


Genetic determinants of steatosis and fibrosis progression in paediatric non-alcoholic fatty liver disease

Christian A. Hudert¹  | Silvia Selinski² | Birgit Rudolph³ | Hendrik Bläker³ | Christoph Loddenkemper³ | Ria Thielhorn⁴ | Nikolaus Berndt⁵ | Klaus Golka² | Cristina Cadenas² | Jörg Reinders² | Stephan Henning⁶ | Philip Bufler⁶ | Peter L. M. Jansen⁷ | Hermann-Georg Holzhütter⁵ | David Meierhofer⁴ | Jan G. Hengstler² | Susanna Wiegand¹

¹Center for Chronically Sick Children, Charité – Universitätsmedizin Berlin, Berlin, Germany

²Systems Toxicology, Leibniz Research Centre for Working Environment and Human Factors at the Technical University Dortmund, Dortmund, Germany

³Institute of Pathology, Charité – Universitätsmedizin Berlin, Berlin, Germany

⁴Max Planck Institute for Molecular Genetics, Mass Spectrometry Facility, Berlin, Germany

⁵Institute for Biochemistry, Charité – Universitätsmedizin Berlin, Berlin, Germany

⁶Department of Pediatric Gastroenterology, Charité – Universitätsmedizin Berlin, Berlin, Germany

⁷Department of Gastroenterology and Hepatology, Academic Medical Center, University of Amsterdam, Amsterdam, The Netherlands

Correspondence

Christian A. Hudert, Center for Chronically Sick Children, Charité – Universitätsmedizin Berlin, Berlin, Germany.
Email: christian.hudert@charite.de

Funding information

This work was supported by the German Federal Ministry of Education and Research (BMBF, Liver Systems Medicine—LiSyM, FKZ 031L0057, 031L0058).

Handling Editor: Frank Tacke

Abstract

Background and Aims: Non-alcoholic fatty liver disease (NAFLD) is the most common chronic liver disease in children and adolescents today. In comparison with adult disease, paediatric NAFLD may show a periportal localization, which is associated with advanced fibrosis. This study aimed to assess the role of genetic risk variants for histological disease pattern and severity in childhood NAFLD.

Methods: We studied 14 single nucleotide polymorphisms (SNP) in a cohort of 70 adolescents with biopsy-proven NAFLD. Genotype was compared to an adult control cohort (n = 200) and analysed in relation to histological disease severity and liver tissue proteomics.

Results: Three of the 14 SNPs were significantly associated with paediatric NAFLD after FDR adjustment, rs738409 (PNPLA3, $P = 2.80 \times 10^{-06}$), rs1044498 (ENPP1, $P = 0.0091$) and rs780094 (GCKR, $P = 0.0281$). The severity of steatosis was critically associated with rs738409 (OR=3.25; 95% CI: 1.72-6.52, FDR-adjusted $P = 0.0070$). The strongest variants associated with severity of fibrosis were rs1260326, rs780094 (both GCKR) and rs659366 (UCP2). PNPLA3 was associated with a portal pattern of steatosis, inflammation and fibrosis. Proteome profiling revealed decreasing levels of GCKR protein with increasing carriage of the rs1260326/rs780094 minor alleles and downregulation of the retinol pathway in rs738409 G/G carriers. Computational metabolic modelling highlighted functional relevance of PNPLA3, GCKR and UCP2 for NAFLD development.

Conclusions: This study provides evidence for the role of PNPLA3 as a determinant of portal NAFLD localization and severity of portal fibrosis in children and adolescents, the risk variant being associated with an impaired hepatic retinol metabolism.

See Editorial on Page 440

Abbreviations: BMI, body mass index; CI, confidence interval; FDR, false discovery rate; GWAS, genome-wide association study; HOMA-IR, homeostasis model assessment for insulin resistance; KEGG, Kyoto Encyclopedia of Genes and Genomes; LC-MS, liquid chromatography-mass spectrometry; LR, likelihood ratio; MAF, minor allele frequency; NAFLD, non-alcoholic fatty liver disease; NAS, NAFLD activity score; NASH CRN, NASH clinical research network; NASH, non-alcoholic steatohepatitis; OR, odds ratio; SNP, single nucleotide polymorphism.

