Introduction

Non-alcoholic fatty liver disease (NAFLD) has become the most common liver disease, the world over and encompasses a spectrum ranging from simple steatosis (non-alcoholic fatty liver or NAFL), non-alcoholic steatohepatitis (NASH), NASH related cirrhosis and hepatocellular carcinoma (HCC) (1,2). In some parts of the world, equally common problem is the alcoholic liver disease (ALD) with a similar spectrum of alcoholic steatosis, alcoholic steatohepatitis, cirrhosis and HCC (3). There are significant inter-individual differences in the severity and progression of liver disease in patients with NAFLD sharing similar metabolic and other environmental risk factors (4). Similarly severity...
Genetic basis in the pathogenesis of NAFLD

NAFLD is a complex disease with multiple factors involved in the pathogenesis that include both environmental and genetic modifiers. Also, NASH and fibrosis progression varies with some patients being ‘rapid progressors’ (12,13). Risk factors for fibrosis progression in patients with NAFLD include metabolic syndrome, presence of type 2 diabetes mellitus, high body mass index (BMI)/obesity, age, steatosis grade, high insulin resistance, baseline biopsy showing inflammation and fibrosis (1,2,13). In a meta-analysis of studies including paired biopsies, Singh et al. showed that rate of fibrosis progression was twice as common in patients with NASH in comparison to patients with NAFL (12).

In addition to environmental and metabolic risk factors, familial clustering, studies involving the monozygotic twins and inter-ethnic differences point towards the role of genetics in the pathogenesis of NAFLD.

Familial clustering in NAFLD

Several studies have shown the clustering of cases of NAFLD among the families (8,14,15). Brouwers et al. compared 157 familial combined hyperlipidemia family members with 20 spouses. Fatty liver was more prevalent in probands (40%) and relatives (35%) compared to spouses (15%) (8). In a study, siblings and parents of overweight/obese children with and without NAFLD were compared with the help of magnetic resonance proton density fat fraction (MR-PDFF) (15). NAFLD was present in 17% of siblings and 37% of parents in non-NAFLD overweight/obese children group in comparison to presence of NAFLD in 59% of siblings and 78% of parents of overweight/obese children with NAFLD. The correlation of liver fat fraction to BMI was stronger in families of children with NAFLD than without NAFLD. The authors found heritability of NAFLD as 1.000 and that of liver fat fraction as 0.386 (15).

Twin studies

Prevalence studies in twins suggest a stronger genetic basis of NAFLD in comparison to that evident from familial clustering studies. Several twin studies have shown higher prevalence of NAFLD in monozygotic twins as compared to dizygotic twins. Makkonen et al. studied 313 twin pairs and observed significantly higher intra-pair correlations in monozygotic than the dizygotic twins for both alanine transaminases (ALT) (0.65 and 0.04 respectively) and fasting serum insulin (0.58 and 0.34 respectively). Heritability of ALT was 55% and that of fasting serum insulin was 61%. ALT and fasting serum insulin correlated with liver fat content in 66 subjects (measured by magnetic resonance spectroscopy) (16). In a study using ultrasonography for NAFLD and carotid intima media thickness in 208 (58 men and 150 women) Hungarian twins (63 monozygotic and 41 dizygotic pairs, aged 43.7±16.7 years), heritability was linked to cardiovascular risk, but not to NAFLD (17). However, it should be noted that ultrasound with its limitations may have categorized some patients with mild steatosis as normal on ultrasound. Loomba et al. studied 60 pairs of twins (42 monozygotic and 18 dizygotic) aged 45.7±22.1 years. Steatosis and fibrosis were measured by MR-PDFF and MR elastography (MRE) in the study subjects. Twenty-six (21.7%) had NAFLD, which correlated between monozygotic twins (r²=0.70; P<0.0001) but not in dizygotic twins. The fibrosis also correlated between monozygotic twins (r²=0.48; P<0.002) but not in dizygotic twins (r²=0.12; P=0.7). The heritability of hepatic steatosis was 0.52 and of hepatic fibrosis was 0.5 (18). In a prospective study, Cui et al. studied 65 twin pairs (45 monozygotic, 20 dizygotic twin pairs, aged 47.1±21.9 years) and found that 20% (n=26) had hepatic steatosis and 8.2% (n=10) had hepatic fibrosis. Steatosis and fibrosis had a significant shared gene effect of 0.756 (95% CI, 0.716–1) (19).

Ethnic differences

Major evidence for having the genetic basis in NAFLD comes from studies showing the ethnic differences in the prevalence and severity of NAFLD. In a study including cryptogenic cirrhosis or NASH related cirrhosis, it was seen that prevalence of cirrhosis among Hispanic and African American patients was 3.1 times higher and 3.9 times lower than European American patients despite similar prevalence of type 2 diabetes mellitus among these groups (20). Wagenknecht et al. compared 795
Hispanic-American and 347 African-American adults, aged 49 [22–84] years. NAFLD was present in 24% of Hispanics versus 10% in African-American diagnosed on the basis of CT scan. NAFLD was independently associated with insulin sensitivity and visceral adipose tissue area in both ethnic groups, proportion of explained variance being higher in Hispanics (21). Browning et al. studied liver steatosis by proton magnetic resonance spectroscopy in 2,287 patients of multi-ethnic population comprising of 32.1% Whites, 48.3% African Americans and 17.5% Hispanics. NAFLD was present in 45% of Hispanics; 33% of Whites and 24% of African Americans. Higher prevalence of NAFLD in Hispanics in comparison to African Americans in spite of the presence of metabolic risk factors in later group suggested the genetic basis for the difference (6). A study analyzing 1,026 adults with biopsy proven NAFLD from the Nonalcoholic Steatohepatitis Clinical Research Network observed that HOMA-IR was not a significant risk factor for NASH among Latinos, but was significant among non-Latino Whites (22). In a systemic review of 34 studies (368,569 patients), NAFLD prevalence was highest in the Hispanics (relative risk 1.09) and lower in African Americans (relative risk 0.72) in comparison to Whites, with no difference in hepatic fibrosis (7). While most of data regarding ethnic differences is available in non-Asians, few studies have addressed this issue in the Asian population. Mohanty et al. showed that Asians had higher grades of histological ballooning in comparison to patients of other ethnicities (OR 2.71, P=0.04). Hispanics on the other hand showed a higher prevalence of Mallory hyaline in comparison to patients of other ethnicities (OR 2.41, P=0.03) and African Americans had lower degree of hepatic steatosis (23). In another study from John Hopkins Hospital, Baltimore, Maryland, authors observed severe hepatic steatosis and a trend towards severe inflammation in Asian-Americans (24). Petersen et al. compared results of oral glucose tolerance test in Caucasians (n=292), Asian-Indians (n=59), Eastern Asians (n=49), African Americans (n=48) and Hispanics (n=34). The prevalence of insulin resistance was 2–3-fold higher in the Asian-Indians as compared to other ethnicities. There was approximately 2-fold increase in hepatic triglyceride content in Asians as compared to Caucasian men (25). The difference in the prevalence and severity among patients with different ethnicities is best explained by the difference in various gene variants.

