

Review

Open Access



Genetic risk factors associated with NAFLD

Dong Yun Kim^{1,2}, Jun Yong Park^{3,4}

¹Department of Medicine, Yonsei University Graduate School of Medicine, Seoul 03722, Korea.

²Department of Medicine, Physician-Scientist Program, Yonsei University Graduate School of Medicine, Seoul 03722, Korea.

³Division of Gastroenterology, Department of Internal Medicine, Yonsei University College of Medicine, Seoul 03722, Korea.

⁴Yonsei Liver Center, Severance Hospital, Seoul 03722, Korea.

Correspondence to: Dr. Jun Yong Park, Division of Gastroenterology, Department of Internal Medicine, Yonsei University College of Medicine, 50-1 Yonsei-ro, Seodaemun-gu, Seoul 03722, Korea. E-mail: drpjy@yuhs.ac

How to cite this article: Kim DY, Park JY. Genetic risk factors associated with NAFLD. *Hepatoma Res* 2020;6:85. <http://dx.doi.org/10.20517/2394-5079.2020.96>

Received: 30 Aug 2020 **First Decision:** 10 Oct 2020 **Revised:** 7 Nov 2020 **Accepted:** 19 Nov 2020 **Published:** 5 Dec 2020

Academic Editor: Stefano Bellentani **Copy Editor:** Cai-Hong Wang **Production Editor:** Jing Yu

Abstract

Non-alcoholic fatty liver disease (NAFLD) is estimated to affect 25% of the worldwide population, and is the leading cause of chronic liver disease in developed countries. Genetic research on NAFLD has included heritability studies, candidate gene studies, familial aggregation studies, and genome-wide association studies (GWAS). Next-generation sequencing approaches, such as whole-genome sequencing and whole-exon sequencing, are emerging as the post-GWAS era of genetic research. However, GWAS remains more practical for elucidating the genetic factors related to NAFLD, which is affected by thousands of common genetic variants and does not follow Mendelian inheritance. In the present review, we summarize the current knowledge regarding five GWAS-identified genetic loci that are associated with NAFLD. We also discuss the relationships between NAFLD-predisposing polymorphisms and cardiovascular disease, and potential applications for these identified genetic loci.

Keywords: Genome-wide association study, non-alcoholic fatty liver disease, PNPLA3, TM6SF2, GCKR, MBOAT7, HSD17B13

INTRODUCTION

Non-alcoholic fatty liver disease (NAFLD) is estimated to affect a quarter of the global population, and is the leading cause of chronic liver diseases in developed countries^[1]. NAFLD etiologies are complex and the factors driving NAFLD progression are not completely understood, although they likely include



© The Author(s) 2020. **Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, sharing, adaptation, distribution and reproduction in any medium or format, for any purpose, even commercially, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.



Table 1. List of five established variants associated with NAFLD

Region ¹	Variant	Reported gene	Effect of the variant	MAF ²				NAFLD spectrum			
				European	Latino	East Asian	African	Simple Steatosis	NASH	Fibrosis	HCC
chr22:43928847	rs738409	<i>PNPLA3</i>	I148M	0.2281	0.5493	0.3816	0.1357	+ [11,22,25,26]	+ [22]	+ [22]	+ [23]
chr19:19268740	rs58542926	<i>TM6SF2</i>	E167K	0.07387	0.03248	0.06969	0.03248	+ [27]	+ [31,32]	+ [30]	+ [30]
chr19:54173068	rs641738	<i>MBOAT7</i>	Linked to 3'-UTR	0.4322	0.3300	0.2382	0.3338	+ [33,35]	+ [35]	+ [35]	+ [33]
chr2:27508073	rs1260326	<i>GCKR</i>	P446L	0.5914	0.6666	0.5099	0.8670	+ [8,36,38]	+ [38]	+ [37]	
chr4:87310241	rs72613567	<i>HSD17B13</i>	Alternate splicing P260S	0.2658	0.09614	0.3266	0.06386	+ [45]	+ [44,45]	+ [45,47]	

¹The region of variant was annotated using the UCSC Genome Browser (GRCh38); ²MAF data was from the genome aggregation database browser (gnomAD, <https://gnomad.broadinstitute.org/>). MAF: minor allele frequency; NAFLD: non-alcoholic fatty liver disease; NASH: non-alcoholic steatohepatitis; HCC: hepatocellular carcinoma

environmental factors (e.g., diet), insulin resistance, increased visceral adiposity, and genetics. Genetic research on NAFLD has included heritability studies^[1,2], candidate gene studies^[3], familial aggregation studies^[4,5], and genome-wide association studies (GWAS)^[6-8]. NAFLD heritability was initially evaluated through a candidate gene study. Such studies are designed to examine the association of a phenotype with single-nucleotide polymorphisms (SNPs) in selected genes; however, they have weak statistical power^[9].

