THE AMERICAN
JOURNAL OF
Gastroenterology

# Genetic variants in PNPLA3 and TM6SF2 predispose to the development of hepatocellular carcinoma in individuals with alcohol-related cirrhosis

SCHOLARONE™
Manuscripts

**Page 1 of 32**                    **American Journal of Gastroenterology**

Stickel et al.                                    Genetic risk of hepatocellular carcinoma

# Genetic variants in *PNPLA3* and *TM6SF2* predispose to the development of hepatocellular carcinoma in individuals with alcohol-related cirrhosis

**Short title:** Genetic risk for hepatocellular carcinoma

Felix Stickel[1*], Stephan Buch[2*], Hans Dieter Nischalke[3], Karl Heinz Weiss[4], Daniel Gotthardt[4], Janett Fischer[5], Jonas Rosendahl[6], Astrid Marot[7], Mona Elamly[7], Markus Casper[8], Frank Lammert[8], Andrew McQuillin[9], Steffen Zopf[10], Ulrich Spengler[3], Silke Marhenke[11], Martha Kirstein[11], Arndt Vogel[11], Florian Eyer[12], Johann von Felden[13], Henning Wege[13], Thorsten Buch[14], Clemens Schafmayer[15], Felix Braun[15], Pierre Deltenre[7ǂ], Thomas Berg[5ǂ], Marsha Y. Morgan[16ǂ], Jochen Hampe[2ǂ]

*FS and SB have contributed equally to the presented work and share premier authorship

ǂPD, TB, MYM and JH have contributed equally to the presented work and share senior authorship

[1]Department of Gastroenterology and Hepatology, University Hospital of Zurich, Switzerland

[2]Medical Department 1, University Hospital Dresden, TU Dresden, Germany

[3]Department of Internal Medicine I, University of Bonn, Germany

[4]Department of Internal Medicine IV, University Hospital Heidelberg, Heidelberg, Germany

[5]Hepatology Section, Division of Gastroenterology and Rheumatology, University Hospital Leipzig, Leipzig, Germany

[6]Department of Gastroenterology, University Hospital Halle/Saale, Germany

[7]Division of Gastroenterology and Hepatology, Centre Hospitalier Universitaire Vaudois, University of Lausanne, Switzerland

[8]Department of Medicine II, Saarland University Medical Center, Homburg, Germany

[9]Molecular Psychiatry Laboratory, Division of Psychiatry, University College London, UK

[10]Medical Department 1, University of Erlangen-Nuremberg, Germany

[11]Department of Gastroenterology, Hepatology and Endocrinology, Hannover Medical School, Germany

[12]Department of Clinical Toxicology, Klinikum rechts der Isar, Technical University of Munich, Germany

[13]First Department of Medicine, University Medical Center Hamburg-Eppendorf, Hamburg

[14]Institute of Laboratory Animal Science, University of Zurich, Schlieren, Switzerland

[15]Department of General and Thoracic Surgery, University Hospital Schleswig-Holstein, Campus Kiel, Germany

[16]UCL Institute for Liver & Digestive Health, Division of Medicine, Royal Free Campus, University College London, United Kingdom

1

**American Journal of Gastroenterology** **Page 2 of 32**

Stickel et al. Genetic risk of hepatocellular carcinoma

**Key words:** candidate genes, genotyping, malignant transformation, host genetics, lipotoxicity, single nucleotide polymorphism

**Correspondence**

Felix Stickel, MD

Department of Gastroenterology and Hepatology, University Hospital Zürich,

Rämistrasse 100, CH-8091 Zürich, Switzerland

Phone +41-31-338 7817

Fax +41-31-335 3519

Felix.stickel@uzh.ch

**Abbreviations:**

| | | |
|---|---|---|
| AUDIT | - | Alcohol Use Disorders Identification Test |
| BMI | - | body mass index |
| ELISA | - | enzyme-linked sorbent assay |
| GWAS | - | genome-wide association study |
| HCC | - | hepatocellular carcinoma |
| MAF | - | minor allele frequency |
| *MBOAT7* | - | membrane bound O-acyltransferase domain containing 7 |
| NAFLD | - | non-alcoholic fatty liver disease |
| *NCAN* | - | neurocan |
| NPV | - | negative predictive value |
| OR | - | odds ratio |
| PAF | - | population-attributable fraction |
| *PNPLA3* | - | patatin-like phospholipase domain-containing 3 |
| SNP | - | single nucleotide polymorphism |
| *TM6SF2* | - | transmembrane 6 superfamily member 2 |

2

Stickel et al.                                                  Genetic risk of hepatocellular carcinoma

**Funding**

**STUDY HIGHLIGHTS**

*What is current knowledge?*

- The gene variant *PNPLA3* rs738409 is an established risk factor for the development of alcohol-related cirrhosis and for the subsequent development of hepatocellular carcinoma (HCC) within this population

- Two other risk variants, *TM6SF2* rs58542926 and *MBOAT7* rs641738, are associated with the development of alcohol-related cirrhosis but their role in the subsequent development of HCC is unclear

*What is new here?*

- While *PNPLA3* rs738409 is confirmed as a risk factor for HCC in alcoholic cirrhotics, the present study identifies *TM6SF2* rs58542926 as an independent risk factor for the development of HCC in people with alcohol-related cirrhosis

- The dual carriage of these risk variants in *PNPLA3* and *TM6SF2* accounts for almost 50% of the attributable risk for the development of HCC in this population

- In contrast to its role as a risk factor for alcoholic cirrhosis, *MBOAT7* rs641738 does not seem to confer risk to HCC development in this risk group

- Genotype-based risk stratification for patients with alcohol-related cirrhosis may guide the development of rational, cost-effective surveillance programmes to facilitate early HCC detection

4

Stickel et al.                                             Genetic risk of hepatocellular carcinoma

## ABSTRACT

**Objectives:** Variants in patatin-like phospholipase domain-containing 3 (*PNPLA3;* rs738409), transmembrane 6 superfamily member 2 (*TM6SF2;* rs58542926) and membrane bound O-acyltransferase domain containing 7 (*MBOAT7;* rs641738) are risk factors for the development of alcohol-related cirrhosis. Within this population, *PNPLA3* rs738409 is also an established risk factor for the development of hepatocellular carcinoma (HCC). The aim of this study was to explore possible risk associations of *TM6SF2* rs58542926 and *MBOAT7* rs641738 with HCC.