Genetic modifiers in NAFLD

Two types of genetic studies are available in patients with NAFLD. Candidate gene studies are hypothesis-testing studies, which are done for a gene with known functions. Candidate gene studies look for difference regarding a polymorphism in cases and controls. A small sample size is needed for candidate gene studies but these studies are not able to find new genetic associations. On the other hand, genome wide association studies (GWAS) are hypothesis-generating studies. GWAS look at thousands to millions of short nucleotide polymorphisms (SNPs). GWAS are done in a large sample size, and are useful to find new genetic associations of a disease. GWAS studies in NAFLD are shown in Table 1 and candidate gene studies are shown in Table 2.

Genetic association of NAFLD with PNPLA3, TM6SF2 and other gene variants

The strongest evidence for the genetic link in NAFLD has been shown to be related to Patatin-like phospholipase domain-containing protein 3 (PNPLA3) also known as adiponutrin (ADPN), acylglycerol O-acyltransferase or calcium-independent phospholipase A2-epsilon (iPLA2-epsilon), an enzyme that in humans is encoded by the PNPLA3 gene located on chromosome 22. The PNPLA3 variant (rs738409 C > G, p.I148M) is a cytosine to guanine nucleotide transversion mutation at codon 148 that causes isoleucine to methionine amino acid change (63) and has been shown to be strongly associated with increased hepatic steatosis. PNPLA3 or Adiponutrin is a triacylglycerol lipase that possesses both lipolytic and lipogenic activities in vitro.

The I148M substitution has been shown to cause both gain and loss of function; loss of function by decreasing lipolytic activity, thus increasing triglyceride accumulation (64). Pirazzi et al. studied very-low-density lipoprotein (VLDL) kinetics in 55 overweight/obese men. The authors showed that PNPLA3 I148M variant affected secretion of apoB-containing lipoproteins, suggestive of a loss-of-function mutation. The authors suggested that PNPLA3 148M promoted intracellular lipid accumulation by reducing the lipiddation of VLDL (65). There is also evidence of increased lipid synthesis by I148M variant. In a study from Austria, it was shown that PNPLA3 also acts as acyl-CoA-dependent lysophosphatidic acid acyltransferase (LPAAT); I148M variant showed elevated LPAAT
Table 1 Showing genome wide association studies (GWAS) in NAFLD

<table>
<thead>
<tr>
<th>Author, year (reference)</th>
<th>n</th>
<th>Gene, SNP</th>
<th>Comments</th>
</tr>
</thead>
<tbody>
<tr>
<td>Romeo, 2008 (26)</td>
<td>2,111</td>
<td>PNPLA3, rs738409[G]</td>
<td>Associated with steatosis and inflammation</td>
</tr>
<tr>
<td>Yuan, 2008 (27)</td>
<td>7,715</td>
<td>CPN1-ERLIN1-CHUK on chromosome 10 and PNPLA3-SAMM50, HNF1A on chromosome 12</td>
<td>Associated with ALT (first 2) and GGT levels</td>
</tr>
<tr>
<td>Rotman, 2010 (28)</td>
<td>1,117</td>
<td>PNPLA3, rs738409, two other SNP near same region</td>
<td>rs738409 G associated with younger age at time of biopsy</td>
</tr>
<tr>
<td>Chalasani, 2010 (29)</td>
<td>226</td>
<td>rs2645424, rs343062, rs1227756, rs6591182, s867304 (multiple genes)</td>
<td>Associated with severity of disease (activity score, inflammation, fibrosis)</td>
</tr>
<tr>
<td>Chambers, 2011 (30)</td>
<td>61,089</td>
<td>PNPLA3, rs738409, two other SNP near same region</td>
<td>Genes involved in biliary transport, glucose, carbohydrate and lipid metabolism, inflammation, immunity and glutathione metabolism were important</td>
</tr>
<tr>
<td>Speliotes, 2011 (31)</td>
<td>7,176</td>
<td>Variants in or near NCAN, GCKR, LYPLAL1, and PNPLA3</td>
<td>Associated with serum lipids as well as glycemic and anthropometric traits</td>
</tr>
<tr>
<td>Kawaguchi, 2012 (32)</td>
<td>529 NAFLD and 932 population controls</td>
<td>Associated with severity</td>
<td></td>
</tr>
<tr>
<td>Feitosa, 2013 (33)</td>
<td>2,767</td>
<td>Variants of the ERLIN1-CHUK-CWF19L1 gene cluster</td>
<td>Associated with steatosis and ALT levels</td>
</tr>
<tr>
<td>Kozlitina, 2014 (34)</td>
<td>2,736</td>
<td>PNPLA3 (rs738409 and rs2281135), rs58542926 in TM6SF2</td>
<td>TM6SF2 variant associated with decreased VLDL secretion</td>
</tr>
<tr>
<td>DiStefano, 2015 (35)</td>
<td>2,300</td>
<td>rs4823173, rs2896019, and rs2281135 (PNPLA3) and rs10401969 in SUGP1</td>
<td>Identified new loci</td>
</tr>
</tbody>
</table>