Following candidate gene studies, GWAS have become the default methodology for testing the associations between diseases (phenotypes of interest) and millions of SNPs throughout the genome. Over the recent years, GWAS have dramatically improved our understanding of the genetic factors related to NAFLD susceptibility, progression, and outcomes^[10]. GWAS have led to the identification of several variants that are significantly associated with NAFLD. For example, one well-known genetic risk factor for NAFLD is a coding variant in patatin-like phospholipase domain containing protein 3 (*PNPLA3*), an I-to-M substitution at position 148 (chr22:43928847, rs738409 C>G)^[11]. This rs738409 variant has been repeatedly found to be associated with NAFLD or elevated hepatic fat content^[7,8,12]. Additionally, NAFLD susceptibility is significantly associated with four other genes: transmembrane 6 superfamily member 2 (*TM6SF2*), membrane-bound O-acyltransferase domain containing 7 (*MBOAT7*), glucokinase regulator (*GCKR*), and hydroxysteroid 17 β -dehydrogenase (*HSD17B13*)^[9].

This was followed by next-generation sequencing, such as whole-genome sequencing and whole-exon sequencing, emerging as post-GWAS era advancements in genetic research. However, unlike monogenic diseases, heritability in complex diseases like NAFLD is affected by thousands of common genetic variants and thus does not follow Mendelian inheritance. GWAS have been used to uncover thousands of genetic variants that influence the risks for complex human traits and diseases, and are thus more appropriate for elucidating the genetic factors related to NAFLD.

In the present review, we describe the five GWAS-identified risk variants that exhibit the most well-established associations with NAFLD [Table 1]. We also identify and discuss genetic associations between NAFLD and cardiovascular diseases, and suggest potential applications of genomic data for precision medicine.

PAPTATIN-LIKE PHOSPHOLIPASE DOMAIN CONTAINING PROTEIN 3

PNPLA3 p.I148M (chr22:43928847, rs738409 C>G) was the first NAFLD-related variant identified using GWAS^[7], and has exhibited a robust and well-replicated association with NAFLD in several studies^[13-15]. *PNPLA3* is highly expressed in the liver and adipose tissues. Its expression is regulated by insulin through a signaling pathway that includes *LXR* and *SREBP-1c*^[16], and is thus increased with feeding in animal

studies^[17]. The PNPLA3 protein hydrolyzes triglycerides and retinyl esters^[18]. The variant rs738409 C>G causes an isoleucine-to-methionine substitution at amino acid position 148 in *PNPLA3*, which results in impaired retinyl ester release and reduced hydrolase activity, causing fat accumulation within hepatocytes, including hepatic stellate cells^[19].

Studies of *PNPLA3* have transformed our knowledge of hepatic steatosis, revealing that lipid remodeling in intracellular lipid droplets is a common pathway underlying NAFLD progression, regardless of the environmental triggers. While the wild-type *PNPLA3* protein is rapidly degraded, the variant protein has no lipase activity, thereby leading to triglyceride accumulation in the liver^[19,20]. This can induce liver damage and inflammation, and can block the release of several extracellular proteins that protect against liver fibrosis, including matrix metalloproteinases and tissue inhibitor of metalloproteinases^[21]. The G allele of rs738409 is significantly associated with NAFLD activity score (NAS, $P = 0.004$), steatosis ($P = 0.03$), lobular inflammation ($P = 0.005$), portal inflammation ($P = 2.5 \times 10^{-4}$), and fibrosis ($P = 7.7 \times 10^{-6}$)^[22]. Moreover, homozygosity of this variant is reportedly linked to a 10-fold increased risk of developing NAFLD-associated HCC in the European population^[23]. Overall, these findings indicate that the G allele of rs738409 increases susceptibility to the whole spectrum of NAFLD - from steatosis to NASH (an inflammation-associated form of NAFLD), fibrosis, and HCC.

The *PNPLA3* gene could also be responsible for the different prevalence rates of NAFLD between ethnic groups. Different populations showed diverse odds ratios (ORs) for the variant rs738409 C>G, ranging from 2.08 to 18.23 [combined OR: 3.41 (2.57-4.52), $P < 0.00001$]^[7,8,11,12]. According to the genome aggregation database browser (gnomAD, <https://gnomad.broadinstitute.org/>), the G allele of rs738409 has a frequency of 27.1% in the general population, but occurs at a lower frequency in persons of African ethnicity (26.1%), and at a higher frequency in persons of Latino ethnicity (54.9%), which may have an impact on NAFLD risk in Latino populations^[7,12]. Accordingly, compared to other ethnic groups, persons of Latino ethnicity are reportedly more likely to progress to more severe forms of NAFLD^[24]. The effect of *PNPLA3* on NAFLD has also been described in East Asian cohorts. In two Japanese cohorts, the G risk allele of the rs738409 variant is significantly associated with NAFLD [OR: 1.66 (1.43-1.94), $P = 1.4 \times 10^{-10}$ and OR: 2.05, $P = 6.8 \times 10^{-14}$]^[11,25]. Similarly, among Korean NAFLD patients, rs738409 in *PNPLA3* is significantly associated with NAFLD [OR: 1.537 (1.383-1.709), $P = 1.74 \times 10^{-15}$]^[26]. Therefore, the rs738409 variant C>G in *PNPLA3* is strongly related to NAFLD progression in both Latino and East Asian cohorts.