**Methods:** Risk variants in *PNPLA3, TM6SF2* and *MBOAT7* were genotyped in 751 cases with alcohol-related cirrhosis and HCC and in 1165 controls with alcohol-related cirrhosis without HCC. Association with the risk of developing HCC was analyzed using multivariate logistic regression.

**Results:** The development of HCC was independently associated with *PNPLA3* rs738409 ($OR_{adjusted}$ 1.84 [95% CI 1.55-2.18], $p=1.85\times10^{-12}$) and *TM6SF2* rs58542926 ($OR_{adjusted}$ 1.66 [1.30-2.13], $p=5.13\times10^{-05}$) using an additive model and after controlling for sex, age, body mass index and type 2 diabetes mellitus; the risk associated with carriage of *MBOAT7* rs641738 ($OR_{adjusted}$ 1.04 [0.88-1.24], $p=0.61$) was not significant. The population-attributable fractions were 43.5% for *PNPLA3* rs738409, 11.5% for *TM6SF2* rs58542926, and 49.9% for carriage of both variants combined.

**Conclusions:** Carriage of *TM6SF2* rs58542926 is an additional risk factor for the development of HCC in people with alcohol–related cirrhosis. Carriage of both *PNPLA3* rs738409 and *TM6SF2* rs58542926 accounts for half of the attributable risk for HCC in this population. Genotyping will allow for more precise HCC risk stratification of patients with alcohol-related cirrhosis, and genotype-guided screening algorithms would optimize patient care.

5

Stickel et al.                                                                     Genetic risk of hepatocellular carcinoma

## INTRODUCTION

The incidence of hepatocellular carcinoma (HCC) is increasing and it is now the fifth most frequent cancer and third most frequent cause of cancer-related mortality world-wide [1-3]. In Western countries, the presence of cirrhosis is a major risk factor for the development of HCC which only rarely evolves in non-cirrhotic livers [1,4,5]. Persistent high alcohol consumption is the leading cause of cirrhosis in these areas. The prognosis of patients with HCC is grave as it is often detected only when it reaches an advanced stage.

Between 5 to 15% of patients with alcohol-related cirrhosis are potentially at risk for developing HCC with a 5-year cumulative risk of around 8% [1]. Identifying disease-modulating factors which render some individuals more susceptible to its development than others is clearly of major importance as it would allow more specific targeting of surveillance programmes. Older age, male sex, and more advanced liver disease at presentation, are known to be important predictors but do not account for more than a small proportion of the observed risk variance. Mounting evidence suggest that host genetic factors may significantly modify the risk for developing HCC. Thus, identifying these factors in patients with alcohol-related cirrhosis is clearly of major importance.

A single nucleotide polymorphism (SNP) rs738409 C>G in patatin-like phospholipase domain-containing 3 (*PNPLA3)* is an established risk factor for the development of alcohol-related cirrhosis in Caucasian populations [6-9]. A recent meta-analysis of individual patient data from candidate gene studies has shown that the *PNPLA3* rs738409 G allele is also an independent risk factor for the development of HCC in this patient population [10,11]. Recently, variants in two further genes *viz.* rs58542926 in transmembrane 6 superfamily member 2 (*TM6SF2)* and rs641738 in membrane bound O-acyltransferase domain containing 7 (*MBOAT7)* were identified as risk factors for the development of alcohol-related cirrhosis in a genome-wide

6

association study (GWAS) [12]. Although a proportion of the cases included in the GWAS had developed HCC, the numbers were too small to allow for a meaningful subset analysis. However, data from two candidate gene studies support a role for variants in or linked to *TM6SF2* as potential risk factors for HCC development. Thus, Nischalke et al. found a significant association between rs2228603 in the gene coding for neurocan (*NCAN*) and the development of HCC in people with alcohol-related cirrhosis [13]; this variant is in close linkage disequilibrium with *TM6SF2* rs58542926. Falleti et al. genotyped *TM6SF2* rs58542926 in 226 cases with alcohol-related cirrhosis of whom 75 (33%) had HCC [14]. The rs58542926 T/*genotype was more frequent in those cases with HCC (adjusted odd ratio [OR] 2.57; p=0.035). They found no interaction between *PNPLA3* rs738409 and *TM6SF2* rs58542926 in relation to HCC risk. These preliminary findings need to be further explored. No studies have been undertaken, to date, to explore the possible risk association between *MBOAT* rs641738 and the development of HCC in patients with alcohol-related cirrhosis.

The aim of the present study was to determine whether variants in *PNPLA3, TM6SF2* and *MBOAT7* are risk factors for the development of HCC, alone or in synergy, in a large, well-characterized population of people with alcohol-related cirrhosis.

## METHODS

### Patient cohorts

For purposes of this study *cases* were defined as people with alcohol-related cirrhosis and HCC and *controls* as people with alcohol-related cirrhosis but without evidence of HCC. Cases and controls were recruited from Hepatology centres across Europe (Switzerland, Germany, United Kingdom) and were of self-reported Swiss German/German/British/Irish ancestry.

The diagnosis of alcohol-related cirrhosis was established as described previously [12], and based on a history of prolonged, sustained alcohol intake of a minimum of 40g/day, if female, and 60g/day, if male, and histological examination of liver tissue; or compatible historical, clinical, laboratory, radiological and endoscopic features. Patients were excluded if they had any other potential cause of liver injury, although a proportion of the participants were overweight. The diagnosis of HCC was based on histological examination of tumour tissue or evidence on imaging, preferably using two modalities, of lesion that are hypervascular in the arterial phase with washout in the portal venous or delayed phases [15]. The severity of underlying alcohol-related liver injury was assessed using Pugh's modification of the Child's grading system [16].

