multi-Ethnic (population case–control)	Biopsy proven cases and 198 controls	*AGTR1* (rs377262, rs3772627, rs3772630 etc.)	rs2276736, rs3772630 and rs3772627 was protective against NAFLD (P =0.010, P =0.016 and P =0.026, respectively) in the Indian ethnic group
Lin *et al.* (2013)	Taipei, Taiwan (population based)	781 Obese children	*PPARGC1A* rs8192678	OR =1.740, 95% CI: 1.149, 2.637, P =0.009
Wong *et al.* (2014)	Chinese (population based)	922 Community subjects	*TM6SF2* variant	*TM6SF2* variant was detected in four subjects only (0.4%, 95% CI: 0–0.9%). Two of the four subjects had NAFLD
Jiang *et al.* (2014)	Chinese (population based)	340 Non–diabetic subjects	*FGF21* rs499765	Effect of rs499765 on NAFLD found significant (P =0.045) after adjustment
Wang *et al.* (2014)	Han Chinese (case–control)	300 Cases and 160 controls	*SREBP-2* (rs2228314 G > C)	(CG+GG) of rs2228314 G>C (both $P < 0.001$)
Niu *et al.* (2014)	Han Chinese (case–control)	390 Patients and 409 controls	*APOC3* (rs2854116 and rs2854117)	OR =1.06, 95% CI: 0.72–1.57, $P > 0.05$ OR = 1.00, 95% CI: 0.68–1.48, $P > 0.05$
Peng *et al.* (2014)	Han Chinese (case–control)	580 Cases and 580 controls	*MTTP* (rs180004, rs057613 and rs305335)	rs1800804 (–164 T/C), rs1057613 ($P = 0.014$)
Wang *et al.* (2015b)	Han Chinese (cohort study)	384 Cases and 384 controls	*TM6SF2* rs5854292, rs58542926	($P = 0.0007$), ($P = 0.0004$) after conditioning on rs738409
Wang *et al.* (2015a)	Han Chinese (case–control)	280 Cases and 281 controls	*APPL1* rs4640525	CG+GG were higher in NAFLD compared with control (all $P < 0.05$)
Boonvisut *et al.* (2016)	Asian and Pacific ethnic groups, (replication-analysis)	3013 Japanese, 119 Palauan, 947 Mongolian, 212 Thai, 40 Chinese	*NCAN-CILP2* rs58542926, rs735273 etc.	rs58542926 in Thai ($P = 0.0008$) Allele T OR = 1.682, 95% CI: 1.289–2.196, $P = 0.00013$
Zhang *et al.* (2016)	Chinese Han (case–control)	59 Cases and 72 control	*APOC3* (rs4225, rs4520, rs5128, rs2070666 and rs2070667)	rs2070666 with NAFLD (OR = 3.683, 95% CI: 1.037–13.084) and with histological steatosis (OR = 4.986, 95% CI: 1.020–24.371)
Wu *et al.* (2016)	Han Chinese (case–control)	384 Cases and 384 controls	*PARVB* rs5764455 and rs6006473	($P = 0.034$), ($P = 0.017$)

OR, odds ratio; CI, confidence interval; P, probability value; NAFLD, non–alcoholic fatty liver disease; SNPs, single nucleotide polymorphisms.

increased mean serum levels of ALT concentration (Lin *et al.* 2013). The polymorphism near *MTTP* gene, which encodes a microsomal triglyceride transfer protein, also influences the susceptibility for developing NAFLD (Peng *et al.* 2014).

Genotype AA of rs5764455 in the *PARVB* gene (Wu *et al.* 2016) TNF-α-238, adiponectin-45, leptin -2548, peroxisome proliferator-activated receptors-161, phosphatidyle thanolamine N-methyltransferase-175 (Zhou *et al.* 2010), *FABP1* rs2241883 C allele, rs1545224 A allele and rs2241883 C variants (Peng *et al.* 2012), GG of rs11235972 in *UCP3* (Xu *et al.* 2013), CG + GG genotype of rs4640525 in *APPL1* gene (Wang *et al.* 2015b) have been reported to be significantly associated with

Table 3. GWASs on NAFLD in Asian populations.