NAFLD, non-alcoholic fatty liver disease; SNP, short nucleotide polymorphism; VLDL, very-low-density lipoprotein.

activity in comparison to the wild-type, thus promoting hepatic lipid synthesis by gain of function (66). PNPLA3 also affects hepatic stellate cells (HSCs), which play an important role in the development of liver fibrosis. HSCs are the main reservoir of the retinoids. Upon activation, HSCs lose retinol content and differentiate into activated myofibroblasts, which produce collagen (fibrosis). PNPLA3 also has retinyl-esterase activity and is involved in retinol metabolism. Expression of PNPLA3 gene and protein remain increased in activated HSCs. While induction of wild type of PNPLA3 is associated with reduced secretion of matrix metalloproteinase 2 and tissue inhibitor of metalloproteinase 1 (1,2), HSCs with I148M show higher expression and release of proinflammatory cytokines (67-69). Donati et al. showed that a different PNPLA3 variant, rs2294918 E434K decreased the effect of the I148M variant (70). Lindén et al. studied antisense oligonucleotide (ASO) mediated silencing of PNPLA3 in a knock-in mice model. The ASO mediated silencing of PNPLA3 led to the reduction of hepatic steatosis, inflammation, NAFLD activity score and fibrosis stage (71). This study provides the first evidence that a PNPLA3 ASO therapy can improve all features of NAFLD including fibrosis.

Studying the ancestry-related and inter-individual differences in hepatic fat content and susceptibility to NAFLD, Romeo et al. conducted a genome-wide association scan of nonsynonymous sequence variations in Hispanics (n=383), African-Americans (n=1,032) and European-Americans (n=696). The authors found an allele of PNPLA3 [rs738409 (G), encoding I148M] to be strongly associated with increased hepatic fat levels and inflammation. This allele was most common in Hispanics, which were most susceptible to NAFLD. The hepatic fat content was more than two times higher in PNPLA3
<table>
<thead>
<tr>
<th>Pathway</th>
<th>Genes (reference)</th>
<th>Polymorphism</th>
<th>Results/comments</th>
</tr>
</thead>
<tbody>
<tr>
<td>Glucose metabolism and insulin resistance</td>
<td>ENPP1 (36)</td>
<td>ENPP1 121Gln; Associated with insulin resistance</td>
<td>Associated with fibrosis</td>
</tr>
<tr>
<td></td>
<td>Insulin-receptor</td>
<td>IRS-1 972Arg; Associated with insulin resistance</td>
<td>Associated with fibrosis</td>
</tr>
<tr>
<td></td>
<td>substrate 1 (36)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>GCKR (37)</td>
<td>GCKR (rs780094 and rs1260326, encoding Pro446Leu); glucokinase regulatory</td>
<td>Associated with increased serum triglycerides and higher liver fibrosis</td>
</tr>
<tr>
<td></td>
<td></td>
<td>protein (GCKR)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>SLC2A1 (38)</td>
<td>SLC2A1, in vitro down-regulation promotes lipid accumulation, increased</td>
<td>Associated with NAFLD, no association with T2DM</td>
</tr>
<tr>
<td></td>
<td></td>
<td>oxidative stress</td>
<td></td>
</tr>
<tr>
<td>Lipid metabolism</td>
<td>PNPLA3 (40)</td>
<td>rs738409</td>
<td>More severe disease</td>
</tr>
<tr>
<td></td>
<td>TM6SF2 (41-43)</td>
<td>TM6SF2 rs10401969 (C); rs58542926 (C/T) E167K</td>
<td>Associated with NAFLD (steatosis, NASH, fibrosis/cirrhosis</td>
</tr>
<tr>
<td></td>
<td>MBOTAT (40,44)</td>
<td>rs641738</td>
<td>Associated with increased hepatic fat content, severe liver disease and increased risk of fibrosis</td>
</tr>
<tr>
<td></td>
<td>Fatty acid</td>
<td>FADS1</td>
<td>Alleles associated with decreased expression of FADS1 were associated with greater steatosis</td>
</tr>
<tr>
<td></td>
<td>desaturase 1 (45)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Lipin 1 (46)</td>
<td>LIPI (rs13412852 C&gt;T)</td>
<td>Associated with severity and fibrosis, TT genotype protective</td>
</tr>
<tr>
<td></td>
<td>Nuclear receptor</td>
<td>NR1I2 (rs7643645 and rs2461823)</td>
<td>Two SNPs associated with NAFLD severity</td>
</tr>
<tr>
<td></td>
<td>subfamily 1 group</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Lmber 2 (also</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>known as</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>pregnane X</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>receptor (47)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>PPAR alpha (48,49)</td>
<td>PPAR alpha (rs1800234, encoding Val227Ala)</td>
<td>PPAR alpha limits TG accumulation by increasing fatty acid oxidation</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Phosphatidylethanolamine</td>
<td>PEMT (rs7946, encoding Val175Met)</td>
<td>Association with NAFLD</td>
</tr>
<tr>
<td></td>
<td>N-methyltransferase (50,51)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>17-beta hydroxysteroid</td>
<td>rs6834314</td>
<td>Associated with increased steatosis but decreases evertiy</td>
</tr>
<tr>
<td></td>
<td>dehydrogenase 13</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(52)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fibrosis</td>
<td>Type-1 angiotensin 2; Receptor (53,54)</td>
<td>AGTR1 (rs3772622)</td>
<td>Associated with steatohepatitis and fibrosis</td>
</tr>
<tr>
<td>Oxidative stress</td>
<td>GCLC (55)</td>
<td>−129 C/T; polymorphism</td>
<td>Associated with steatohepatitis compared (OR 12.14)</td>
</tr>
<tr>
<td></td>
<td>SOD2 (56)</td>
<td>SOD2 A16V (rs4880)</td>
<td>Associated with NAFLD fibrosis, risk factor for severe alcoholic disease</td>
</tr>
<tr>
<td></td>
<td>Uncoupling protein 2 (UCP2) (57)</td>
<td>UCP2 (−866 G &gt; A, rs695366)</td>
<td>Associated with reduced risk of NASH</td>
</tr>
</tbody>
</table>