Hengstler and Wiegand contributed equally.

KEYWORDS

genetics, metabolic modelling, paediatric NAFLD, proteomics

1 | INTRODUCTION

Since the marked increase in obesity prevalence, non-alcoholic fatty liver disease (NAFLD) has become the most frequent chronic liver disease in adults and children alike.¹ NAFLD is strongly associated with obesity as well as insulin resistance and has been proposed as the hepatic manifestation of the metabolic syndrome. However, marked differences in the prevalence of NAFLD according to ethnicity have highlighted the importance of genetic influence in addition to environmental factors. In an autopsy study of 742 children in the US, NAFLD was five times more likely to occur in children with Hispanic ethnicity as compared to Black ethnicity.² This finding was further reflected in adult US citizens,³ and ultimately in general population prevalence of NAFLD, with highest rates registered in South America, while prevalence remains low in Africa. Heritability studies, including twin studies,^{4,5} have shown the increased familial risk of NAFLD for first-degree relatives and monozygotic twins. Overall, the estimation of heritability of NAFLD is thought to be roughly 50%.⁶ Another indicator of the interplay of environment and genetic factors is the high interindividual variability of disease progression of NAFLD. While exact data on the natural history of the disease are sparse due to the lack of prospective cohorts with sequential liver biopsies, it is estimated that only one-quarter of NAFLD patients develop non-alcoholic steatohepatitis (NASH).⁷ Furthermore, the well-known differing histological patterns in paediatric NAFLD, presenting with a portal/zone 1-predominant type or an "adult" zone 3-predominant type,⁸ suggest a distinct genetic impact on the development of disease.⁹ This phenomenon is of particular importance, since the portal/zone 1 pattern is more frequent in younger children and associated with an increased risk of advanced fibrosis.^{10,11} Genetic variants associated with the portal/zone 1-predominant type in paediatric NAFLD have not yet been identified.

Since the advent of genome-wide association studies, multiple genes have been proposed to influence the development and progression of NAFLD. These include variants of genes involved in insulin response, fatty acid metabolism, oxidative stress control and endotoxin response, inflammatory cytokines and modifiers of fibrogenesis (for overview, see Table 1 and Table S1). The missense variant of PNPLA3 is now seen as a main contributor to NAFLD and modifies all aspects of disease progression. Although variants of several genes, particularly PNPLA3 and GCKR, have been associated with childhood NAFLD, only few studies have looked at histological disease severity.¹² It would be of high clinical relevance to identify genetic markers that allow discrimination of patients at risk of early progressive disease. To reach this goal, the Berlin adolescence NAFLD cohort (BaNA) has recently been established. For all patients of this cohort, liver biopsies are available that have been categorized according to the severity of steatosis, inflammation

Key points

- PNPLA3 148M is the strongest determinant of paediatric NAFLD and is associated with histological disease severity and periportal zonation.
- PNPLA3 148M is associated with downregulated retinol metabolism on proteomic analysis.
- GCKR variant regulates the actual liver protein level and is associated with the presence of fibrosis.
- Computational modelling reveals mechanistic significance of PNPLA3, GCKR and UCP2.

and fibrosis. DNA of all patients was analysed for 14 SNPs previously shown to be associated with adult NAFLD, and the results were analysed in relation to histological scores and the proteome of the biopsies.

2 | PATIENTS AND METHODS

2.1 | Paediatric NAFLD cohort

The study protocol conformed to the guidelines of the Declaration of Helsinki and was approved by the local institutional review board at Charité University Medicine Berlin. Informed consent was obtained from all parents or guardians. Patients were recruited from the paediatric obesity outpatient clinic and paediatric gastroenterology outpatient clinic of Charité between June 2014 and July 2017. Standard serologic tests were obtained for exclusion of alpha-1-antitrypsin deficiency, coeliac disease, autoimmune hepatitis, viral hepatitis A, B and C, active CMV or EBV infection, and Wilson's disease. Overweight or obese children and adolescents aged 10 to 17 years with suspected non-alcoholic steatohepatitis (NASH) and a clinical indication for liver biopsy were evaluated for enrolment in the study. Exclusion criteria were as follows: age ≥ 18 years, any concurrent liver disease, severe underlying chronic disease (eg, cardiopulmonary or autoimmune disease), alcohol consumption >20 g alcohol per day and pregnancy. In all patients included in our study, anthropometric measures (height, weight, waist circumference) were taken, laboratory analysis including a hepatic panel and complete blood count was performed, and metabolic (serum lipid profile, HOMA-IR, lactate, pyruvate, uric acid) parameters were assessed. For comparison, we studied 200 adult individuals without reported NAFLD (Neuss Bladder Cancer case-control series) that have already been used as a Caucasian reference group for SNP analyses in previous studies.¹³