Reference	Study design/sample size	Population	Inclusion method for cases	Results
Kawaguchi *et al.* (2012)	Case–control: 543 cases and 942 controls	Japanese population	Histologically confirmed NAFLD cases	NAFLD: *PNPLA3* rs738409 (OR = 1.66, 95% CI: 1.43–1.94, $P = 1.4 \times 10^{-10}$) NASH: *PNPLA3* rs738409 (OR = 1.96, 95% CI: 1.47–2.62, $P = 4.8 \times 10^{-6}$)
Kitamoto *et al.* (2013)	Case–control: 392 cases and 934 controls Replication sample: 344 cases and 1012 controls	Japanese population	Biopsy, CT, MRI proven cases	*PNPLA3:* rs738409 (OR = 2.05, 95% CI: 1.70–2.47, $P = 6.8 \times 10^{-14}$) After adjusted with rs738409; *PNPLA3* (rs2896019, and rs381062) *SAMM50* (rs738491, rs3761472 and rs2143571) *PARVB* (rs6006473, rs5764455 and rs6006611) [OR range=1.84–2.02, $P < 2.0 \times 10^{-10}$]
Chatterjee *et al.* (2015)	Population based: 20 ethnic groups (376 samples)	Indian population	Additive risk scores of NAFLD for three Hapmap populations	Allelic and genotypic OR of eight loci OR range = 1.16–6.69
Kawaguchi *et al.* (2018)	Case–control 902 case and 7672 controls	Japanese population	Histologically confirmed NAFLD cases	NAFLD: *PNPLA3* rs2896019 (OR = 1.85, 95% CI: 1.67–2.05, $P = 2.3 \times 10^{-31}$) NAFLD: *GCKR* rs1260326 (OR = 1.38, 95% CI: 1.25–1.53, $P = 9.6 \times 10^{-10}$) NAFLD: *GATAD2A* rs4808199 (OR = 1.37, 95% CI: 1.23–1.53, $P = 2.3 \times 10^{-8}$) NASH-HCC: *DYSF* rs17007417 (OR = 2.74, 95% CI: 1.84–4.06, $P = 5.2 \times 10^{-7}$)

CT, computed tomography; MRI, magnetic resonance imaging.

susceptibility to NAFLD, suggesting their role in developing the disease. *TM6SF2* gene was significantly increasing the susceptibility of NAFLD independent of the *PNPLA3* and *NCAN* gene (Wang *et al.* 2015a). Moreover, higher levels of fasting insulin, HOMA scores, triglycerides (Bhatt *et al.* 2013), total cholesterol, ALT (Jiang *et al.* 2014; Wu *et al.* 2016), AST (Wu *et al.* 2016) and hs-CRP in NAFLD subjects indicate that these biochemical parameters also have a considerable role in disease progression. Oxidative stress and smoking could also be important risk factors for NAFLD as a study by Yang *et al.* (2011) found significant association of these factors with NAFLD.

GWAS of NAFLD

Four GWASs have been conducted among Asian populations (Kawaguchi *et al.* 2012, 2018; Kitamoto *et al.* 2013; Chatterjee *et al.* 2015) of which three were on Japanese populations (case–control design) and one study was conducted on an Indian population (using population-based design) (table 3).

The GWAS conducted on Japanese population had observed a strong association of *PNPLA3* (rs738409) with the histological classifications of NAFLD ($P = 1.4 \times 10^{-10}$) (Kawaguchi *et al.* 2012). In addition, study has also found marked differences in rs738409 genotype distributions between type-4 subgroup corresponding to NASH and the other three subgroups (OR = 1.96, 95% confidence interval (CI): 1.47–2.62, $P = 4.86 \times 10^{-6}$) (Kawaguchi *et al.* 2012). A subgroup analysis of NAFLD patients against controls showed a significant association of rs738409 with the NAFLD type (OR = 2.18, 95% CI: 1.81–2.63, $P = 1.7 \times 10^{-16}$). The association of *PNPLA3* (rs738409) was also reported with high levels of hyaluronic acid, HbA1c and iron deposition in NAFLD patients (table 3).

The second GWAS related to NAFLD was conducted using the Japanese SNP database on 392 cases and 934 control samples (Kitamoto *et al.* 2013). Genome scan data from control-1 ($n = 934$) were genotyped using the Illumina Human-Hap550 Bead Chip. The rs738409 variant in the *PNPLA3* gene was found to be most strongly associated with NAFLD after adjustments, and other variants such as rs2896019 and rs381062 in the *PNPLA3* gene, rs738491, rs3761472 and rs2143571 in the *SAMM50* gene, rs6006473, rs5764455 and rs6006611 in the *PARVB* gene also showed significant associations with NAFLD. These SNPs were also found to be significantly associated with

Table 4. Studies on validation of GWAS findings with NAFLD in Asian populations.