Table 2 (continued)
rs738409 [G] homozygotes than in non-carriers. Another allele of PNPLA3 [rs6006460 (T), encoding S453I] was associated with lower hepatic fat content in African-Americans (26). Palmer et al. studied variants associated with NAFLD in persons with European ancestry, Africans and Hispanics. The authors found that steatosis was 0.20–0.34 heritable in Africans and Hispanic-American families. The variants in or near PNPLA3, neurocan (NCAN, rs2228603), glucokinase regulator (GCKR; rs780094), protein phosphatase 1, regulatory subunit 3B (PPP1R3B, rs4240624) were significantly associated with NAFLD in African Americans while PNPLA3 and PPP1R3B were significantly associated with NAFLD in Hispanic Americans (72).

Table 1 (26-35) shows the GWAS studies in NAFLD. Rotman et al. analyzed data of 1,117 (894 adults and 223 children) individuals with histologically confirmed NAFLD. After adjustment for age, sex, diabetes mellitus (DM) and alcohol consumption, rs738409 C/G (PNPLA3-I148M) was associated with steatosis, portal inflammation, lobular inflammation, NAFLD activity score and fibrosis. Three SNPs on chromosome 10 near CHUK gene were independently associated with fibrosis. In children, no SNP was associated with histological severity (28). A meta-analysis of 16 studies showed that PNPLA3 rs738409 GG homozygous had 73% higher fat content when compared with CC polymorphism. Also, GG homozygous had greater risk of higher necroinflammatory scores and greater risk of developing fibrosis when compared with CC homozygous (3.24- and 3.2-fold respectively). NASH was more common in GG than CC homozygous [odds ratio (OR) 3.488; 95% CI, 1.859–6.545, data from 2,124 patients] (73).

Genetic variation in PNPLA3 is also associated with risk of HCC in NAFLD. Burza et al. showed for the first time, that PNPLA3 I148M allele is associated with risk of HCC. The authors compared bariatric surgery group to control group (conventional treatment for obesity) and found significantly higher incidence of HCC in the PNPLA3 I148M variant group in comparison to the control group (log-rank P value =0.001) (74). Liu et al. compared PNPLA3 rs738409 genotype in 100 European Caucasians with NAFLD related HCC to 275 controls with histological NAFLD. The multivariate analysis adjusted for age, gender, DM, BMI, and cirrhosis found that additive risk for HCC was 2.26 OR for each copy of gene. The GG homozygotes exhibited a 5-fold [1.47–17.29], P=0.01, increased risk over CC (75). Singal et al. performed a meta-analysis of 24 studies, which included 9,915 patients. PNPLA3 was associated with severity of fibrosis, OR being 1.32 (95% CI, 1.20–1.45). PNPLA3 increased fibrosis risk across all etiologies. Nine studies (n=2,937) showed increased risk of HCC in patients with cirrhosis (OR 1.40; 95% CI, 1.12–1.75), related to NASH or ALD, and not in other etiologies (76).

Transmembrane 6 superfamily 2 (TM6SF2) is a gene on chromosome 19p12, which functions as a lipid transporter. It is expressed predominantly in liver and intestine. The TM6SF2 encodes a protein with transmembrane domains. TM6SF2 is localized in the endoplasmic reticulum and ER-Golgi intermediate compartment of human liver cells. Functional studies were done in human hepatoma Huh7 and HepG2 cells, using siRNA inhibition and overexpression techniques. TM6SF2 siRNA inhibition was associated with reduced secretion of triglyceride rich lipoproteins and increased cellular triglyceride concentration and lipid droplet content, while TM6SF2 over expression reduced steatosis (77-79). TM6SF2 activity is required for normal

<table>
<thead>
<tr>
<th>Pathway</th>
<th>Genes, (reference)</th>
<th>Polymorphism</th>
<th>Results/comments</th>
</tr>
</thead>
<tbody>
<tr>
<td>Immune response</td>
<td>TNF (58,59)</td>
<td>TNF G238A (rs361525); G308A (rs1800629)</td>
<td>Susceptibility for insulin resistance, NAFLD and NASH</td>
</tr>
<tr>
<td></td>
<td>CD14 (60)</td>
<td>CD14 C (−159) T polymorphism</td>
<td>Patients with TT genotype had a 2.6-fold increased risk of developing NAFLD</td>
</tr>
<tr>
<td>Others</td>
<td>Cyclin-dependent kinase inhibitor 1A (also known as P21) (61)</td>
<td>CDKN1A (rs762623)</td>
<td>Associated with development but not progressive liver disease</td>
</tr>
<tr>
<td></td>
<td>KLF6 (62)</td>
<td>rs3750861, KLF6–IVS1, −27 G &gt; A</td>
<td>Associated with fibrosis</td>
</tr>
</tbody>
</table>

NAFLD, non-alcoholic fatty liver disease.
VLDL secretion and impaired TM6SF2 function causally contributes to NAFLD (32). The Glu167Lys missense mutation was shown to alter serum lipid profiles in humans and the knockdown of TM6SF2 in mice was shown to have increase liver triglyceride content and decreased VLDL secretion (79). Prill et al. studied in vitro disease model based on 3D spheroids from human hepatocytes and found that the TM6SF2E167K variant increased hepatocyte fat content by reducing APOB particle secretion (80).