### DNA preparation and genotyping

DNA was prepared using the FlexiGene chemistry (Qiagen, Hilden, Germany). For preparation of genotyping, DNA samples were evaluated by gel electrophoresis and adjusted to 20-30 ng/µl DNA content using the Picogreen fluorescent dye (Molecular Probes – Invitrogen, Carlsbad, Ca, USA) on a robotic platform using TECAN liquid handling equipment. One microliter of genomic DNA was amplified with the GenomiPhi (Amersham, Uppsala, Sweden) whole genome amplification kit and fragmented at 99°C for three minutes. Five ng of DNA

8

were dried overnight in TwinTec hardshell 384well plates (Eppendorf, Hamburg, Germany) at room temperature. Genotyping of *PNPLA3* rs738409 (Assay ID HCV7241_10 ), *TM6SF2* rs58542926 (Assay ID HCV89463510_10), and *MBOAT7* rs641738 (Assay ID HCV8716820_10) was performed using TaqMan® SNP Genotyping Assays and chemistries (Applied Biosystems, Foster City, CA, USA) on an automated platform with TECAN Freedom EVO and 384well TEMO liquid handling robots (TECAN, Männedorf, Switzerland) as described before.[12]

All process data were logged and administered with a database-driven LIMS. Reactions were completed and read in a 7900 HT TaqMan sequence detector system (Applied Biosystems, Foster City, CA, USA). The amplification reaction was carried out with the TaqMan universal master mix. Thermal cycling conditions consisted of 1 cycle for 10 minutes at 95°C, followed by 45 cycles for 15 seconds at 95°C, and 45 cycles for 1 minute at 60°C.

**Statistical analysis**

Statistical analyses of data were performed using SPSS v.24.0 (IBM Corp., Armonk, NY, USA). Comparisons of the frequencies of the genotypes at the three loci were performed on 3x2 contingency tables between cases and controls. Genotypic and allelic tests of association were performed using $\chi^2$ statistics or Fisher's exact tests as appropriate. For all tests, nominal p-values are reported. The association between SNPs and the risk of HCC was estimated by calculating beta coefficients and their standard errors using logistic regression models adjusted for age, sex, BMI and diabetic status, assuming log-additive genetic effects. The interactions between polymorphic loci were examined by testing allelic and genotypic interaction terms in the logistic regression models. Results are expressed as odds ratios (OR) and 95% confidence intervals (CI).

Stickel et al.                                                    Genetic risk of hepatocellular carcinoma

*Population Attributable Fraction*

The population-attributable fraction (PAF) provides an epidemiological estimate of the proportion of the disorder that is attributable to the risk factor. Thus, it is an estimate of how much lower the frequency of HCC would be in the patients with alcohol-related cirrhosis, if the risk genotype(s) were eliminated from the population.

The *PAF* was estimated for heterozygous and homozygous carriage using the formula:

$$(x - 1)/x$$

*where x = (1- p)2 + 2p(1- p)OR1 + p2OR2; p is the allele frequency in alcohol-related cirrhosis, and OR1 and OR2 are the ORs associated with hetero- and homozygosity adjusted by sex, age, body mass index (BMI) and type 2 diabetes mellitus, respectively.[17]*

Combined PAF estimates were calculated based on the individual PAFs for each associated SNP, assuming no multiplicative interaction amongst the SNPs.[17]

**Ethics**

The study was approved by the ethics committees of the participating institutions; all included subjects provided written informed consent prior to inclusion into the study.

Stickel et al.                                          Genetic risk of hepatocellular carcinoma

### RESULTS

The cases were significantly older than the controls; proportionately more were male, overweight and had comorbid type 2 diabetes mellitus (**Table 1**). However, the cases were significantly less likely to have decompensated liver disease (**Table 1**).

Genotyping was completed for all three SNPs with call rates greater than 95%; there were significant differences in the genotyping call rates between cases and controls but the distributions of the missing data were random and not skewed. All markers followed Hardy-Weinberg equilibrium (cutoff HWE $P > 0.05$) in both the cases and controls (**Table 2**).

Allelic and genotypic associations of *PNPLA3* rs738409 with HCC in alcohol-related cirrhosis were highly significant ($P_{allelic}= 3.64 \times 10^{-20}$, $P_{geno}= 6.14 \times 10^{-18}$). The odds ratio (OR) for homozygous carriers of the risk allele *PNPLA3* rs738409 G was 3.26 (95% CI: 2.50-4.25) compared to the risk in heterozygous carriers of 1.70 (95% CI: 1.37-2.10). The allelic and genotypic associations of *TM6SF2* rs58542926 with HCC in alcohol-related cirrhosis were also highly significant at $P_{allelic}= 8.93 \times 10^{-7}$ and $P_{genotypic}= 7.74371 \times 10^{-6}$, respectively. The OR for homozygous carriage of the risk allele *TM6SF2* rs58542926 T was 3.32 (95% CI: 1.76-6.24), while the OR for heterozygous carriage was 1.46 (95% CI: 1.15-1.84). In contrast, neither heterozygous nor homozygous carriage of the *MBOAT7* rs641738 T allele was associated with HCC risk; the allelic and genotypic levels of significance were $P_{allelic}=0.639$ and $P_{genotypic}=0.676$ respectively while corresponding ORs for homozygous and heterozygous carriage of the T allele were 1.07 (95% CI: 0.83-1.38) and 0.96 (95% CI: 0.77-1.20).

The minor allele frequencies (MAF) of all three loci in the controls in the present study were in line with those observed in the previous GWAS in alcohol-related cirrhosis [12], but the frequencies of *PNPLA3* rs738409 G and *TM6SF2* rs58542926 T were higher in the cases. The MAF for *MBOAT7* rs641738 T was similar in cases and controls.

11

Stickel et al. Genetic risk of hepatocellular carcinoma

The robustness of the association of *PNPLA3* rs738409 and *TM6SF2* rs58542926 and the development of HCC was confirmed after adjustment for other known risk factors viz. age, sex, BMI and type 2 diabetes mellitus (**Table 3**). The lack of association with *MBOAT7* rs641738 did not change following adjustment for these confounders.

The association of *PNPLA3* rs738409 and *TM6SF2* rs58542926 with HCC was consistent across the separate cohorts from Switzerland, Germany, Britain and Ireland, as was the lack of association with regard to *MBOAT7* rs641738 (**Supplementary Figures 1A-C**).

The PAF% for *PNPLA3* rs738409 was 43.5% and for *TM6SF2* rs58542926 11.4% yielding a combined PAF% of both risk loci of 49.9% (**Table 4**).