Reference	Study design/(population)	Sample size	Targeted genes	Results
Yang *et al.* (2011)	Population based (Chinese)	436 Cases and 467 controls	*GCKR*	rs780094 (OR = 1.607, 95% CI: 1.139–2.271, $P = 7.2 \times 10^{-3}$)
Xie *et al.* (2013)	Population based (Chinese)	1440 Men	*PNPLA3*, *TNFα*, *TNFRSFB*, *IL6Rα*, *UGT1A1*, *SOD2*, *ABCG2*, *SLC17A3* and *SLC2A9*	rs887829 in *UGT1A1* and rs4880 in *SOD2* (OR combined = 2.81, 95% CI: 1.66–4.76, $P < 0.001$)
Kitamoto *et al.* (2014)	Case–control (Japanese)	540 Cases and 1012 controls	*LYPLAL1*, *ZP4*, *GCKR*, *HSD17B13*, *PALLD*, *PPP1R3*, *FDFT1* and *NCAN*	*GCKR*: rs780094 (OR = 1.37, 95% CI: 1.12–1.68, $P = 0.0024$) *TRIB1*: rs2954021 (OR = 1.53, 95% CI: 1.25–1.88, $P = 4.5 \times 10^{-5}$)
Lin *et al.* (2014)	Population based (Taiwanese Han Chinese)	797 Obese children	*PNPLA3*, *NCAN*, *LYPLAL1*, *GCKR*, and *PPP1R3B*	*GCKR*: rs780094 (OR = 1.99, 95% CI: 1.196, 3.335; $P = 0.008$)
Shang *et al.* (2015)	Case–control (Chinese)	162 Cases and 865 controls	*GCKR*, *PDGFA*, *FDFT1*, *COL13A1*, *NCAN* and *PNPLA3*	*PNPLA3* with NAFLD (OR = 1.55, 95% CI: 1.13–2.11, $P = 0.006$), and steatosis (OR = 3.77, 95% CI: 1.78–7.98, $P = 0.001$)
Chen *et al.* (2015)	Case–control (Chinese Han population)	340 Cases and 452 controls	*SAMM50*	rs738491 (OR = 1.507, 95% CI: 1.035–2.195, $P = 0.032$) rs2143571 (OR = 1.761, 95% CI: 1.232–2.517, $P = 0.002$) rs3761472 (OR = 1.483, 95% CI: 1.039–2.115, $P = 0.030$)
Kasturiratne *et al.* (2015)	Case–control (Sri Lankan)	946 Cases and 1213 controls	*PNPLA3*, *LYPLAL1*, *GCKR*, *PPP1R3B*, *NCAN*, *APOC3*, *ADIPOR2* and *STAT3*	*PNPLA3*: rs738409 (OR = 1.25, 95% CI: 1.08–1.44, $P = 0.003$) *STAT3*: rs9891119 (OR = 1.15, 95% CI: 1.02–1.3, $P = 0.028$) *APOC3*: (rs2854117, rs2854116) (OR =0.86, 95% CI: 0.77–1, $P = 0.019$)
Wang *et al.* (2016)	Community based case–control (Han Chinese)	384 Cases and 384 controls	*PNPLA3*, *NCAN*, *GCKR*, *LYPLAL1* and *TM6SF2*	*PNPLA3*: rs738409 (OR = 1.52, 95% CI: 1.19 –1.96, $P = 0.00087$) *TM6SF2*: rs58542926 (OR = 2.11, 95% CI: 1.34–3.39, $P = 0.0016$)

decreased serum triglycerides and increased AST and ALT in NAFLD patients (table 3).

Recently, a GWAS was conducted by Kawaguchi *et al.* (2018) in Japanese populations ($n = 902$: cases –476 NASH and 58 NASH-HCC patients; and controls –7672 controls). This study confirms the significant association of *PNPLA3* (rs2896019), *GCKR* (rs1260326) and *GATAD2A* (rs4808199) with NAFLD (table 3). Moreover, they have also identified a new association of rs17007417 in *DYSF* with NASH-HCC.