As TM6SF2 controls hepatic lipid efflux, decreased effect by deletions or mutations reduces secretion of lipoprotein, causing hepatocellular triglyceride accumulation (81). TM6SF2 rs58542926 T-allele mediated hepatic retention of triglycerides and cholesterol predispose to NAFLD related fibrosis, whereas C-allele carriage promotes VLDL excretion, thus increasing risk of CVD or atherosclerosis while protecting the liver (82). In an exome-wide association study, Kozlitina et al. identified a TM6SF2 variant that conferred susceptibility to nonalcoholic fatty liver disease (34). Dongiovanni et al. confirmed that patients with TM6SF2 E167K variant have severe fatty liver disease (more steatosis, inflammation, ballooning and fibrosis) and lower circulating lipids (thus having reduced risk of myocardial infarction). E167K carriers had higher ALT and lower lipid levels (P<0.05), as well as a lower incidence of cardiovascular events (41). Liu et al. studied two histological NAFLD cohorts consisting of steatosis, steatohepatitis, fibrosis and cirrhosis (n=1,074). The authors found significant association of TM6SF2 to advanced fibrosis/cirrhosis. This association was independent of age, BMI, T2DM and PNPLA3 rs738409 genotype (42). In contrast to association of PNPLA3 to HCC; there is less data on TM6SF2 and HCC risk. A recent study by Yang et al. included 1,020 HCC, 2,021 controls with chronic liver disease and 2,484 healthy individuals in discover cohort and 249 alcoholic cirrhosis and 268 hepatitis C cirrhosis in prospective cohort. The authors found significant association of PNPLA3 and TM6SF2 to risk of HCC. PNPLA3 SNP was also significantly associated with HCC in non-fibrotic liver (OR =2.19; 95% CI, 1.22–3.92, P=0.007) (83).

**Other genetic modifiers in NAFLD**

The Table 2 shows some of the candidate gene studies in NAFLD. The genetic associations based on candidate gene studies can be either specific to NAFLD (related to lipid metabolism, insulin resistance and glucose metabolism) or non-specific related to inflammation, oxidative stress and fibrosis. ENPP1 (ectoenzyme nucleotide pyrophosphate phosphodiesterase 1) and IRS-1 (insulin receptor substrate-1) polymorphisms affect insulin sensitivity. Dongiovanni et al. compared 702 patients with biopsy-proven NAFLD and 310 controls. The ENPP1-121Gln and 972Arg-IRS-1 polymorphisms were independently associated with fibrosis in multivariate analysis. Both polymorphisms were associated with a marked reduction of AKT activation status, reflecting insulin resistance in obese patients with NAFLD (36). Glucokinase regulatory protein (GCKR) is associated with lipid and glucose metabolism and GCKR C > T SNP (OR 2.06) has been shown to be independently associated with significant liver fibrosis (37). A study observed that variants of solute carrier family 2 [(facilitated glucose transporter) member 1] (SLC2A1) were associated with NAFLD, and in vitro down-regulation SLC2A1 promoted lipid accumulation and oxidative stress (38). The transcription factor 7-like 2 (TCF7L2) polymorphism predisposes to diabetes by modulating beta-cell function and modulates lipid levels in familial dyslipidemia. Musso et al. showed that TCF7L2 polymorphism predisposed to NAFLD and significantly impacted liver injury, glucose homeostasis, postprandial lipoprotein and adipokine responses to fat ingestion (39). Membrane bound O-acyltransferase domain-containing 7 (MBOAT7), which is also known as LPIAT1, is a protein involved in the acyl chain remodeling of phospholipids. Recently, it has been shown that MBOAT7 is a multispanning transmembrane protein with six transmembrane domains (84). MBOAT7 catalyzes transfer of polyunsaturated fatty acids, thus maintaining fluidity of cell membranes. MBOAT7 is involved in the re-acylation of phospholipids. The rs641738 gene variant leads to a reduced MBOAT7 expression favoring increase in free arachidonic acid, which is a driver of hepatic inflammation (85). Mancina et al. showed that rs641738 C > T variant of MBOAT7 increased hepatic fat content, severity and fibrosis in comparison to subjects without the variant (44).

Ma et al. analyzed role of the SNP rs6834314 and its nearest gene, 17-beta hydroxysteroid dehydrogenase 13 (HSD17B13) in 768 adult Caucasians patients with NAFLD. The enzyme is lipid droplet-associated retinol dehydrogenase. The minor allele of rs6834314 was significantly associated with increased steatosis, but decreased severity (inflammation, ballooning, Mallory-Denck bodies) and liver enzyme levels (52).

The fatty acid transport protein 5 (FATP5) is involved in hepatic lipid and bile metabolism. Oxidative stress
plays an important role in pathogenesis of NASH and genes affecting oxidative stress have been shown to be associated with NASH. GCLC is involved in glutathione synthesis and variant is associated with steatohepatitis as compared to steatosis (55). Manganese-dependent superoxide dismutase (MnSOD) plays a role in protecting cells from oxidative stress. Al-Serri et al. showed that SOD2 C47T polymorphism was associated with more fibrosis in NASH (56). Uncoupling protein 2 (UCP2) is involved in mitochondrial lipid fluxes and reactive oxygen species production by the respiratory chain and it was observed that UCP2-866 A/A genotype was associated with increased hepatic UCP2 expression and reduced risk of NASH, particularly in subjects with normal glucose (57). Genes associated with inflammation also modify susceptibility of NAFLD/NASH. Valenti et al. analyzed 99 patients with NAFLD, 238 TNF-alpha polymorphism was higher in patients with NAFLD than in controls (31% vs. 15%; P=0.0001), and also these patients had higher insulin resistance indices (58).