TRANSMEMBRANE 6 SUPERFAMILY 2

Transmembrane 6 superfamily member 2 (TM6SF2) is a protein that localizes to the endoplasmic reticulum-Golgi apparatus of hepatocytes, and is involved in the increased hepatocytic secretion of triglyceride-rich lipoproteins *via* the pathway of very-low-density lipoprotein secretion. The *TM6SF2* polymorphism rs58542926 C>T (chr19:19268740, C>T) involves a C-to-T substitution at nucleotide 499, encoding a glutamate-to-lysine change at codon 167 (E167K). The variant rs58542926 leads to reduced *TM6SF2* expression, and is thus associated with increased hepatic lipid content. In a multi-ancestry study, the rs58542926 polymorphism was related to increased serum liver enzyme [alanine transaminase (ALT)] levels and a decreased serum lipid profile (total cholesterol and triglycerides)^[27]. Interestingly, rs58542926 has also been linked to a decreased risk of cardiovascular events based on the decreased circulating lipoprotein levels^[28,29]. The relationship between *TM6SF2* and serum liver enzyme (ALT) has also been identified in other large cohorts ($n > 80,000$)^[27]. Moreover, the variant rs58542926 has been associated with increased liver fibrosis ($P = 5.57 \times 10^{-5}$), independently of *PNPLA3* I148M^[30].

The G risk allele of rs58542926 occurs with a lower frequency (0.06969, gnomAD) in East Asia; thus, we explored the association of rs58542926 with NAFLD in East Asian studies. A Korean study reported that the co-existence of the risk alleles rs738409 and rs58542926 was associated with an increased risk of NASH

[OR: 2.03 (1.50-2.73), $P < 0.001$] and significant fibrosis [OR: 1.61 (1.19-2.17), $P < 0.002$], after adjustment for confounding variables^[31]. Similar findings were reported in a Chinese cohort^[32]. These results indicate that the rs58542926 variant in *TM6SF2* is associated with NAFLD, even in East Asia, where the allele frequency is low.

MEMBRANE-BOUND O-ACYLTRANSFERASE DOMAIN CONTAINING 7

Membrane-bound O-acyltransferase domain containing 7 (MBOAT7) is a protein involved in phosphatidylinositol remodeling with arachidonic acid in the Lands cycle. MBOAT7 is mainly expressed in the liver, including in hepatic sinusoidal cells, hepatic stellate cells, and hepatocytes. In several studies, the T allele of rs641738 in *MBOAT7* (chr19:54173068, rs641738 C>T) has been reported to increase the risk of developing the whole spectrum of NAFLD. Each T allele was associated with an increased risk of the development of hepatic steatosis [OR: 1.42 (1.07-1.91), $P = 0.015$], NASH [OR: 1.18 (1.00-1.40), $P = 0.05$], significant fibrosis [OR: 1.30 (1.06-1.70), $P = 0.012$], and HCC without advanced fibrosis [OR: 2.10 (1.33-3.31)]^[33-35].

GLUCOKINASE REGULATOR

Glucokinase regulator (GCKR) controls *de novo* lipogenesis by regulating the glucose influx into hepatocytes, which boosts the lipogenic pathway by providing further substrate for liver biosynthesis. Several variants in the *GCKR* gene are reportedly associated with NAFLD^[8,36]. The rs1260326 variant (chr2:27508073, C>T) encoding P446L has been considered as a causal variant for this association. In NAFLD patients, the T allele of rs1260326 is significantly associated with the hepatic fibrosis stage as compared to the F1 stage [OR: 2.06 (1.02-1.14), $P = 0.0008$]^[37]. The rs780094 variant in *GCKR* has also been significantly associated with computed tomography-proven and biopsy-proven NAFLD in a genome-wide association study (OR: 1.45, $P = 2.59 \times 10^{-8}$)^[8], and in a meta-analysis of five studies comprising of 2,091 NAFLD cases and 3,003 controls [OR: 1.25 (1.14-1.36), $P < 0.00001$]^[38].

The T risk allele of rs1260326 is associated with higher *GCKR* expression^[39]. Unlike the wild-type *GCKR* protein, the *GCKR* P446L protein is not sustained by fructose-6-phosphate, resulting in enhanced hepatic uptake of glucose, glucokinase activity^[40], and *de novo* lipogenesis^[41]. Interestingly, the risk allele of rs1260326 is also associated with decreased serum glucose levels and reduced T2DM risk^[42,43].

17 β -HYDROXYSTEROID DEHYDROGENASE TYPE 13

Several recent GWAS have identified a protective variant against NAFLD: rs72613567:TA in hydroxysteroid 17 β dehydrogenase 13 (*HSD17B13*)^[44,45]. This variant seems to be associated with decreased serum aspartate aminotransferase (AST, $P = 6.2 \times 10^{-10}$) and serum ALT ($P = 4.2 \times 10^{-12}$), and ameliorated progressive NASH among persons of European descent ($n = 46,544$). Furthermore, this splice variant (chr4:87310241, rs72613567: TA) in *HSD17B13* protects against chronic liver diseases, including both NAFLD and alcoholic liver disease, and this finding has been replicated in two studies^[45,46].

The protective effect of *HSD17B13* is mediated by reduced activity of the enzyme, which is involved in the conversion of retinol to retinoic acid^[47]. Retinoic acid reportedly suppresses fibrosis in NAFLD. This means that the protective effect of *HSD17B13* is not due to changes in hepatic fat accumulation, but rather caused by the enzymatic activity of lipid droplet-associated retinol dehydrogenase activity.