Another genomewide study on 376 individuals comprising 20 Indian ethnic groups (table 3) studied the risk allele frequencies of 34 SNPs, phylogenetic relationship of study populations, allelic and genotypic risk scores of eight disease-associated loci and an empirical null distribution of risk (Chatterjee *et al.* 2015). This study has reported associations of 29 of 34 loci using a pair-wise binomial proportion test, and found that the caste populations showed higher risk scores for NAFLD than Indo-European tribal populations. Further, Tibeto-Burman speaking populations showed additive risk scores similar to Caucasians, but higher than Chinese, Japanese and African populations. This Indian study suggested that tribal populations have significantly lower predicted risk score of NAFLD than caste populations suggesting that Indian caste populations have a high genetic predisposition of developing

NAFLD as compared to tribal populations (Chatterjee *et al.* 2015).

Genetic validation studies of NAFLD

The validations of GWAS findings have been examined for 26 genes from eight studies conducted in Asian populations (predominantly in the Chinese population in six studies) (table 4). These studies validated the genetic associations of *PNPLA3* (Kasturiratne *et al.* 2015; Shang *et al.* 2015; Wang *et al.* 2016), *GCKR* (Yang *et al.* 2011; Kitamoto *et al.* 2014; Lin *et al.* 2014), *TRIB1* (Kitamoto *et al.* 2014), *STAT3* and *APOC3* (Kasturiratne *et al.* 2015), *SAMM50* (Chen *et al.* 2015) and *TM6SF2* (Wang *et al.* 2016) with the NAFLD (table 4). The variants associated with NAFLD were also associated with quantitative traits related to NAFLD, for instance, rs780094 was associated with AST levels (Lin *et al.* 2014), rs2143571 and rs3761472 were associated with increased triglycerides, AST and ALT levels (Chen *et al.* 2015), *PNPLA3* (rs738409) with a lower serum triglyceride level and *PPP1R3B* with HOMA-IR (Kasturiratne *et al.* 2015). Elevated serum uric acid is independently associated with NAFLD regardless of IR and with metabolic disorders, especially central obesity (Xie *et al.* 2013). The genetic variants of *ABCG2* and *TNFRSF1B* were reported to be associated with serum uric acid in NAFLD cases (Xie *et al.* 2013).

Limitations of the studies

The diagnosis of NAFLD was primarily based on ultrasonography in most of the studies (Bhatt *et al.* 2011; Peng *et al.* 2012; Lee *et al.* 2014; Oniki *et al.* 2015; Nishioji *et al.* 2015; Song *et al.* 2016), as it is less invasive, and not on liver biopsy which is considered as a gold standard (Hernaez *et al.* 2011). Limited sample sizes (ranging from 59 to 162) and low statistical power is common in Asian studies on genetics of NAFLD (Bhatt *et al.* 2011; Lee *et al.* 2014; Shang *et al.* 2015; Zhang *et al.* 2016). Further, selection of controls from the hospital-based settings instead of general population in case–control design leads to biased findings confounded by population structure (Li *et al.* 2012; Chan *et al.* 2013; Xu *et al.* 2013; Pan *et al.* 2015).

In conclusion, The *PNPLA3* gene is a major locus replicated and validated with NAFLD, NASH as well as with elevated liver enzymes (ALT, AST and GGT) in Asian populations. There is a need for conducting large genomewide or whole genome association studies to fill the research gap in this region. Apart from *PNPLA3*, other noticeable loci are *APOC3*, *PPARG*, *NCAN* and *GCKR* that emerge out to be important biological markers associated with NAFLD in Asian populations.

Acknowledgement

The authors would like to acknowledge the University of Delhi's Research and Development Grant for funding this research work.

References

Aly F. Z. and Kleiner D. E. 2011 Update on fatty liver disease and steatohepatitis. *Adv. Anat. Pathol.* **18**, 294–300.

Amarapurkar D., Kamani P., Patel N., Gupte P., Kumar P., Agal S. *et al.* 2007 Prevalence of non-alcoholic fatty liver disease: population based study. *Ann. Hepatol.* **6**, 161–163.

Angulo P. 2002 Treatment of nonalcoholic fatty liver disease. *Ann. Hepatol.* **1**, 12–19.

Anstee Q. M., Targher G. and Day C. P. 2013 Progression of NAFLD to diabetes mellitus, cardiovascular disease or cirrhosis. *Nat. Rev. Gastroenterol. Hepatol.* **10**, 330–344.

Bellentani S. and Tiribelli C. 2001 The spectrum of liver disease in the general population: lesson from the Dionysos study. *J. Hepatol.* **35**, 531–537.