Iron overload is thought to be associated with severe forms of NAFLD; however, not all studies have found an association with HFE gene mutations (86-88). Apolipoproteins are proteins that bind to lipids to form lipoproteins, thus helping in transport of lipids and also have affinity to some receptors. While some studies have found an association of Apolipoprotein C3 gene variants with NAFLD, other studies have not shown an association (89,90).

**Genetic basis in the pathogenesis of ALD**

As stated earlier, severity of liver disease differs among patients consuming similar type and quantity of alcohol over a similar duration. Even though alcoholic steatosis develops in majority, most patients consuming unsafe amount of alcohol escape from liver injury and do not develop significant liver disease (severe alcoholic hepatitis or cirrhosis). In fact, of all patients who consume alcohol to excessive amounts, cirrhosis develops in majority, most patients consuming unsafe amounts of alcohol escape from liver injury and do not develop significant liver disease (severe alcoholic hepatitis or cirrhosis) (97). Environmental factors other than heavy alcohol consumption which affect progression of ALD include gender (females are more susceptible), obesity and coexistence of viral hepatitis but are still not able to explain the significant inter-individual differences in ALD (3). Thus, genetic and epigenetic factors may play an important role in the pathogenesis of ALD.

In a study comparing 580 cases and 279 controls (defined as free of significant liver disease despite similar alcohol consumption), cases were significantly more likely to report death of father due to ALD (odds ratio 2.53) (93). Reed et al. analyzed medical records of 15,924 twin-pairs from twin registry. The monozygotic twins had higher concordance of alcoholism (26.7 vs. 12.2 for dizygotic) and cirrhosis in comparison to dizygotic twins (6.9 vs. 5.3 for dizygotic, P<0.001) (11). In a meta-analysis of 12 twin and five adoption studies, heritability of alcohol use disorder was almost 50% (9). Analysing the inter-ethnic differences, Stinson et al. reported higher risk for alcoholic cirrhosis related mortality in Hispanics in comparison to African Americans non-Hispanics and white non-Hispanics (10).

**Genetic modifiers in ALD**

Multiple studies have shown PNPLA3 gene polymorphisms to be associated with ALD. Tian et al. showed that rs738409 variant of PNPLA3 was strongly associated with ALD (unadjusted OR =2.25) (94). Stickel et al. studied impact of rs738409 gene variant on manifestation of ALD in two German cohorts that included 1,043 alcoholic patients with or without ALD and in 376 at-risk drinkers. The rs738409 genotype GG was strongly over represented in patients with cirrhosis (OR 2.79) and in alcoholic patients with elevated ALT levels (OR 2.33). The population attributable risk of cirrhosis in carriers of PNPLA3 rs738409 (G) was estimated to be 26.6% (95). Several studies have shown that PNPLA3 is a risk factor for HCC in ALD (96-98). A study observed that patients with cirrhosis and HCC were more likely to be G/G homozygotes, and this happened more commonly in patients with alcohol/metabolic cirrhosis as compared to viral cirrhosis (96). Friedrich et al. studied I148M polymorphism in 421 Caucasian patients. The G allele of the I148M variant was significantly more common in patients with ALD and HCC. Also, transplant free survival was lower in these patients (97).

Data regarding the association of TM6SF2 gene variant with ALD is less robust in comparison to NAFLD. Buch et al. performed a GWAS for alcohol related cirrhosis in 712 cases and 1,426 controls; the authors also validated results in two independent European cohorts (1,148 cases and 922 controls) and found that membrane bound O-acyltransferase domain containing 7 (MBOAT7), TM6SF2 and PNPLA3...
gene variants were associated with alcohol related cirrhosis. As all these genes are involved in lipid metabolism, it appears that lipid turnover is also important in the pathogenesis of alcohol related cirrhosis (99). Mancina et al. studied 416 at-risk alcohol drinkers retrospectively. The authors observed that PNPLA3, CD14 and TM6SF2 were associated with prevalence of alcoholic cirrhosis but only PNPLA3 and CD14 (and not TM6SF2) were associated with incidence of alcoholic cirrhosis (100). Further, in an Eastern European population, TM6SF2 rs58542926 and MBOAT7 rs641738 were not found to be related to alcohol related cirrhosis (101).

The other genetic polymorphisms shown to be associated with ALD are shown in Table 3 (94-110) and include alcohol metabolizing genes like alcohol dehydrogenase (ADH), aldehyde dehydrogenase (ALDH) and Cytochrome P450 2E1 (CYP2E1). ADH oxidizes alcohol to acetaldehyde, and acetaldehyde is further oxidized to acetate by ALDH. Both these steps require NAD+ as cofactor and both these reactions lead to a reduced NAD+/NADH ratio, which favors fatty acid synthesis and fat accumulation. CYP2E1 catalyzes ethanol oxidation to acetaldehyde, generating significant amount of reactive oxygen species, oxidative stress and inflammation. Formation of acetaldehyde and oxidative stress inhibit Peroxisome Proliferator Activated Receptor (PPAR) alpha transcriptional activity, decreasing fatty acid oxidation (85). The genetic variants of ADH, ALDH and CYP2E1 have also been shown to be associated alcoholic cirrhosis (102-106). The effect of CYP2E1 variants increases in the presence of other genetic variants involved in detoxification of reactive oxygen species (105,106). Genes involved in inflammation like cytotoxic T-lymphocyte associated protein 4, tumor necrosis factor and interleukin 1 beta are also associated with higher risk of cirrhosis (108-110).

Figure 1 shows genetic modifiers of NAFLD and ALD.