ASSOCIATION BETWEEN NAFLD-PREDISPOSING POLYMORPHISMS AND CARDIOVASCULAR DISEASE

Several studies report that cardiovascular disease (CVD) is the most common cause of mortality among NAFLD patients^[48-50]. This can be explained by the fact that NAFLD and CVD share common pathological

pathways, such as inflammation, endothelial dysfunction, and oxidative stress^[51,52]. Therefore, we explored the relationship between CVD and two of the most well-validated NAFLD-associated loci: *PNPLA3* and *TM6SF2*.

A meta-analysis study, including 60,801 coronary heart disease (CHD) cases and 123,504 controls, determined that the rs738409 variant in *PNPLA3* showed a protective effect against CVD [OR: 0.92 (0.87-0.97), $P = 0.002$]^[53]. In another study, the G allele of rs738409 was inversely related to CHD in 576 patients who underwent elective coronary angiography ($P = 0.02$)^[54]. However, this trend of association between the rs738409 G risk allele and CHD has been inconsistent. Another study found no association between the G risk allele of rs738409 and CHD [OR: 0.98 (0.95-1.02), $P = 0.79$]^[55]. Interestingly, in a study including 1,103 premature CHD patients and 1,469 healthy controls, the presence of the G allele of rs738409 was associated with increased risk of premature CHD development among T2DM patients [OR 1.20 (1.011-1.421), $P = 0.042$]^[56]. In addition to CHD, the G risk allele of rs738409 was reportedly linked to a greater risk of increased thickness of the carotid artery intima-media in 162 patients with biopsy-proven NAFLD [OR: 2.94 (1.12-7.70), $P = 0.02$], and this finding was validated in 267 patients with biopsy-proven or clinical NAFLD^[57].

The rs58542926 variant in *TM6SF2*, which is associated with fatty liver, is also reportedly protective against CVD by lowering serum lipid levels (total cholesterol, LDL-cholesterol, and triglycerides)^[58]. This means that the T allele of rs58542926 in *TM6SF2* confers protection against CVD at the expense of a higher risk of NAFLD. A meta-analysis confirmed that the rs58542926 T allele is associated with a tendency of decreased CVD risk [OR: 0.951 (0.92-0.98), $P = 0.005$]^[53]. Moreover, in a smaller cross-sectional study, this allele was related to a decreased risk of carotid artery plaques [OR: 0.49 (0.25-0.94)]^[29].

Despite these studies, the shared genetic causality between CVD and NAFLD remains unclear. The four genes that are most commonly reported to be associated with NAFLD - *PNPLA3*, *TM6SF2*, *MBOAT7*, and *GCKR* - have been analyzed by weighted fixed-effects statistical modeling, revealing no association of NAFLD with CVD [OR: 1.00 (0.99-1.91), $P = 0.93$]^[59]. These findings indicate that more complex relationships may exist between NAFLD and CVD. Exploring the roles of NAFLD-associated variants in CVD risk demonstrates that biologically relevant insights can be obtained by identifying individual pleiotropic loci that affect different outcomes. Clearly, more research is needed.

APPLICATIONS FOR PRECISION MEDICINE

There are several other genetic variants associated with NAFLD [Supplementary Table 1]. These associations with SNPs from GWAS are generally reported as P values and/or effect sizes. However, these metrics do not fully reflect the SNP's ability to differentiate between the control and the phenotype of interest. To apply genetic information for disease prediction, it is important to not only focus on the statistical power of a variant but also on the measurement of area under the ROC curve, which summarizes the true and false positive rates for a binary outcome^[60].

Another emerging metric is the development of polygenic risk scores (PRSs), which reflect the risk accumulation based on multiple SNPs, and can be calculated as a weighted sum of the disease risk alleles carried by an individual^[61]. PRS use would be a sensible approach in the study of NAFLD, as both common and rare SNPs are related to NAFLD risk, irrespective of clinical risk factors^[62]. However, there is little evidence of the clinical application of PRSs. Similarly, Krawczyk *et al.*^[63] reported that the summed number of risk alleles (0-5) for three genes (*PNPLA3*, *TM6SF2*, and *MBOAT7*) was significantly correlated with the individual risks of increased hepatic triglyceride content and elevated serum liver enzyme levels (AST and ALT). This study further suggests that the historical concept of genetic risk score (GRS) - which includes the contemporary concept of a continuous spectrum of NAFLD risk - could also be useful for predicting