Bhatt S. P., Nigam P., Misra A., Guleria R., Luthra K., Vaidya M. *et al.* 2011 SREBP-2 1784 G/C genotype is associated with non-alcoholic fatty liver disease in north Indians. *Dis. Markers* **31**, 371–377.

Bhatt S. P., Nigam P., Misra A., Guleria R., Luthra K., Pandey R. M. *et al.* 2013 Association of peroxisome proliferator activated receptor-γ gene with non-alcoholic fatty liver disease in Asian Indians residing in north India. *Gene* **512**, 143–147.

Boonvisut S., Nakayama K., Makishima S., Watanabe K., Miyashita H., Lkhagvasuren M. *et al.* 2016 Replication analysis of genetic association of the *NCAN-CILP2* region with plasma lipid levels and non-alcoholic fatty liver disease in Asian and Pacific ethnic groups. *Lipids Health Dis.* **15**, 8.

Browning J. D., Szczepaniak L. S., Dobbins R., Nuremberg P., Horton J. D., Cohen J. C. *et al.* 2004 Prevalence of hepatic steatosis in an urban population in the United States: impact of ethnicity. *Hepatology* **40**, 1387–1395.

Brunt E. M. and Tiniakos D. G. 2010 Histopathology of nonalcoholic fatty liver disease. *World J. Gastroenterol.* **16**, 5286–5296.

Chan W. K., Tan A. T., Vethakkan S. R., Tah P. C., Vijayanathan A. and Goh K. L. 2013 Non-alcoholic fatty liver disease in diabetics–prevalence and predictive factors in a multiracial hospital clinic population in Malaysia. *J. Gastroenterol. Hepatol.* **28**, 1375–1383.

Chatterjee A., Basu A., Chowdhury A., Das K., Sarkar-Roy N., Majumder P. P. *et al.* 2015 Comparative analyses of genetic risk prediction methods reveal extreme diversity of genetic predisposition to nonalcoholic fatty liver disease (NAFLD) among ethnic populations of India. *J. Genet.* **94**, 105–113.

Chen L., Lin Z., Jiang M., Lu L., Zhang H., Xin Y. *et al.* 2015 Genetic variants in the SAMM50 gene create susceptibility to nonalcoholic fatty liver disease in a Chinese Han population. *Hepat. Mon.* **15**, e31076.

Dassanayake A. S., Kasturiratne A., Rajindrajith S., Kalubowila U., Chakrawarthi S., De Silva A. P. *et al.* 2009 Prevalence and risk factors for non-alcoholic fatty liver disease among adults in an urban Sri Lankan population. *J. Gastroenterol. Hepatol.* **24**, 1284–1288.

Fan J. G., Zhu J., Li X. J., Chen L., Lu Y. S., Li L. *et al.* 2005 Fatty liver and the metabolic syndrome among Shanghai adults. *J. Gastroenterol. Hepatol.* **20**, 1825–1832.

He S., McPhaul C., Li J. Z., Garuti R., Kinch L., Grishin N. V. *et al.* 2010 A sequence variation (I148M) in *PNPLA3* associated with nonalcoholic fatty liver disease disrupts triglyceride hydrolysis. *J. Biol. Chem.* **285**, 6706–6715.

Hernaez R., Lazo M., Bonekamp S., Kamel I., Brancati F. L., Guallar E. *et al.* 2011 Diagnostic accuracy and reliability of ultrasonography for the detection of fatty liver: a meta-analysis. *Hepatology* **54**, 1082–1090.

Hotta K., Yoneda M., Hyogo H., Ochi H., Mizusawa S., Ueno T. *et al.* 2010 Association of the rs738409 polymorphism in *PNPLA3* with liver damage and the development of nonalcoholic fatty liver disease. *BMC Med. Genet.* **11**, 172.

Jiang S., Zhang R., Li H., Fang Q., Jiang F., Hou X. *et al*. 2014 The single nucleotide polymorphism rs499765 is associated with fibroblast growth factor 21 and nonalcoholic fatty liver disease in a Chinese population with normal glucose tolerance. *J. Nutrigenet. Nutrigenomics* **7**, 121–129.

Kasturiratne A., Akiyama K., Niriella M. A., Takeuchi F., Isono M., Dassanayake A. S. *et al.* 2015 Association of genetic variants with non-alcoholic fatty liver disease in an urban Sri Lankan community. *Liver Int.* **35**, 676–679.