### Table 3 Showing gene association studies in ALD

<table>
<thead>
<tr>
<th>Pathway</th>
<th>Genes, (reference)</th>
<th>Polymorphism</th>
<th>Results/comments</th>
</tr>
</thead>
<tbody>
<tr>
<td>Alcohol, metabolism</td>
<td>ADH (102,103)</td>
<td>ADH2<em>1 and ADH3</em>2</td>
<td>More frequent in cirrhosis</td>
</tr>
<tr>
<td></td>
<td>ALDH (104)</td>
<td>ALDH2<em>1; ALDH2</em>1/1</td>
<td>More frequent in cirrhosis</td>
</tr>
<tr>
<td></td>
<td>Cytochrome P450 (105,106)</td>
<td>CYP2E1 5B; CYP2E1*c2</td>
<td>Interaction with other genes which are involved in detoxification of reactive oxygen species (GSTM1, GABRG2) increased risk in cirrhosis</td>
</tr>
<tr>
<td>Lipid metabolism</td>
<td>PNPLA3 (94-99)</td>
<td>PNPLA3 rs738409 C/G (amino acid change I148M)</td>
<td>Severity of disease and HCC</td>
</tr>
<tr>
<td></td>
<td>TM6SF2 (99-101)</td>
<td>TM6SF2 rs10401969 (C); rs58542926 (C/T) E167K</td>
<td>Severity of disease and HCC, not all studies have shown an association</td>
</tr>
<tr>
<td></td>
<td>MBOAT7 (99)</td>
<td>rs626283 (C)</td>
<td>Associated with cirrhosis</td>
</tr>
<tr>
<td>Oxidative stress</td>
<td>SOD2 (107)</td>
<td>SOD2 A16V (rs4880)</td>
<td>Risk factor for severe alcoholic disease</td>
</tr>
<tr>
<td>Immune response</td>
<td>TNF (108)</td>
<td>TNF G238A (rs361525); G308A (rs1800629)</td>
<td>Susceptibility for alcoholic cirrhosis</td>
</tr>
<tr>
<td></td>
<td>IL1B</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>CTLA4 (110)</td>
<td>CTLA4 G/G</td>
<td>Associated with cirrhosis</td>
</tr>
</tbody>
</table>

ALD, alcoholic liver disease.

Epigenetics include a process that alters gene activity without changing the DNA sequence. These changes can be transmitted to daughter cells by cell division. Epigenetic modifications are caused by alterations in DNA methylation, modifications in histone proteins and by micro RNAs (miR) (111-113). Epigenetic changes have been shown to affect both NAFLD (114) and ALD (85); however, knowledge and data regarding epigenic changes is limited at present. miRNAs are transcribed in cell nucleus and transported to the cytoplasm, where they are processed into mature miRNAs. miRNAs are 19-22 nucleotide non-coding sequences that bind to the complementary sequence of messenger RNA molecules and regulate gene expression.
by silencing or inhibition of translation. miRNAs play a role in many cellular processes. miRNA dysregulation has been shown to be associated with several liver diseases including ALD, NAFLD, viral hepatitis, fibrosis and HCC (111,112). miRNAs remain stable in the circulation (blood and urine) and thus represent potential biomarkers for diseases (115). Following miRNAs are shown to be upregulated in ALD (in humans or animal models); miR-155, miR-34a, miR-212, miR-21, miR-181a, miR-217, miR-223. Following miRNAs are downregulated in ALD; miR-122, miR-29, miR-199a, miR-125b, miR-126, miR-200a, miR-375. These miRNAs are involved in increased intestinal permeability, liver injury, inflammation, steatosis, oxidative stress, fibrosis, cirrhosis and HCC, in humans or animal models (85,116-118). Data in NAFLD patients suggests that following miRNAs are upregulated in NAFLD: miR-21, miR-34a, miR-182 whereas miR-122 is downregulated (114). Other epigenetic mechanisms which affect NAFLD and ALD include DNA methylation. In a study of DNA methylation of 4 CpG islands (CpG99, CpG71, CpG26 and CpG101) in regulatory regions of PNPLA3, SAMM50, PARVB variant 1, and PARVB variant 2, Kitamoto et al. showed hypomethylation of CpG26 (PARVB variant 1) and hypermethylation of CpG99 in the regulatory region of PNPLA3 which was associated with fibrosis in NAFLD (119). Cordero et al. showed that liver fat accumulation induced by a high-fat-sucrose diet in male Wistar rats was prevented by methyl donor supplementation (120). Zeybel et al. showed that differential DNA methylation at specific CpGs within genes affecting fibrogenesis distinguished mild from severe fibrosis in both NAFLD and ALD (121). Hardy et al. compared biopsy proven NAFLD patients with controls; differential DNA methylation at PPARγ promoter was detected within the pool of cell-free DNA of plasma. Similar changes were present in patients with alcoholic cirrhosis (122).

Figure 1 Venn diagram showing genetic factors associated with non-alcoholic fatty liver disease and alcoholic liver disease. ADH, alcohol dehydrogenase; ALDH, Aldehyde dehydrogenases; Cyclin-dependent kinase inhibitor 1A, CTLA4: cytotoxic T-lymphocyte associated protein 4, ENPP1, ectoenzyme nucleotide pyrophosphate phosphodiesterase 1; FADS1, fatty acid desaturase 1; GCKR, glucokinase regulatory protein; GCLC, Glutamate-Cysteine Ligase Catalytic Subunit; HSD17B13, 17-beta hydroxysteroid dehydrogenase 13; KLF6, Kruppel-like factor; LPIN1, Lipin 1; MBOAT7, membrane-bound O-acyltransferase domain-containing 7; PNPLA3, patatin-like phospholipase domain-containing protein 3; PPAR alpha, peroxisome proliferator-activated receptor alpha; PEMT, phosphatidylethanolamine N-methyltransferase; PXR, pregnane X receptor; SLC2A1, solute carrier family 2 member 1; SOD2, superoxide dismutase 2; TM6SF2, transmembrane 6 superfamily 2; TNF, tumor necrosis factor.
expression (123). The aberrant histone modifications promote development of insulin resistance and diabetes mellitus (124). cAMP responsive element-binding protein (CREBH) is a hepatocyte specific transcription factor, which is localized in the endoplasmic reticulum (ER) membrane. CREBH is activated by ER stress or inflammation, which enters into the cell nucleus and activates expression of genes involved in acute-phase response, gluconeogenesis, lipogenesis, fatty acid oxidation, and lipolysis. Thus, modulation of CREBH acetylation can lead to altered lipid homeostasis associated with NAFLD (123,125). The activation of deacetylase sirtuin-1 has potential against the physiological mechanisms related to NAFLD (126). Alcohol also has been shown to affect acetylation and phosphorylation of histones in rat hepatocytes (127).