Personalized Medicine

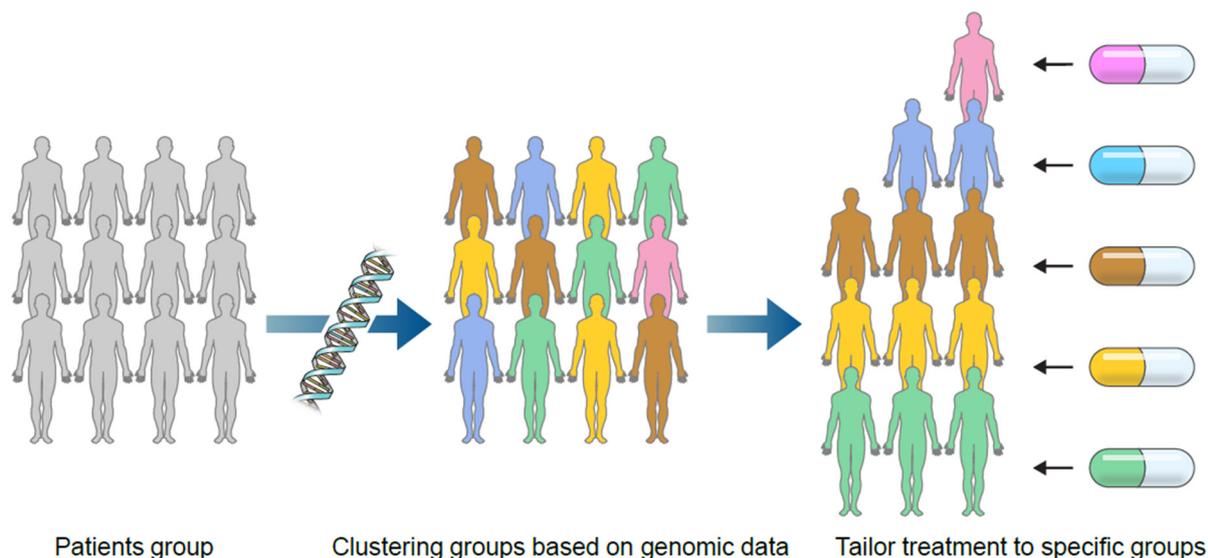


Figure 1. Toward personalized medicine based on genetic information. Personalized medicine makes it possible to tailor a treatment as individualized as the disease. This approach can be applied to solve tackling diseases that have far eluded effective treatments

NAFLD development and progression. Several NAFLD risk scoring models that incorporate both genetic and clinical information have also been proposed^[64]. For example, Hyysalo *et al.*^[65] developed a model for predicting NASH, which combines clinical variables and genetic information, based on European cohorts with biopsy-proven NAFLD. In another study, NAFLD-HCC was identified based on genotype information, age, sex, obesity, T2DM, and severe fibrosis, showing an AUROC of 0.96 ± 0.04 (89% specificity and 96% sensitivity)^[33].

Additional research is needed before PRSs, GRSs, or prediction models using both clinical variables and genetic information can be effectively applied for NAFLD investigation in clinical practice. This knowledge of genetic loci is potentially useful for risk stratification in patients with NAFLD. Moreover, considering that there are presently no approved drugs for NAFLD treatment^[66], there is an urgent need for more research and development of therapeutic targeting of the products of these genes in NAFLD patients with specific genetic variants that could provide insight into personalized treatments for NAFLD [Figure 1].

TRANSLATIONAL IMPLICATIONS AND CHALLENGES

As discussed above, several attempts have been made to predict NAFLD and/or NASH using genetic information alone or in combination with clinical information. The results of these efforts can be applied to the development of a new scoring model with better diagnostic performance compared to the previous models. Notably, a prediction model developed by combining serum metabolites, serum biochemical parameters, and genotype information was reported to discriminate NASH from NAFL with a good diagnostic performance^[67]. Such a model would be appealing for clinical translation, considering that the current gold standard for NASH diagnosis is liver biopsy, which is an invasive method. However, there are several challenges hindering the clinical translation of genetic information in NAFLD.

Firstly, most studies performed to evaluate the diagnostic accuracy of predictive models for NAFLD and/or NASH risk have been based on a cross-sectional design. Although these results can be useful, it is not an optimal study design for investigating models based on genetic information. Indeed, unlike classical factors such as biochemical results (AST, ALT), genetic variants have the strength of being stable over time. Thus,

if an ideal prediction model based on genetic information is properly established, it could be possible to stratify NAFLD before it develops or progresses, thus enabling intervention at an early stage or early age. However, to properly apply this concept, results should be accumulated from multiple longitudinal studies.

Secondly, another important issue to consider when utilizing genetic information is the interaction between environment and gene. For example, with regards to rs738409 in the *PNPLA3* gene, it has been reported that the variant's effect is especially amplified in the setting of obesity^[68]. This suggests that adiposity (environment) can influence how specific genetic information influences the full spectrum of NAFLD. Considering this interaction between gene and environment, a prediction model based exclusively on genetic information may not exhibit sufficient predictive power, while a model with integration of relevant NAFLD-associated clinical factors would be more likely to reach significant predictive power.

Overall, prediction models that use both genetic information and relevant clinical factors derived from longitudinal studies can achieve sufficient predictive power for NAFLD risk stratification at the individual level.

CONCLUSION

The identification of genes associated with NAFLD development and progression is expected to provide important insights into its pathophysiology, as well as to guide disease risk stratification and further new opportunities for timely therapeutic intervention. Several genetic variants have been implicated in NAFLD development and progression, and here we focused on the five genes whose associations with NAFLD have been most extensively replicated. These genetic risk variants can improve the accuracy of NAFLD diagnosis, and may also be useful for the identification of high-risk NAFLD patients who have unfavorable prognoses. An understanding of these NAFLD-associated genetic risk factors will help identify individuals at risk, and potentially guide the provision of appropriate treatments based on an individual's risk and likelihood of disease progression.

DECLARATIONS

Acknowledgements

The authors thank Medical Illustration & Design, part of the Medical Research Support Services of Yonsei University College of Medicine, for all artistic support related to this work.