Kawaguchi T., Sumida Y., Umemura A., Matsuo K., Takahashi M., Takamura T. *et al.* 2012 Genetic polymorphisms of the human *PNPLA3* gene are strongly associated with severity of non-alcoholic fatty liver disease in Japanese. *PLoS One* **7**, e38322.

Kawaguchi T, Shima T, Mizuno M, Mitsumoto Y, Umemura A, Kanbara Y. *et al.* 2018 Risk estimation model for non-alcoholic fatty liver disease in the Japanese using multiple genetic markers. *PLoS One* **13**, e0185490.

Kemmer N., Neff G. W., Franco E., Osman-Mohammed H., Leone J., Parkinson E. *et al.* 2013 Nonalcoholic fatty liver disease epidemic and its implications for liver transplantation. *Transplantation* **96**, 860–862.

Kitamoto T., Kitamoto A., Yoneda M., Hyogo H., Ochi H., Nakamura T. *et al.* 2013 Genome-wide scan revealed that polymorphisms in the *PNPLA3*, *SAMM50*, and *PARVB* genes are associated with development and progression of nonalcoholic fatty liver disease in Japan. *Hum. Genet.* **132**, 783–792.

Kitamoto A., Kitamoto T., Nakamura T., Ogawa Y., Yoneda M., Hyogo H. *et al.* 2014 Association of polymorphisms in GCKR and TRIB1 with nonalcoholic fatty liver disease and metabolic syndrome traits. *Endocr. J.* **61**, 683–689.

Lee S. S., Byoun Y. S., Jeong S. H., Woo B. H., Jang E. S., Kim J. W. *et al.* 2014 Role of the *PNPLA3* I148M polymorphism in nonalcoholic fatty liver disease and fibrosis in Korea. *Dig. Dis. Sci.* **59**, 2967–2974.

Li Y., Xing C., Tian Z. and Ku H. C. 2012 Genetic variant I148M in *PNPLA3* is associated with the ultrasonography-determined steatosis degree in a Chinese population. *BMC Med. Genet.* **13**, 113.

Lin Y. C., Chang P. F., Hu F. C., Yang W. S., Chang M. H. and Ni Y. H. 2011 A common variant in the *PNPLA3* gene is a risk factor for non-alcoholic fatty liver disease in obese Taiwanese children. *J. Pediatr.* **158**, 740–744.

Lin Y. C., Chang P. F., Chang M. H. and Ni Y. H. 2013 A common variant in the peroxisome proliferator-activated receptor-γ coactivator-1α gene is associated with nonalcoholic fatty liver disease in obese children. *Am. J. Clin. Nutr.* **97**, 326–331.

Lin Y. C., Chang P. F., Chang M. H. and Ni Y. H. 2014 Genetic variants in GCKR and *PNPLA3* confer susceptibility to non-alcoholic fatty liver disease in obese individuals. *Am. J. Clin. Nutr.* **99**, 869–874.

Machado M., Marques-Vidal P. and Cortez-Pinto H. 2006 Hepatic histology in obese patients undergoing bariatric surgery. *J. Hepatol.* **45**, 600–606.

Nishioji K., Mochizuki N., Kobayashi M., Kamaguchi M., Sumida Y., Nishimura T. *et al*. 2015 The impact of PNPLA3 rs738409 genetic polymorphism and weight gain $\geq$10 kg after age 20 on non-alcoholic fatty liver disease in non-obese Japanese individuals. *PLoS One* **20**, e0140427.

Niu T. H., Jiang M., Xin Y. N., Jiang X. J., Lin Z. H. and Xuan S. Y. 2014 Lack of association between apolipoprotein C3 gene polymorphisms and risk of nonalcoholic fatty liver disease in a Chinese Han population. *World J. Gastroenterol.* **20**, 3655–3662.

Oniki K., Saruwatari J., Izuka T., Kajiwara A., Morita K., Sakata M. *et al.* 2015 Influence of the *PNPLA3* rs738409 polymorphism on non-alcoholic fatty liver disease and renal function among normal weight subjects. *PLoS One* **10**, e0132640.

Pan Q., Zhang R. N., Wang Y. Q., Zheng R. D., Mi Y. Q., Liu W. B. *et al*. 2015 Linked PNPLA3 polymorphisms confer susceptibility to nonalcoholic steatohepatitis and decreased viral load in chronic hepatitis B. *World J. Gastroenterol.* **21**, 8605–8614.