In an example of epigenetic change transmitted to offspring, Bruce et al. showed that maternal fat intake during gestation in female mice contributed to development of NAFLD in offspring (128). Female mice were fed with either a high-fat (HF) or control chow (C) diet before and during gestation, and during lactation. The offspring were fed either a C or a HF diet after weaning, thus generating four offspring groups: HF/HF, HF/C, C/HF, C/C. The liver histology was normal in C/C and HF/C offspring at 15 weeks. The C/HF offspring developed NAFL while HF/HF offspring developed NASH. Histological analysis at 30 weeks showed NAFLD in HF/C and C/HF groups, whereas the HF/HF had a more severe form of NASH. Thus, exposure to HF diet in utero and during lactation contributed to severe form of disease. Hepatic mitochondrial electron transport chain enzyme complex activity was reduced in offspring from HF-fed mothers. Also, lipogenesis, oxidative stress, and inflammatory pathways were up regulated at 15 weeks in offsprings of HF mothers (128).

**Clinical implications of genetic and epigenetic modifiers in NAFLD and ALD**

Genetic information can change occurrence of a disease in several ways. First, it may lead to behaviour change of a subject, thus decreasing risk of disease. Second, disease can be diagnosed timely or can be prevented by timely intervention. Third, pharmacotherapy can be tailored as per genetic information. As fibrosis progression in both NAFLD and ALD spans many years, a timely diagnosis/intervention of at risk subjects can prevent development of cirrhosis. As PNPLA3 is associated with higher risk of HCC, the knowledge of genetic modifiers may also help in identifying patients who would need strict HCC surveillance (75). A lower threshold may be kept in patients with TM6SF2 variants for cardiovascular disease screening (82). Genetic and epigenetic modifiers may alter response to pharmacotherapy and thus may help in modifying the treatment options. Dongiovanni et al. looked at response to statins in 107 patients with NAFLD. Use of statins was significantly associated with protection from steatosis, NASH, and fibrosis, but this effect was stronger in patients without I148M PNPLA3 risk variant. Thus, I148M PNPLA3 variant limited efficacy of statins (129). Scorletti et al. studied 103 patients with NAFLD, randomized to omega-3 fatty acids (DHA + EPA) or placebo for 15–18 months. Fifty-five men and 40 women completed the study. PNPLA3 148M/M variant influenced the changes in liver fat and DHA tissue enrichment (130). Patients with PNPLA3 polymorphism have been shown to respond better to life style modifications (131). Since pathogenesis of NAFLD is complex and is affected by several environmental and genetic factors, a single genetic variant is unlikely to have very strong role in risk prediction; however, a combination score of several genetic variants may provide useful insight into disease progression in future. Whether genetic information can modify disease risk of NAFLD or ALD, is yet to be seen. How genetic information can modify diet was shown by Nielsen et al. The authors conducted a double-blinded randomized controlled trial to examine the short- and long-term effects of genetic information on nutrition and dietary intake of caffeine, vitamin C, added sugars and sodium. The intervention group (n=92) received genetic information based dietary advice and the control group (n=46) received general dietary recommendations without genetic information for 12-month. At 12-month, the participants with a risk type of ACE gene in the intervention group significantly reduced sodium intake compared to the control group (132). The genetic information about risk of future development of NAFLD or ALD should lead to change of risk behavior and decrease of modifiable risk factors like diet, exercise, avoidance of weight gain or weight loss (if obese), decrease in amount/frequency of alcohol consumption.

**Conclusions**

A significant variability in disease severity of NAFLD and ALD occur due to genetic modifiers. Epigenetics changes may explain the phenotypic variability in patients with
similar gene polymorphism. Knowledge of genetic and epigenetic modifiers should lead to development of new therapeutic targets and more selected therapy for patients with NAFLD and ALD. Also, better risk prediction regarding development of cirrhosis and HCC should be feasible in future with better understanding of these factors.

Acknowledgments

Mr. Yogesh Saini (Research Coordinator).

Funding: None.

Footnote

Provenance and Peer Review: This article was commissioned by the Guest Editor (Ashwani K. Singal) for the series “Non-alcoholic Fatty Liver Disease and Alcoholic Liver Disease” published in Translational Gastroenterology and Hepatology. The article was sent for external peer review organized by the Guest Editor and the editorial office.

Conflicts of Interest: Both authors have completed the ICMJE uniform disclosure form (available at http://dx.doi.org/10.21037/tgh.2019.09.06). The series “Non-alcoholic Fatty Liver Disease and Alcoholic Liver Disease” was commissioned by the editorial office without any funding or sponsorship. The authors have no other conflicts of interest to declare.

Ethical Statement: The authors are accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved.

Open Access Statement: This is an Open Access article distributed in accordance with the Creative Commons Attribution-NonCommercial-NoDerivs 4.0 International License (CC BY-NC-ND 4.0), which permits the non-commercial replication and distribution of the article with the strict proviso that no changes or edits are made and the original work is properly cited (including links to both the formal publication through the relevant DOI and the license). See: https://creativecommons.org/licenses/by-nc-nd/4.0/.

References

14. Struben VM, Hespenheide EE, Caldwell SH.


doi: 10.21037/tgh.2019.09.06

Cite this article as: Choudhary NS, Duseja A. Genetic and epigenetic disease modifiers: non-alcoholic fatty liver disease (NAFLD) and alcoholic liver disease (ALD). Transl Gastroenterol Hepatol 2021;6:2.