2.2 | Histological assessment

All liver biopsies were evaluated and scored by pathologists with many years of experience in hepatic pathology. All histological scoring was carried out blinded to clinical data. Discrepancies were discussed, and a consensus was reached on a final score. Formalin-fixed, paraffin-embedded biopsy samples were scored by using 2- μ m sections stained with haematoxylin and eosin, chromotrope aniline blue trichrome, Masson trichrome and Gomori silver. All biopsy specimens were considered technically adequate for evaluation. Liver biopsies were staged and graded by using the histological scoring system for non-alcoholic fatty liver disease (NAFLD) by the NASH Clinical Research Network (NASH CRN).¹⁴ We defined $F \geq 1$ as "any," $F \geq 2$ as "progressive" and $F = 3$ as "advanced" fibrosis. Moreover, to account for specific features of paediatric NASH, we defined zonal steatosis pattern as 1 = "predominantly zone 1," 3 = "predominantly zone 3," 4 = "panlobular" or 0 = "azonal" distribution) and scored portal inflammation

(0 = none; 1 = mild; 2 = more severe than mild). Additionally, an individual score for periportal septa (0 = no portal fibrosis or enlarged portal field without infiltrative septa; 1 = enlarged portal field with short infiltrative septa <50% of the porto-portal distance; 2 = long infiltrative septa >50% of the porto-portal distance, may present with or without porto-portal bridging) was introduced.

2.3 | Analysis of polymorphisms

Genotypes of rs738409[C/G], rs2294918[A/G], rs1260326[C/T], rs780094[C/T], rs58542926[C/T], rs13412852[C/T], rs4880[T/C], rs3750861[C/T], rs641738[C/T], rs659366[C/T], rs1044498[A/C], rs1801278[A/G], rs12979860[C/T] and rs12137855[C/T] were detected viaTaqMan[®] SNP Genotyping Assay (Applied Biosystems, Darmstadt, Germany).¹⁵ Experimental details are given in Table S2A and Methods S1. All SNPs were in Hardy-Weinberg equilibrium (Table S2B).

TABLE 1 Single nucleotide polymorphisms (SNPs) analysed in the present study, SNP function, associated genes and frequency among cases and controls

SNP	Gene	Gene	Function	Chromosome : Location	Controls MAF in %	Cases MAF in %
rs738409[C/G]	PNPLA3	Patatin-like phospholipase domain-containing 3	Missense	22:43928847	[G]: 23	49
rs2294918[A/G] ^a	PNPLA3	Patatin-like phospholipase domain-containing 3	Missense	22:43946236	[A]: 39	28
rs1260326[C/T]	GCKR	Glucokinase (hexokinase 4) regulator	Downstream variant 500B, missense	2:27508073	[T]: 41	52
rs780094[C/T]	GCKR	Glucokinase (hexokinase 4) regulator	Intron	2:27518370	[T]: 39	52
rs58542926[C/T]	TM6SF2	Transmembrane 6 superfamily member 2	Missense	19:19268740	[T]: 5	9
rs13412852[C/T]	LPIN1	Lipin 1	Intron	2:11774815	[T]: 35	34
rs4880[T/C] ^b	SOD2	Superoxide dismutase 2, mitochondrial	Missense, utr variant 5'	6:159692840	[C]: 48	53
rs3750861[C/T]	KLF6	Kruppel-like factor 6	Intron	10:3782241	[T]: 8	8
rs641738[C/T]	MBOAT7, TMC4	Membrane-bound O-acyltransferase domain-containing 7, transmembrane channel-like 4	Downstream variant 500B, missense, nc transcript variant	19:54173068	[T]: 44	44
rs659366[C/T]	UCP2	Uncoupling protein 2 (mitochondrial, proton carrier)	Upstream variant 2KB	11:7398370	[T]: 40	32
rs1044498[A/C]	ENPP1	Ectonucleotide pyrophosphatase/phosphodiesterase 1	Missense	6:131851228	[C]: 8	19
rs1801278[A/G] ^c	IRS1	Insulin receptor substrate 1	Missense	2:226795828	[A]: 6	6
rs12979860[C/T]	IFNL4	Interferon, lambda 4 (gene/pseudogene)	Intron	19:39248147	[T]: 36	40
rs12137855[C/T]	LYPLAL1	Lysophospholipase-like 1	Intron	1:219275036	[T]: 19	21