Authors' contributions

Wrote and reviewed the manuscript: Kim DY, Park JY

Availability of data and materials

Not applicable.

Financial support and sponsorship

This research was supported by the Bio & Medical Technology Development Program of the National Research Foundation (NRF) funded by the Ministry of Science & ICT (NRF-2020M3A9E4038694).

Conflicts of interest

All authors declared that there are no conflicts of interest.

Ethical approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

Copyright

© The Author(s) 2020.

REFERENCES

1. Younossi Z, Anstee QM, Marietti M, et al. Global burden of NAFLD and NASH: trends, predictions, risk factors and prevention. *Nat Rev Gastroenterol Hepatol* 2018;15:11.
2. Eslam M, George J. Genetic and epigenetic mechanisms of NASH. *Hepatol Int* 2016;10:394-406.
3. Daly AK, Ballestri S, Carulli L, Loria P, Day CP. Genetic determinants of susceptibility and severity in nonalcoholic fatty liver disease. *Expert Rev Gastroenterol Hepatol* 2011;5:253-63.
4. Abdelmalek MF, Liu C, Shuster J, Nelson DR, Asal NR. Familial aggregation of insulin resistance in first-degree relatives of patients with nonalcoholic fatty liver disease. *Clin Gastroenterol Hepatol* 2006;4:1162-9.
5. Willner IR, Waters B, Patil SR, Reuben A, Morelli J, Riely CA. Ninety patients with nonalcoholic steatohepatitis: insulin resistance, familial tendency, and severity of disease. *Am J Gastroenterol* 2001;96:2957-61.
6. Chalasani N, Guo X, Loomba R, et al. Genome-wide association study identifies variants associated with histologic features of nonalcoholic fatty liver disease. *Gastroenterology* 2010;139:1567-76.e6.
7. Romeo S, Kozlitina J, Xing C, et al. Genetic variation in PNPLA3 confers susceptibility to nonalcoholic fatty liver disease. *Nat Genet* 2008;40:1461-5.
8. Speliotes EK, Yerges-Armstrong LM, Wu J, et al. Genome-wide association analysis identifies variants associated with nonalcoholic fatty liver disease that have distinct effects on metabolic traits. *PLoS Genet* 2011;7:e1001324.
9. Lohmueller KE, Pearce CL, Pike M, Lander ES, Hirschhorn JN. Meta-analysis of genetic association studies supports a contribution of common variants to susceptibility to common disease. *Nat Genet* 2003;33:177-82.
10. Eslam M, Valenti L, Romeo S. Genetics and epigenetics of NAFLD and NASH: clinical impact. *J Hepatol* 2018;68:268-79.
11. Kawaguchi T, Sumida Y, Umemura A, et al. Genetic polymorphisms of the human PNPLA3 gene are strongly associated with severity of non-alcoholic fatty liver disease in Japanese. *PLoS One* 2012;7:e38322.
12. Xu R, Tao A, Zhang S, Deng Y, Chen G. Association between patatin-like phospholipase domain containing 3 gene (PNPLA3) polymorphisms and nonalcoholic fatty liver disease: a HuGE review and meta-analysis. *Sci Rep* 2015;5:9284.
13. Dongiovanni P, Romeo S, Valenti L. Genetic factors in the pathogenesis of nonalcoholic fatty liver and steatohepatitis. *Biomed Res Int* 2015;2015:460190.
14. Dongiovanni P, Donati B, Fares R, et al. PNPLA3 I148M polymorphism and progressive liver disease. *World J Gastroenterol* 2013;19:6969.
15. Singal AG, Manjunath H, Yopp AC, et al. The effect of PNPLA3 on fibrosis progression and development of hepatocellular carcinoma: a meta-analysis. *Am J Gastroenterol* 2014;109:325.
16. Cohen JC, Horton JD, Hobbs HH. Human fatty liver disease: old questions and new insights. *Science* 2011;332:1519-23.
17. Huang Y, He S, Li JZ, et al. A feed-forward loop amplifies nutritional regulation of PNPLA3. *Proc Natl Acad Sci U S A* 2010;107:7892-7.
18. Huang Y, Cohen JC, Hobbs HH. Expression and characterization of a PNPLA3 protein isoform (I148M) associated with nonalcoholic fatty liver disease. *J Biol Chem* 2011;286:37085-93.
19. BasuRay S, Smagris E, Cohen JC, Hobbs HH. The PNPLA3 variant associated with fatty liver disease (I148M) accumulates on lipid droplets by evading ubiquitylation. *Hepatology* 2017;66:1111-24.
20. Luukkonen PK, Nick A, Hölttä-Vuori M, et al. Human PNPLA3-I148M variant increases hepatic retention of polyunsaturated fatty acids. *JCI Insight* 2019;4:e127902.
21. Pingitore P, Dongiovanni P, Motta BM, et al. PNPLA3 overexpression results in reduction of proteins predisposing to fibrosis. *Hum Mol Genet* 2016;25:5212-22.
22. Rotman Y, Koh C, Zmuda JM, Kleiner DE, Liang TJ, CRN N. The association of genetic variability in patatin-like phospholipase domain-containing protein 3 (PNPLA3) with histological severity of nonalcoholic fatty liver disease. *Hepatology* 2010;52:894-903.
23. Liu YL, Patman G, Leathart J, et al. Carriage of the PNPLA3 rs738409 C>G polymorphism confers an increased risk of non-alcoholic fatty liver disease associated hepatocellular carcinoma. *J Hepatol* 2014;61:75-81.
24. Carrion AF, Ghanta R, Carrasquillo O, Martin P. Chronic liver disease in the Hispanic population of the United States. *Clin Gastroenterol Hepatol* 2011;9:834-41.
25. Kitamoto T, Kitamoto A, Yoneda M, et al. 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* 2013;132:783-92.
26. Chung GE, Lee Y, Yim JY, et al. Genetic polymorphisms of PNPLA3 and SAMM50 are associated with nonalcoholic fatty liver disease in a Korean population. *Gut Liver* 2018;12:316-23.
27. Kozlitina J, Smagris E, Stender S, et al. Exome-wide association study identifies a TM6SF2 variant that confers susceptibility to nonalcoholic fatty liver disease. *Nat Genet* 2014;46:352-6.
28. Holmen OL, Zhang H, Fan Y, et al. Systematic evaluation of coding variation identifies a candidate causal variant in TM6SF2 influencing