Palmer N. D., Musani S. K., Yerges-Armstrong L. M., Feitosa M. F., Bielak L. F., Hernaez R. *et al.* 2013 Characterization of European ancestry nonalcoholic fatty liver disease-associated variants in individuals of African and Hispanic descent. *Hepatology* **58**, 966–975.

Peng X. E., Wu Y. L., Lu Q. Q., Hu Z. J. and Lin X. 2012 Two genetic variants in *FABP1* and susceptibility to non-alcohol fatty liver disease in a Chinese population. *Gene* **500**, 54–58.

Peng X. E., Wu Y. L., Lu Q. Q., Hu Z. J. and Lin X. 2014 MTTP polymorphisms and susceptibility to non-alcoholic fatty liver disease in a Han Chinese population. *Liver Int.* **34**, 118–128.

Romeo S., Kozlitina J., Xing C., Pertsemlidis A., Cox D., Pennacchio L. A. *et al.* 2008 Genetic variation in *PNPLA3* confers susceptibility to nonalcoholic fatty liver disease. *Nat. Genet.* **40**, 1461–1465.

Rotman Y., Koh C., Zmuda J. M., Kleiner D. E., Liang T. J. and NASH CRN. 2010 The association of genetic variability in patatin-like phospholipase domain containing protein 3 (*PNPLA3*) with histological severity of nonalcoholic fatty liver disease. *Hepatology* **52**, 894–903.

Sanyal A. J. 2011 NASH: a global health problem. *Hepatol. Res.* **41**, 670–674.

Schwimmer J. B., Celedon M. A., Lavine J. E., Salem R., Campbell N., Schork N. J. *et al.* 2009 Heritability of nonalcoholic fatty liver disease. *Gastroenterology* **136**, 1585–1592.

Shang X. R., Song J. Y., Liu F. H., Ma J. and Wang H. J. 2015 GWAS-identified common variants with nonalcoholic fatty liver disease in Chinese children. *J. Pediatr. Gastroenterol. Nutr.* **60**, 669–674.

Song G., Xiao C., Wang K., Wang Y., Chen J., Yu Y. *et al.* 2016 Association of patatin-like phospholipase domain-containing protein 3 gene polymorphisms with susceptibility of nonalcoholic fatty liver disease in a Han Chinese population. *Medicine (Baltimore)* **95**, e4569.

Speliotes E. K., Yerges-Armstrong L. M., Wu J., Hernaez R., Kim L. J., Palmer C. D. *et al.* 2011 Genome-wide association analysis identifies variants associated with nonalcoholic fatty liver disease that have distinct effects on metabolic traits. *PLoS Genet.* **7**, e1001324.

Wagenknecht L. E., Scherzinger A. L., Stamm E. R., Hanley A. J., Norris J. M., Chen Y. D. *et al.* 2009 Correlates and heritability of nonalcoholic fatty liver disease in a minority cohort. *Obesity (Silver Spring)* **17**, 1240–1246.

Wang C. W., Lin H. Y., Shin S. J., Yu M. L., Lin Z. Y., Dai C. Y. *et al.* 2011 The *PNPLA3* I148M polymorphism is associated with insulin resistance and nonalcoholic fatty liver disease in a normoglycaemic population. *Liver Int.* **31**, 1326–1331.

Wang Y., Tong J., Chang B., Wang B. F., Zhang D. and Wang B. Y. 2014 Relationship of SREBP-2 rs2228314 G>C polymorphism with nonalcoholic fatty liver disease in a Han Chinese population. *Genet. Test Mol. Biomarkers* **18**, 653–657.

Wang X., Liu Z., Peng Z. and Liu W. 2015a The TM6SF2 rs58542926 T allele is significantly associated with non-alcoholic fatty liver disease in Chinese. *J. Hepatol.* **62**, 1438–1439.

Wang B., Wang B., Wang Y., Wen B., Liu S., Sang L. *et al.* 2015b Association of APPL1 gene polymorphism with non-alcoholic fatty liver disease susceptibility in a Chinese Han population. *Clin. Lab.* **61**, 1659–1666.

Wang X., Liu Z., Wang K., Wang Z., Sun X., Zhong L. *et al.* 2016 Additive effects of the risk alleles of *PNPLA3* and TM6SF2 on non-alcoholic fatty liver disease (NAFLD) in a Chinese population. *Front Genet.* **7**, 140.