MAF, minor allele frequency.

Location according to GRCh38.p7.

^a[A] is the minor allele (about 43% in Europeans), and the risk allele [G] is the major allele (about 57% in Europeans).

^b[C] corresponding to the exchange to [Ala] is the minor allele (about 48% in Europeans), and [T] corresponding to [Val] is the major allele (about 52% in Europeans).

^c[A] is the minor allele (about 7% in Europeans), and [G] is the major allele (about 93% in Europeans).

TABLE 2 Study group characteristics

Variable	Cases		Controls	
	n	%	n	%
Gender				
Female	16	23%	77	39%
Male	54	77%	123	62%
Total	70	100%	200	100%
Age				
Prepubertal/early puberty	28	40%	0	0%
Pubertal/post-pubertal	42	60%	0	0%
Total	70	100%	200	100%
Range				
Range	10 - 17		18 - 70.5	
Interquartile range	13 - 16		31.25 - 60.85	
Mean	14.11		46.73	
SD	2.15		16.03	
BMI				
Range	21.4 - 50.4			
Interquartile range	30.9 - 37.8			
Mean	34.63			
SD	6.11			
Total	70			
BMI z-score				
Range	1.29 - 3.86			
Interquartile range	2.47 - 3.14			
Mean	2.76			
SD	0.58			
Total	70			
Portal Inflammation				
None	25	36%		
Mild	37	53%		
More severe than mild	8	11%		
Total	70	100%		
Lobular Inflammation				
None	21	30%		
<2 foci/ 200x field	39	56%		
2-4 foci/200x field	10	14%		
Total	70	100%		
Ballooning				
Absent	36	51%		
Few	27	39%		
Prominent	7	10%		
Total	70	100%		
NAS				
NAS < 5	46	66%		
NAS ≥ 5	24	34%		
Total	70	100%		

(Continues)

TABLE 2 (Continued)

Variable	Cases		Controls	
	n	%	n	%
Steatosis				
5–33% of hepatocytes	13	19%		
34–66% of hepatocytes	24	34%		
>66% of hepatocytes	33	47%		
Total	70	100%		
Steatosis Zonation				
Zone 1 (periportal)	11	16%		
Zone 3 (centrolobular)	28	40%		
Panlobular	19	27%		
Azonal	12	17%		
Total	70	100%		
Fibrosis				
No fibrosis	19	27%		
Portal/periportal or zone 3	19	27%		
Portal/periportal and zone 3	14	20%		
Bridging fibrosis	18	26%		
Total	70	100%		
Periportal Septa				
No septa	35	50%		
Short infiltrative septa	12	17%		
Long infiltrative septa or porto-portal bridging	23	33%		
Total	70	100%		
ALT				
Range	51 - 338			
Interquartile range	62 - 125			
Mean	108.20			
SD	68.49			
Total	70			
HOMA-IR				
Range	1.3 - 21.0			
Interquartile range	4.6 - 8.95			
Mean	7.04			
SD	3.73			
Total	68			

Prepubertal or early puberty: girls 10–12, boys 10–13 years old.

Pubertal or post-pubertal: girls 13–17, boys 14–17 years old.

SD: standard deviation.

2.4 | Proteome profiling

Liver biopsy specimens were extracted under denaturing conditions and digested by trypsin for subsequent analysis by mass spectrometry.¹⁶ For one patient, no proteome data were available. The software tools MaxQuant¹⁷ and gene set enrichment analysis (GSEA)¹⁸ were used for peptide identification and pathway analyses respectively (see Methods S1 for details).