The proportion of cases and controls who carried either no, one out of two, or both risk variants in *PNPLA3* and *TM6SF2* illustrated the aggregated risk due to co-carriage of risk loci (**Figure 1)**. Thus, only 10% of the controls but approximately 20% of the cases carried both risk alleles.

12

**DISCUSSION**

Elucidating the genetic background of chronic liver diseases is important as it will enable an individual's risk to develop progressive liver damage to be assessed more precisely, and will allow greater insight into currently unknown molecular aspects of the underlying pathophysiology. Of particular importance is the definition of the risk factors predisposing an individual with an *a priori* cancer risk, i.e. cirrhosis, to the development of life-threatening malignancy particularly if defining the risk factors would allow development of a strategy for more individualized screening and targeted prevention.

The present study provides data on the largest cohort of patients with HCC on a background of alcohol-related cirrhosis undertaken to date and provides further confirmation that the *PNPLA3* rs738409 G allele is the strongest independent genetic risk factor for HCC in this patient population [10, 11]. The present analysis further expands the panel of genetic risk factors for HCC to include *TM6SF2* rs58542926, albeit with lesser effect as reflected by its lower OR and PAF%. It also showed a lack of association with *MBOAT7* rs641738. There is only limited information on the possible risks associated with carriage of *TM6SF2* rs58542926 and *MBOAT7* rs641738 and the development of HCC with which to compare the findings of the present study.

Nischalke et al. in a relatively small, multicentre study found a significant risk association between carriage of rs2228603 in *neurocan* (*NCAN*) and the development of HCC in patients with alcohol-related cirrhosis [13]. This genetic variant had been linked the development of hepatic steatosis in a previous GWAS in nonalcoholic fatty liver disease (NAFLD) [18]. Subsequent genetic fine mapping showed that *NCAN* rs2228603 lies within 50 kb and is in strong linkage disequilibrium with *TM6SF2* rs58542926 [19]. A subsequent exome-wide

13

analysis conclusively identified *TM6SF2* rs58542926 as the more likely risk factor for progressive NAFLD in this genomic region [20]. Falleti *et al.* [14] analyzed a cohort of 511 patients with cirrhosis of mixed aetiology, of whom 101 had HCC, and found an association between the risk of developing HCC and carriage of the *TM6SF2* rs58542926 T allele in the 226 patients with alcohol-related cirrhosis (Adjusted OR=2.57; p=0.035). They found no interaction between *PNPLA3* rs738409 and *TM6SF2* rs58542926 in relation to HCC risk.

A recent candidate gene association study from Italy and the UK investigated the association between variants within *MBOAT7* and HCC in a mixed cohort of patients with NAFLD, chronic hepatitis C and alcohol-related liver disease the majority of whom did not have cirrhosis [21]. The authors reported a significant association between *MBOAT7* rs641738 and HCC in non-cirrhotic patients with NAFLD, but only in the Italian cohort. The findings were not however replicated in a validation cohort of UK patients. The reasons for this discrepancy likely relates to the fact that development of HCC in people with non-cirrhotic alcohol- and viral-related liver disease in Western countries is very rare and hence the number of cases was small. Thus, the inability of these workers to replicate their findings may reflect the inconsistent of the chosen phenotypes studied and/or low statistical power.

The lack of association between HCC risk and *MBOAT*rs641738 is further supported by a recent candidate gene study in people with chronic hepatitis C which demonstrated that the influence of this variant is restricted to the transition from absence of fibrosis to mild fibrosis [22]. In this same study, no difference was observed in the MAF of *MBOAT7*rs641738 in a small subgroup of 75 people with HCV-related HCC within an entire cohort of 1,706 indicating a lack of statistical power. However, the present study was adequately powered to detect an allelic odds ratio of 1.18 with 80% power and hence the finding of no association between

*MBOAT7* rs641738 and the development of HCC in alcohol–related cirrhosis is robust [23] and was consistent across all study subpopulation (**Supplementary Figure 1C**).

In the present study carriage of risk alleles in *PNPLA3* and *TM6SF2* accounted for almost 50% of the population-attributable genetic risk for the development of HCC in patients with alcohol-related cirrhosis. The combined PAF% is usually lower than the sum of each individual PAR% but not in this instance. The odds ratio was, in fact, greater than 3 for homozygous carriage and ~1.5 or heterozygous carriage for both variants, and hence surprisingly high for a germline genetic risk factor in a complex disorder.

The possibility that these genetic markers could be included in a stratification score to identify patients at high risk for HCC has to be considered. Hypothetically, screening intervals could be tailored according to the genetic signature in a given patient, applying longer intervals in those minimal inherited predisposition, and shorter periods for those who carry both risk alleles.  However, any such proposals would need to be tested in prospective surveillance studies designed to assess possible patient benefit.

Anstee *et al*. [24] have recently looked at the clinical utility of *PNPLA3* genotyping for stratifying the risk for developing HCC in patients with NAFLD in a reanalysis of the data published by Liu *et al.* [25].  They found that use of *PNPLA3* genotyping alone to *positivel*y predict the risk of developing HCC was unlikely to be tenable. However, the *negative predictive value (NPV)* was substantially greater, suggesting that genotyping this SNP might have utility for selecting those individuals least likely to develop HCC and therefore least likely to benefit from surveillance.

Using a similar approach the diagnostic test performance for *PNPLA3* and *TM6SF2* genotyping for the detection of HCC risk in the present study cohort showed a NPV of 72.5% for the combination *PNPLA3* rs738409 GG and *TM6SF2* rs58542926 \*T. Based on these figures, a

15

schema for individual screening intervals is proposed with the caveat that any adoption must be subject to appropriate prospective evaluation (**Table 5**).