Wei J. L., Leung J. C., Loong T. C., Wong G. L., Yeung D. K., Chan R. S. *et al.* 2015 Prevalence and severity of non-alcoholic fatty liver disease in non-obese patients: a population study using proton-magnetic resonance spectroscopy. *Am. J. Gastroenterol.* **110**, 1306–1314.

Wong V. W. 2013 Nonalcoholic fatty liver disease in Asia: a story of growth. *J. Gastroenterol. Hepatol.* **28**, 18–23.

Wong V. W., Wong G. L., Choi P. C., Chan A. W., Li M. K., Chan H. Y. *et al.* 2010 Disease progression of non-alcoholic fatty liver disease: a prospective study with paired liver biopsies at 3 years. *Gut* **59**, 969–974.

Wong V. W., Wong G. L., Tse C. H. and Chan H. L. 2014 Prevalence of the TM6SF2 variant and non-alcoholic fatty liver disease in Chinese. *J. Hepatol.* **61**, 708–709.

Wu G., Wang K., Xue Y., Song G., Wang Y., Sun X. *et al.* 2016 Association of rs5764455 and rs6006473 polymorphisms in *PARVB* with liver damage of nonalcoholic fatty liver disease in Han Chinese population. *Gene* **575**, 270–275.

Xu Y. P., Liang L., Wang C. L., Fu J. F., Liu P. N., Lv L. Q. *et al.* 2013 Association between UCP3 gene polymorphisms and nonalcoholic fatty liver disease in Chinese children. *World J. Gastroenterol.* **19**, 5897–5903.

Xie Y., Wang M., Zhang Y., Zhang S., Tan A., Gao Y. *et al.* 2013 Serum uric acid and non-alcoholic fatty liver disease in non-diabetic Chinese men. *PLoS One* **8**, e67152.

Yang Z., Wen J., Tao X., Lu B., Du Y., Wang M. *et al.* 2011 Genetic variation in the GCKR gene is associated with non-alcoholic fatty liver disease in Chinese people. *Mol. Biol. Rep.* **38**, 1145–1150.

Yang Z., Wen J., Li Q., Tao X., Ye Z., He M. *et al.* 2012 *PPARG* gene Pro12Ala variant contributes to the development of non-alcoholic fatty liver in middle-aged and older Chinese population. *Mol. Cell. Endocrinol.* **348**, 255–259.

Yoneda M., Hotta K., Nozaki Y., Endo H., Uchiyama T., Mawatari H. *et al.* 2009 Association between angiotensin II type 1 receptor polymorphisms and the occurrence of nonalcoholic fatty liver disease. *Liver Int.* **29**, 1078–1085.

Zain S. M., Mohamed Z., Mahadeva S., Rampal S., Basu R. C., Cheah P. L. *et al.* 2013 Susceptibility and gene interaction study of the angiotensin II type 1 receptor (*AGTR1*) gene polymorphisms with non-alcoholic fatty liver disease in a multi-ethnic population. *PLoS One* **8**, e58538.

Zhang Y., Cai W., Song J., Miao L., Zhang B., Xu Q. *et al.* 2014 Association between the *PNPLA3* I148M polymorphism and non-alcoholic fatty liver disease in the Uygur and Han ethnic groups of northwestern China. *PLoS One* **9**, e108381.

Zhang Y., Cai W., Song J., Miao L., Zhang B., Xu Q. *et al.* 2015 The impact of *PNPLA3* rs738409 genetic polymorphism and weight gain $\geq$10 kg after age 20 on non-alcoholic fatty liver disease in non-obese Japanese individuals. *PLoS One* **10**, e0140427.

Zhang R. N., Zheng R. D., Mi Y. Q., Zhou D., Shen F., Chen G. Y. *et al.* 2016 *APOC3* rs2070666 is associated with the Hepatic Steatosis independently of *PNPLA3* rs738409 in Chinese Han patients with nonalcoholic fatty liver diseases. *Dig. Dis. Sci.* **61**, 2284–2293.

Zhou Y. J., Li Y. Y., Nie Y. Q., Yang H., Zhan Q., Huang J. *et al.* 2010 Influence of polygenetic polymorphisms on the susceptibility to non-alcoholic fatty liver disease of Chinese people. *J. Gastroenterol. Hepatol.* **25**, 772–777.

Corresponding editor: H. A. Ranganath