- total cholesterol and myocardial infarction risk. *Nat Genet* 2014;46:345-51.
29. Dongiovanni P, Petta S, Maglio C, et al. Transmembrane 6 superfamily member 2 gene variant disentangles nonalcoholic steatohepatitis from cardiovascular disease. *Hepatology* 2015;61:506-14.
 30. Liu YL, Reeves HL, Burt AD, et al. TM6SF2 rs58542926 influences hepatic fibrosis progression in patients with non-alcoholic fatty liver disease. *Nat Commun* 2014;5:1-6.
 31. Koo BK, Joo SK, Kim D, et al. Additive effects of PNPLA3 and TM6SF2 on the histological severity of non-alcoholic fatty liver disease. *J Gastroenterol Hepatol* 2018;33:1277-85.
 32. Wang X, Liu Z, Wang K, et al. Additive effects of the risk alleles of PNPLA3 and TM6SF2 on non-alcoholic fatty liver disease (NAFLD) in a Chinese population. *Front Genet* 2016;7:140.
 33. Donati B, Dongiovanni P, Romeo S, et al. MBOAT7 rs641738 variant and hepatocellular carcinoma in non-cirrhotic individuals. *Sci Rep* 2017;7:1-10.
 34. Buch S, Stickel F, Trépo E, et al. A genome-wide association study confirms PNPLA3 and identifies TM6SF2 and MBOAT7 as risk loci for alcohol-related cirrhosis. *Nat Genet* 2015;47:1443-8.
 35. Mancina RM, Dongiovanni P, Petta S, et al. The MBOAT7-TMC4 variant rs641738 increases risk of nonalcoholic fatty liver disease in individuals of European descent. *Gastroenterology* 2016;150:1219-30.e6.
 36. Valenti L, Alisi A, Nobili V. Unraveling the genetics of fatty liver in obese children: additive effect of P446L GCKR and I148M PNPLA3 polymorphisms. *Hepatology* 2012;55:661-3.
 37. Petta S, Miele L, Bugianesi E, et al. Glucokinase regulatory protein gene polymorphism affects liver fibrosis in non-alcoholic fatty liver disease. *PLoS One* 2014;9:e87523.
 38. Zain SM, Mohamed Z, Mohamed R. A common variant in the glucokinase regulatory gene rs780094 and risk of nonalcoholic fatty liver disease: a meta-analysis. *J Gastroenterol Hepatol* 2015;30:21-7.
 39. Rodríguez ML, Silva LF, Vangipurapu J, et al. Functional variant in the GCKR gene affects lactate levels differentially in the fasting state and during hyperglycemia. *Sci Rep* 2018;8:1-8.
 40. Beer NL, Tribble ND, McCulloch LJ, et al. The P446L variant in GCKR associated with fasting plasma glucose and triglyceride levels exerts its effect through increased glucokinase activity in liver. *Hum Mol Genet* 2009;18:4081-8.
 41. Anstee QM, Day CP. The genetics of NAFLD. *Nat Rev Gastroenterol Hepatol* 2013;10:645-55.
 42. Sparsø T, Andersen G, Nielsen T, et al. The GCKR rs780094 polymorphism is associated with elevated fasting serum triacylglycerol, reduced fasting and OGTT-related insulinaemia, and reduced risk of type 2 diabetes. *Diabetologia* 2008;51:70-5.
 43. Onuma H, Tabara Y, Kawamoto R, et al. The GCKR rs780094 polymorphism is associated with susceptibility of type 2 diabetes, reduced fasting plasma glucose levels, increased triglycerides levels and lower HOMA-IR in Japanese population. *J Hum Genet* 2010;55:600-4.
 44. Abul-Husn NS, Cheng X, Li AH, et al. A protein-truncating HSD17B13 variant and protection from chronic liver disease. *N Engl J Med* 2018;378:1096-106.
 45. Pirola CJ, Garaycoechea M, Flichman D, et al. Splice variant rs72613567 prevents worst histologic outcomes in patients with nonalcoholic fatty liver disease. *J Lipid Res* 2019;60:176-85.
 46. Ma Y, Belyaeva OV, Brown PM, et al. HSD17B13 is a hepatic retinol dehydrogenase associated with histological features of non-alcoholic fatty liver disease. *Hepatology* 2019;69:1504.
 47. Ma Y, Belyaeva OV, Brown PM, et al. 17-beta hydroxysteroid dehydrogenase 13 is a hepatic retinol dehydrogenase associated with histological features of nonalcoholic fatty liver disease. *Hepatology* 2019;69:1504-19.
 48. Haddad TM, Hamdeh S, Kanmanthareddy A, Alla VM. Nonalcoholic fatty liver disease and the risk of clinical cardiovascular events: a systematic review and meta-analysis. *Diabetes Metab Syndr* 2017;11:S209-16.
 49. Younossi ZM, Stepanova M, Rafiq N, et al. Nonalcoholic steatofibrosis independently predicts mortality in nonalcoholic fatty liver disease. *Hepatol Commun* 2017;1:421-8.
 50. Targher G, Byrne CD, Lonardo A, Zoppini G, Barbui C. Non-alcoholic fatty liver disease and risk of incident cardiovascular disease: a meta-analysis. *J Hepatol* 2016;65:589-600.
 51. Liu H, Lu HY. Nonalcoholic fatty liver disease and cardiovascular disease. *World J Gastroenterol* 2014;20:8407.
 52. Francque SM, van der Graaff D, Kwanten WJ. Non-alcoholic fatty liver disease and cardiovascular risk: Pathophysiological mechanisms and implications. *J Hepatol* 2016;65:425-43.
 53. Simons N, Isaacs A, Koek GH, Kuč S, Schaper NC, Brouwers MCGJ. PNPLA3, TM6SF2, and MBOAT7 genotypes and coronary artery disease. *Gastroenterology* 2017;152:912-3.
 54. Rüschenbaum S, Schwarzkopf K, Friedrich-Rust M, et al. Patatin-like phospholipase domain containing 3 variants differentially impact metabolic traits in individuals at high risk for cardiovascular events. *Hepatol Commun* 2018;2:798-806.
 55. Lauridsen BK, Stender S, Kristensen TS, et al. Liver fat content, non-alcoholic fatty liver disease, and ischaemic heart disease: Mendelian randomization and meta-analysis of 279 013 individuals. *Eur Heart J* 2017;39:385-93.
 56. Posadas-Sánchez R, López-Urbe ÁR, Posadas-Romero C, et al. Association of the I148M/PNPLA3 (rs738409) polymorphism with premature coronary artery disease, fatty liver, and insulin resistance in type 2 diabetic patients and healthy controls. The GEA study. *Immunobiology* 2017;222:960-6.
 57. Petta S, Valenti L, Marchesini G, et al. PNPLA3 GG genotype and carotid atherosclerosis in patients with non-alcoholic fatty liver disease. *PLoS One* 2013;8:e74089.
 58. Pirola CJ, Sookoian S. The dual and opposite role of the TM6SF2-rs58542926 variant in protecting against cardiovascular disease and conferring risk for nonalcoholic fatty liver: a meta-analysis. *Hepatology* 2015;62:1742-56.