The distinct genetic profiles of the alcohol-related cirrhosis risk genes with regards to liver carcinogenesis also have intriguing functional implications. The *PNPLA3* rs738409 GG variant is considered a loss-of-function mutation and as such is less able to hydrolyze triglycerides from lipid droplets leading to lipid trapping in hepatocytes and increased lipotoxicity. Since progression of fibrosis is only weakly associated with the degree of steatosis [26], additional effects beyond mere lipid accumulation are suspected. Thus, *PNPLA3* rs738409 has been functionally linked to increased levels of circulating pro-inflammatory mediator intracellular adhesion molecule 1 (ICAM-1) [27], which is also overexpressed in human HCC stem cells and circulating HCC tumor cells [28]. Furthermore, in patients with NAFLD *PNPLA3*rs738409 G allele was found to be associated with decreased levels of adiponectin which has important anti-inflammatory, anti-fibrotic and anti-carcinogenic properties [29]. Recent data in cell culture and mice xenografts also suggest an effect of *PNPLA3* rs738409 on the activation of the interleukin (IL)-6-Janus kinase (JAK2)/Signal Transducers and Activators of Transcription (STAT3) pathway *via* fatty acid-mediated inflammation [30]. Therefore, functional studies linking the 'lipid droplet disease' caused by PNPLA3 malfunction to cancer may likely unravel interesting insights into the pathophysiology of HCC.

While the promoting effect of *PNPLA3* rs738409 on HCC evolution remains incompletely understood, even less is known about the possible role of genetic variation of *TM6SF2* in this process. TM6SF2 is involved in the assembly and secretion of very low-density lipoproteins, similar to microsomal triglyceride transfer protein (MTTP), and thus a loss of function mutation would be compatible with lipid accumulation in hepatocytes. Accordingly, Sookoian et al. [31] performed allele-specific expression analyses of cDNA isolated from liver tissue

16

from patients with NASH and confirmed that expression levels of the rs58542926 T allele are reduced to about 56% of that of the wild type C allele. These findings suggest that carriage of the *TM6SF2* T risk allele is associated with decreased hepatic gene and protein expression, leading to a loss of function similar to the mutant *PNPLA3* variant. In turn, the wild type TM6SF2 protein has been associated with increased serum VLDL levels and an elevated cardiovascular risk [32], but less liver injury [33]. Whether and how there may be cross-talk between TM6SF2 and signaling pathways pertinent for HCC development beyond mere intracellular lipid accumulation and lipotoxicity remains to be elucidated.

This study has its limitations, primarily the fact that it was undertaken retrospectively and was restricted to Caucasian people with alcohol-related cirrhosis. Thus the findings do not necessarily define the genetic risks for HCC development in people with liver disease of other aetiologies or to other ethnic groups. However, it also has its strengths, including the size of the study population, and the fact similar effect sizes were observed across the component geographical cohorts. Thus, the presented data provide robust evidence for a strong effect of *PNPLA3 rs738409* on HCC risk in patients with alcohol-related cirrhosis. In addition, *TM6SF2* rs58542926 is identified and confirmed as a significant host genetic risk factor for HCC, while *MBOAT7* rs641738 is not. In combination, *PNPLA3 rs738409* and *TM6SF2* rs58542926 account for almost 50% of the population-attributable risk for the development of HCC, and thus, should be evaluated further for stratification of patients at risk in screening programs. In addition, sight should not be lost of the fact that a significant proportion of the genetic variance still remains to be identified.

17

Stickel et al. Genetic risk of hepatocellular carcinoma

## Competing Interests

None of the authors has any conflict of interest in relation to the contents of this manuscript.

## REFERENCES

1. Fattovich G, Stroffolini T, Zagni I, Donato F. Hepatocellular carcinoma in cirrhosis: incidence and risk factors. *Gastroenterology* 2004;**127**:S35-50.

2. McGlynn KA, London WT. The global epidemiology of hepatocellular carcinoma: present and future. *Clin Liver Dis* 2011;**15**:223-43.

3. Nordenstedt H, White DL, El-Serag HB. The changing pattern of epidemiology in hepatocellular carcinoma. *Dig Liver Dis* 2010;**42**(Suppl 3):S206-14.

4. Ascha MS, Hanouneh IA, Lopez R, et al. The incidence and risk factors of hepatocellular carcinoma in patients with nonalcoholic steatohepatitis. *Hepatology* 2010;**51**:1972-8.

5. Seitz HK, Stickel F. Molecular mechanisms of alcohol-mediated carcinogenesis. *Nat Rev Cancer* 2007;**7**:599-612.

6. Stickel F, Moreno C, Hampe J, Morgan MY. Genetics of alcohol dependence and alcohol-related liver disease. *J Hepatol* 2017;**66**:195-211.

7. Chamorro AJ, Torres JL, Miron-Canelo JA, et al. Systematic review with meta-analysis: the I148M variant of patatin-like phospholipase domain-containing 3 gene (*PNPLA3*) is significantly associated with alcoholic liver cirrhosis. *Aliment Pharmacol Ther* 2014;**40**:571-81.

8. Stickel F, Buch S, Lau K, et al. Genetic variation in the *PNPLA3* gene is associated with alcoholic liver injury in Caucasians. *Hepatology* 2011;**53**:86-95.

9. Stickel F, Hampe J, Trepo E, et al. *PNPLA3* genetic variation in alcoholic steatosis and liver disease progression. *Hepatobiliary Surg Nutr* 2015;**4**:152-60.

10. 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**:325-34.

19

11. Trepo E, Nahon P, Bontempi G, et al. Association between the *PNPLA3* (rs738409 C>G) variant and hepatocellular carcinoma: Evidence from a meta-analysis of individual participant data. *Hepatology* 2014;**59**:2170-7.

12. Buch S, Stickel F, Trepo 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.

13. Nischalke HD, Lutz P, Kramer B, et al. A common polymorphism in the *NCAN* gene is associated with hepatocellular carcinoma in alcoholic liver disease. *J Hepatol* 2014;**61**:1073-9.

14. Falleti E, Cussigh A, Cmet S, et al. *PNPLA3* rs738409 and *TM6SF2* rs58542926 variants increase the risk of hepatocellular carcinoma in alcoholic cirrhosis. *Dig Liver Dis* 2016;**48**:69-75.

15. EASL–EORTC. Clinical Practice Guidelines: Management of hepatocellular carcinoma. *J Hepatol* 2012;**56**:908–43.

16. Pugh RN, Murray-Lyon IM, Dawson JL, et al. Transection of the oesophagus for bleeding oesophageal varices. *Br J Surg* 1973;**60**:646-9.