2.5 | Statistical methods

Case-control genotype differences were analysed by chi-squared tests and logistic regression models. FDR was used to adjust for multiple testing. A backward selection on all SNPs assuming an additive (trend) model was done with SAS/STAT, version 9.3 (SAS Institute Inc, Cary, NC). Differences within NAFLD cases regarding age, gender, histopathological findings, proteins and genetic

TABLE 3 PNPLA3 missense SNP rs738409, ENPP1 missense SNP rs1044498 and GCKR intron SNP rs780094 showed FDR-adjusted significant differences between cases and controls. A, Crude and FDR-adjusted *P*-values for the most important genetic models. FDR-adjusted *P*-values for the recessive and dominant mode of inheritance are given in Table S4. *P*-values below 0.05 are printed red and bold face. B, Unadjusted *P*-values, odds ratios (OR) and 95% confidence intervals (CI) for the genotype differences, additive (trend) and multiplicative (alleles) mode of inheritance. SNPs and *P*-values <0.05 that were significant in the case-control analysis are printed red

A		Likelihood ratio (LR) test		Genotype trend test		Allele Wald test	
Test	Gene	Unadjusted	FDR adjusted	Unadjusted	FDR adjusted	Unadjusted	FDR adjusted
rs738409	PNPLA3	3.36×10 ⁻⁰⁸	4.71×10 ⁻⁰⁷	2.00×10 ⁻⁰⁷	2.80×10 ⁻⁰⁶	1.62×10 ⁻⁰⁸	2.27×10 ⁻⁰⁷
rs1044498	ENPP1	0.0013	0.0092	0.0013	0.0091	0.0007	0.0050
rs780094	GCKR	0.0182	0.0847	0.0060	0.0281	0.0070	0.0329
rs58542926	TM6SF2	0.0332	0.1163	0.0891	0.1967	0.0750	0.1751
rs1260326	GCKR	0.0466	0.1306	0.0164	0.0575	0.0198	0.0600
rs2294918	PNPLA3	0.0711	0.1660	0.0254	0.0712	0.0214	0.0600
rs12137855	LYPLAL1	0.1068	0.2135	0.4446	0.6224	0.4499	0.6298
rs641738	MBOAT7, TMC4	0.1795	0.3141	0.9881	0.9881	0.9883	0.9883
rs4880	SOD2	0.2680	0.3751	0.2812	0.4921	0.2755	0.4821
rs659366	UCP2	0.2481	0.3751	0.0983	0.1967	0.1000	0.2000
rs13412852	LPIN1	0.6307	0.6748	0.8381	0.9778	0.8369	0.9764
rs3750861	KLF6	0.6405	0.6748	0.9678	0.9881	0.9675	0.9883
rs12979860	IFNL4	0.6611	0.6748	0.3998	0.6219	0.3699	0.5754
rs1801278	IRS1	0.6748	0.6748	0.6752	0.8593	0.6847	0.8714

The likelihood ratio (LR) test was used for the genotype model (global genotype differences). The genotype trend test corresponds to the additive model for increasing number of risk alleles in the genotype. The allele Wald test corresponds to the multiplicative model.

B		Cases		Controls		LR test		Wald test		95% CI	
SNP	Genetic model	Genotype	Cases	Controls	OR	P-value	L	U	L	U	
rs738409	Genotypes	C/C	20	118	2.58	Reference	1.38	4.92	1.38	4.92	
PNPLA3	Genotypes	C/G	31	71	10.19	0.0035	4.31	25.37	4.31	25.37	
		G/G	19	11	3.05	2.40×10 ⁻⁰⁷	2.02	4.71	2.02	4.71	
rs2294918	Trend	Increasing [G]	71	307	3.21	2.00×10 ⁻⁰⁷	2.14	4.82	2.14	4.82	
		Multiplicative	69	93	Reference	Reference	Reference	Reference	Reference	Reference	
PNPLA3	Genotypes	A/A	5	33	2.15	1.62×10 ⁻⁰⁸	0.82	6.73	0.82	6.73	
		A/G	29	89	3.05	0.1450	1.18	9.46	1.18	9.46	
rs2294918	Trend	Increasing [G]	39	155	1.61	0.0323	1.07	2.47	1.07	2.47	
		Multiplicative	101	245	1.64	0.0254	1.08	2.52	1.08	2.52	