59. Brouwers MC, Simons N, Stehouwer CD, Koek GH, Schaper NC, Isaacs A. Relationship between nonalcoholic fatty liver disease susceptibility genes and coronary artery disease. *Hepatol Commun* 2019;3:587-96.
60. Kraft P, Wacholder S, Cornelis MC, et al. Beyond odds ratios - communicating disease risk based on genetic profiles. *Nat Rev Genet* 2009;10:264-9.
61. Chatterjee N, Shi J, Garcia-Closas M. Developing and evaluating polygenic risk prediction models for stratified disease prevention. *Nat Rev Genet* 2016;17:392.
62. Dongiovanni P, Stender S, Pietrelli A, et al. Causal relationship of hepatic fat with liver damage and insulin resistance in nonalcoholic fatty liver. *J Intern Med* 2018;283:356-70.
63. Krawczyk M, Liebe R, Lammert F. Toward genetic prediction of nonalcoholic fatty liver disease trajectories: PNPLA3 and beyond. *Gastroenterology* 2020;158:1865-80.e1.
64. Vespasiani-Gentilucci U, Gallo P, Dell'Unto C, Volpentesta M, Antonelli-Incalzi R, Picardi A. Promoting genetics in non-alcoholic fatty liver disease: combined risk score through polymorphisms and clinical variables. *World J Gastroenterol* 2018;24:4835.
65. Hyysalo J, Männistö VT, Zhou Y, et al. A population-based study on the prevalence of NASH using scores validated against liver histology. *J Hepatol* 2014;60:839-46.
66. Kim DY, Park JY. Overview of emerging treatment of non-alcoholic fatty liver disease: more than one drug needed? *Hepatobiliary Surg Nutr* 2019;8:522.
67. Zhou Y, Orešič M, Leivonen M, et al. Noninvasive detection of nonalcoholic steatohepatitis using clinical markers and circulating levels of lipids and metabolites. *Clin Gastroenterol Hepatol* 2016;14:1463-72.e6.
68. Romeo S, Sentinelli F, Dash S, et al. Morbid obesity exposes the association between PNPLA3 I148M (rs738409) and indices of hepatic injury in individuals of European descent. *Int J Obes* 2010;34:190-4.