17. Tomlinson IP, Webb E, Carvajal-Carmona L, et al. A genome-wide association study identifies colorectal cancer susceptibility loci on chromosomes 10p14 and 8q23.3. *Nat Genet* 2008;**40**:623-30.

18. 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.

19. 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**:4309.

Stickel et al. Genetic risk of hepatocellular carcinoma

20. 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.

21. Donati B, Dongiovanni P, Romeo S, et al. *MBOAT7* rs641738 variant and hepatocellular carcinoma in non-cirrhotic individuals. *Sci Rep* 2017;**7**:4492.

22. Thabet K, Asimakopoulos A, Shojaei M, et al. *MBOAT7* rs641738 increases risk of liver inflammation and transition to fibrosis in chronic hepatitis C. *Nat Commun* 2016;**7**:12757.

23. Dupont WD, Plummer WD. Power and sample size calculations for studies involving linear regression. *Control CliniTrials* 1998;**19**:589-601.

24. Anstee QM, Liu YL, Day CP, Reeves HL. Reply to: HCC and liver disease risk in homozygous PNPLA3 p.I148M carriers approach monogenic inheritance. *J Hepatol* 2015;**62**:982-3.

25. Liu YL, Patman GL, Leathart JB, 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.

26. Valenti L, Al-Serri A, Daly AK, et al. Homozygosity for the patatin-like phospholipase-3/adiponutrin I148M polymorphism influences liver fibrosis in patients with nonalcoholic fatty liver disease. *Hepatology* 2010;**51**:1209-17.

27. Pare G, Ridker PM, Rose L, et al. Genome-wide association analysis of soluble ICAM-1 concentration reveals novel associations at the *NFKBIK, PNPLA3, RELA,* and *SH2B3* loci. *PLoS Genet* 2011;**7**:e1001374.

28. Liu S, Li N, Yu X, et al. Expression of intercellular adhesion molecule 1 by hepatocellular carcinoma stem cells and circulating tumor cells. *Gastroenterology* 2013;**144**:1031-41.

Stickel et al.                                                              Genetic risk of hepatocellular carcinoma

29. Valenti L, Rametta R, Ruscica M, et al. The I148M *PNPLA3* polymorphism influences serum adiponectin in patients with fatty liver and healthy controls. *BMC Gastroenterol* 2012;**12**:111.

30. Liu Z, Chen T, Lu X, et al. Overexpression of variant *PNPLA3* gene at I148M position causes malignant transformation of hepatocytes via IL-6-JAK2/STAT3 pathway in low dose free fatty acid exposure: a laboratory investigation in vitro and in vivo. *Am J Transl Res* 2016;**8**:1319-38.

31. Sookoian S, Castaño GO, Scian R, et al. Genetic variation in transmembrane 6 superfamily member 2 and the risk of nonalcoholic fatty liver disease and histological disease severity. *Hepatology* 2015;**61**:515–25.

32. 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.

33. Sookoian S, Pirola CJ. Meta-analysis of the influence of *TM6SF2 E167K* variant on plasma concentration of aminotransferases across different populations and diverse liver phenotypes. *Sci Rep* 2016;**6**:27718.

Stickel et al.                                        Genetic risk of hepatocellular carcinoma

**LEGENDS TO FIGURES**


**Figure 1**

Proportion of cases with alcohol–related cirrhosis and HCC by carriage of the risk variants in

*PNPLA3* and *TM6SF2*.


**Supplementary figures**

**Figure 1A**

Forrest plot displaying a consistent association of *PNPLA3* rs738409 with HCC across different

national cohorts. The association is similar in all three cohorts.


**Figure 1B**

Forrest plot displaying association of *TM6SF2* rs58542926 with HCC in different national

cohorts. The association is similar in all three cohorts.


**Figure 1C**

Forrest plots ahowing no significant association between *MBOAT7* rs641738 with HCC in any

of the three cohorts.

**Table 1**

Characteristics of cases and controls included in the study

| Variable | Total (n=1916) | Cases (n= 751) | Controls (n=1165) | Significance (p) |
|---|---|---|---|---|
| **Proportion male** (n: %) | 1496 (78) | 679 (90) | 817 (70) | < 0.00001 |
| **Age (yr)** | 58±11 | 61±10 | 55±10 | < 0.00001 |
| **BMI (kg/m$^2$)** | 26.6±4.7 | 27.9±4.7 | 25.9±4.6 | < 0.00001 |
| **BMI (%)** <25 | 36% | 26% | 43% | |
| 25-30 | 42% | 42% | 42% | |
| ≥30 | 21% | 32% | 15% | < 0.00001 |
| **Child Pugh Grade** [16] **(%)** A | 35% | 48% | 27% | |
| B | 40% | 36% | 42% | |
| C | 26% | 16% | 31% | < 0.00001 |
| **Type II diabetes mellitus** (n: %) | 27 (1.4) | 42 (5.6) | 18 (1.5) | < 0.00001 |

Data are reported as mean ± standard deviation or as number (%)

**Table 2**

Genotypic and allelic frequencies of *PNPLA3, TM6SF2* and *MBOAT7* in people with alcohol-related cirrhosis with HCC (cases) and without (controls)

| SNP | Total (n=1916) | | | | | | Unadjusted OR (CI 95%) | Adjusted OR (CI 95%) | P trend | $P_{allelic}$ | $P_{genotypic}$ |
|---|---|---|---|---|---|---|---|---|---|---|---|
| *PNPLA3* | | CC | CG | GG | MAF | fCG (%) | fGG (%) | CG vs CC | GG vs CC | | | |
| Cases | 751 | 205 | 346 | 200 | 0.497 | 46% | 27% | 1.70 [1.37-2.10] | 3.26 [2.50-4.25] | 5.45E-19 | 3.64E-20 | 6.14E-18 |
| Controls | 1165 | 508 | 505 | 152 | 0.347 | 43% | 13% | | | | | |
| *TM6SF2* | | CC | CT | TT | MAF (%) | fCT (%) | fTT (%) | CT vs TT | CC vs TT | | | |
| Cases | 751 | 558 | 164 | 29 | 0.148 | 22% | 4% | 1.46 [1.15-1.84] | 3.32 [1.76-6.24] | 1.22E-06 | 8.93E-07 | 7.74371E-06 |
| Controls | 1165 | 957 | 193 | 15 | 0.096 | 17% | 1% | | | | | |
| *MBOAT7* | | CC | CT | TT | MAF (%) | fCT (%) | fTT (%) | CT vs TT | CC vs TT | | | |
| Cases | 751 | 203 | 363 | 185 | 0.488 | 48% | 25% | 0.96 [0.77-1.20] | 1.07 [0.83-1.38] | 0.377 | 0.639 | 0.676 |
| Controls | 1165 | 314 | 583 | 268 | 0.480 | 50% | 23% | | | | | |

MAF, minor allele frequency; OR, odds ratio

**Table 3**

Multivariate analysis of factors associated with hepatocellular carcinoma in people with alcohol-related cirrhosis

| SNP | *Adjusted OR (95% CI) | Significance (p) |
|---|---|---|
| *PNPLA3* rs738409:G | 1.84 (1.55-2.18) | $1.85 \times 10^{-12}$ |
| *PNPLA3* rs738409:CG | 2.07 [1.58-2.73] | $1.87 \times 10^{-07}$ |
| *PNPLA3* rs738409:GG | 3.26 [2.31-4.60] | $1.87 \times 10^{-11}$ |
| *TM6SF2* rs58542926:T | 1.66 (1.30-2.13) | $5.13 \times 10^{-05}$ |
| *TM6SF2* rs58542926:CT | 1.67 [1.23-2.26] | $1.03 \times 10^{-03}$ |
| *TM6SF2* rs58542926:TT | 2.86 [1.34-6.12] | $6.75 \times 10^{-03}$ |
| *MBOAT7* rs641738:T | 1.04 (0.88-1.24) | 0.61 |
| *MBOAT7* rs641738:CT | 0.90 [0.67-1.20] | .473 |
| *MBOAT7* rs641738:TT | 1.09 [0.78-1.53] | .619 |

SNP: single nucleotide polymorphism; OR: odds ratio; CI: confidence interval

* adjusted for age, gender, BMI, type II diabetes mellitus.

**Table 4**

Population-attributable fraction of each variant with regard to the genetic risk of HCC in alcohol-related cirrhosis alone, and in combination.

| Genetic variant | *Population Attributable Fraction (%) |
|---|---|
| *PNPLA3 r*s738409 | 43.5 |
| *TM6SF2* rs58542926 | 11.5 |
| *PNPLA3 + TM6SF2* (combined) | 49.9 |

* Estimated for heterozygous and homozygous carriage in one equation.

**Table 5**
Proposed screening intervals for HCC of people with alcohol-related cirrhosis based on the genetic signature

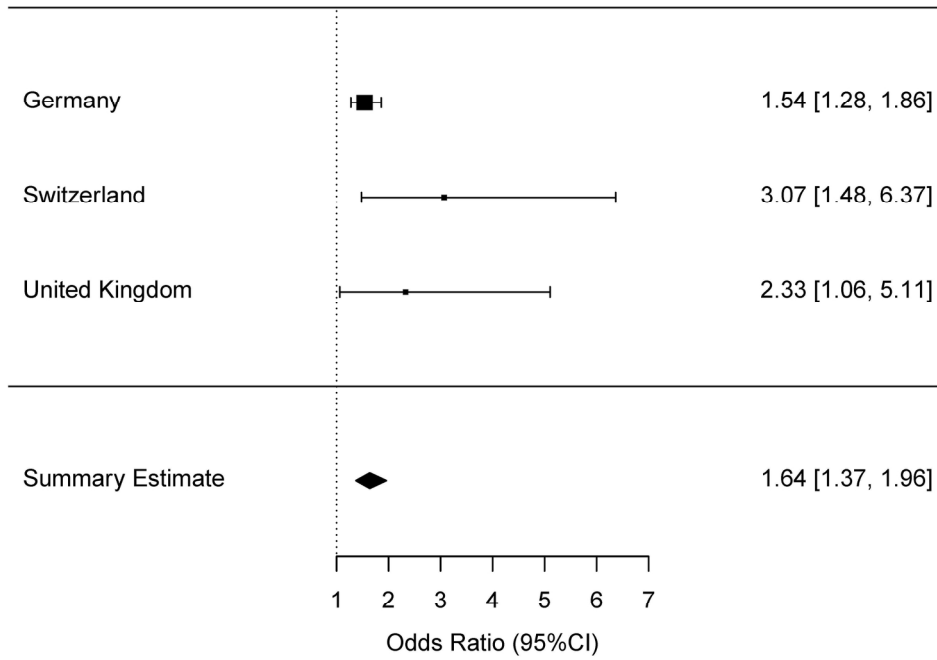| Genetic background | Screening interval |
|---|---|
| *PNPLA3 r*s738409 WT (CC), *TM6SF2* rs58542926:CC | 24 months |
| *TM6SF2* rs58542926:T\* and *PNPLA3 r*s738409 CC | 12 months |
| *TM6SF2* rs58542926:CC *and PNPLA3 r*s738409 GG | 6 months |
| *PNPLA3 r*s738409 GG *and TM6SF2* rs58542926:T\* | 3 months |

**Figure 1**



■ Alcohol-related cirrhosis without HCC (1165)
■ Alcohol-related cirrhosis with HCC  (n=751)
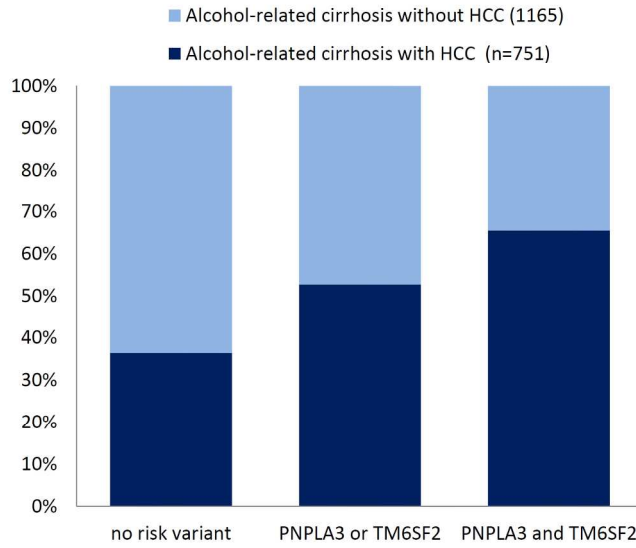
| Population (n) | No risk variants (%) | PNPLA3 or TM6SF2 (%) | PNPLA3 and TM6SF2 (%) |
|---|---|---|---|
| Alcohol-related  cirrhosis without HCC (751) | 35.6 | 55.0 | 9.9 |
| Alcohol-related  cirrhosis with HCC (1165) | 20.4 | 61.0 | 18.8 |
| Total (n) | 568 | 1092 | 256 |

Proportion of cases with alcohol–related cirrhosis and HCC by carriage of the risk variants in PNPLA3 and TM6SF2
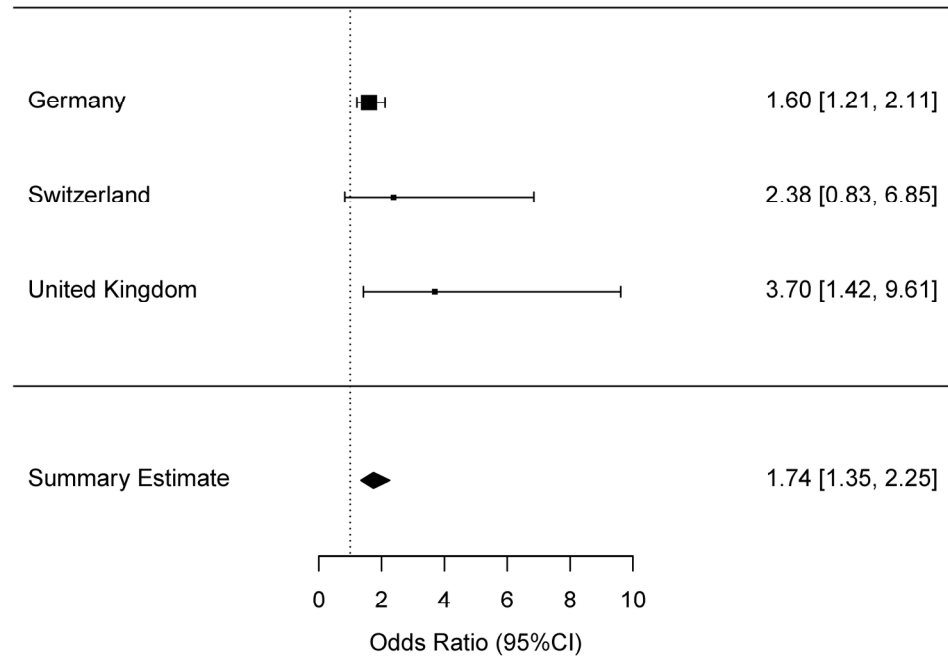
297x420mm (300 x 300 DPI)

Association p−value= $4.99 \times 10^{-08}$
Heterogeneity p−value= 0.134

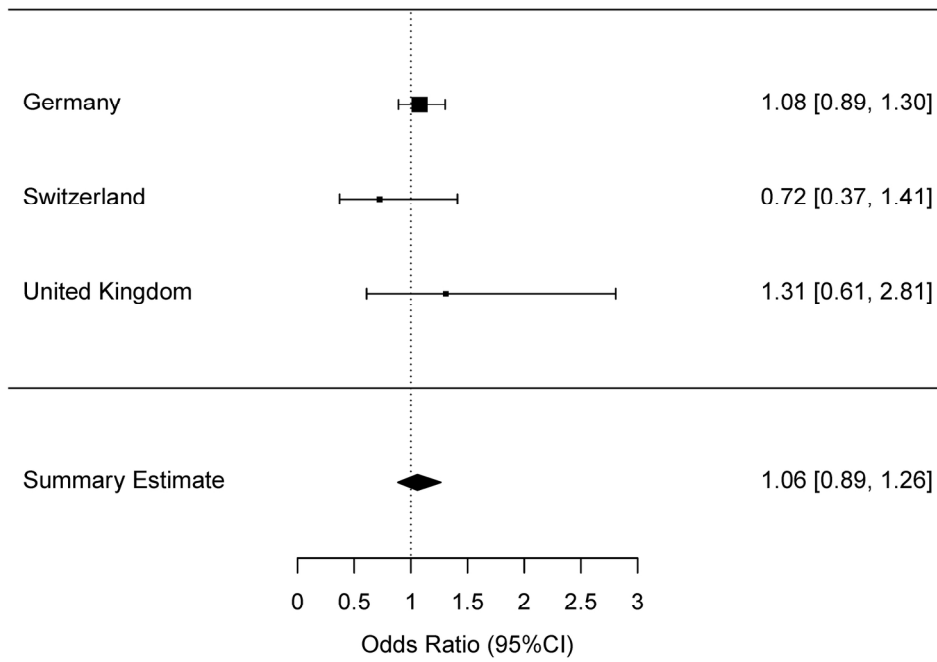| | | |
|---|---|---|
| Germany | | 1.54 [1.28, 1.86] |
| Switzerland | | 3.07 [1.48, 6.37] |
| United Kingdom | | 2.33 [1.06, 5.11] |
| Summary Estimate | | 1.64 [1.37, 1.96] |

Odds Ratio (95%CI)

177x177mm (300 x 300 DPI)

Association p−value= $2.47 \times 10^{-05}$
Heterogeneity p−value= 0.215

| | | |
|---|---|---|
| Germany | | 1.60 [1.21, 2.11] |
| Switzerland | | 2.38 [0.83, 6.85] |
| United Kingdom | | 3.70 [1.42, 9.61] |
| Summary Estimate | | 1.74 [1.35, 2.25] |

Odds Ratio (95%CI)

0  2  4  6  8  10

177x177mm (300 x 300 DPI)

Association p−value= 0.532
Heterogeneity p−value= 0.456

| | | |
|---|---|---|
| Germany | | 1.08 [0.89, 1.30] |
| Switzerland | | 0.72 [0.37, 1.41] |
| United Kingdom | | 1.31 [0.61, 2.81] |
| Summary Estimate | | 1.06 [0.89, 1.26] |

0    0.5    1    1.5    2    2.5    3

Odds Ratio (95%CI)

177x177mm (300 x 300 DPI)