(Continues)

TABLE 3 (Continued)

SNP	Genetic model	Genotype	Controls	Cases	LR test	Wald test P-value	OR	95% CI	
								L	U
rs1260326 GCKR	Genotypes	C/C	66	33%	21%	Reference			
		C/T	105	53%	53%	0.2024	1.55	0.80	3.12
	Trend	T/T	29	15%	26%	0.0154	2.73	1.22	6.24
		Increasing [T]				0.0164	1.65	1.10	2.50
rs780094 GCKR	Genotypes	[C]	237	59%	48%	Reference			
		[T]	163	41%	52%	0.0198	1.58	1.08	2.34
	Trend	Increasing [T]				Reference			
		Multiplicative				0.0182	1.72	0.89	3.44
rs58542926 TM6SF2	Genotypes	C/C	71	36%	21%	Reference			
		C/T	102	51%	53%	0.1149	1.72	0.89	3.44
	Trend	T/T	27	14%	26%	0.0058	3.16	1.40	7.23
		Increasing [T]				0.0060	1.77	1.18	2.69
rs13412852 LPIN1	Genotypes	[C]	244	61%	48%	Reference			
		[T]	156	39%	52%	0.0070	1.70	1.16	2.52
	Trend	Increasing [T]				Reference			
		Multiplicative				0.0332	2.72	1.17	6.22
rs4880 SOD2	Genotypes	C/C	184	92%	83%	Reference			
		C/T	14	7%	17%	0.0176	2.72	1.17	6.22
	Trend	T/T	2	1%	0%	0.9888	-	-	-
		Increasing [T]				0.0891	1.89	0.89	3.94
rs3750861 KLF6	Genotypes	[C]	382	96%	91%	Reference			
		[T]	18	5%	9%	0.0750	1.99	0.91	4.20
	Trend	Increasing [T]				Reference			
		Multiplicative				0.6307	0.78	0.43	1.40
rs4880 SOD2	Genotypes	C/C	83	42%	46%	Reference			
		C/T	93	47%	40%	0.4092	0.78	0.43	1.40
	Trend	T/T	24	12%	14%	0.8567	1.08	0.45	2.46
		Increasing [T]				0.8381	0.96	0.64	1.43
rs3750861 KLF6	Genotypes	[C]	259	65%	66%	Reference			
		[T]	141	35%	34%	0.8369	0.96	0.64	1.43
	Trend	Increasing [C]				Reference			
		Multiplicative				0.2680	0.88	0.45	1.75
rs3750861 KLF6	Genotypes	C/C	54	27%	26%	Reference			
		C/T	102	51%	43%	0.7147	0.88	0.45	1.75
	Trend	T/T	44	22%	31%	0.2824	1.50	0.72	3.17
		Increasing [C]				0.2812	1.23	0.84	1.82
rs3750861 KLF6	Genotypes	[C]	210	53%	47%	Reference			
		[T]	190	48%	53%	0.2755	1.24	0.84	1.82
	Trend	Increasing [C]				Reference			
		Multiplicative				0.6405	1.18	0.53	2.47
rs3750861 KLF6	Genotypes	C/C	171	86%	84%	Reference			
		C/T	27	14%	16%	0.6686	1.18	0.53	2.47
	Trend	T/T	2	1%	0%	0.9888	-	-	-
		Multiplicative				0.9888	-	-	-

(Continues)

TABLE 3 (Continued)

SNP	Genetic model	Genotype	Controls	Cases	LR test	Wald test P-value	OR	95% CI	
								L	U
rs641738	Trend	Increasing [T]				0.9678	1.01	0.48	2.01
	Multiplicative	[C]	369	129	92%	Reference			
MBOAT7, TMC4	Genotypes	C/C	59	25	36%	Reference			
		C/T	108	29	54%	0.1506	0.63	0.34	1.18
	Trend	Increasing [T]				0.7277	1.14	0.53	2.43
	Multiplicative	[C]	226	79	56%	Reference			
rs659366	Genotypes	C/C	71	32	46%	Reference			
		C/T	98	31	44%	0.2322	0.70	0.39	1.26
UCP2	Trend	Increasing [T]				0.0983	0.71	0.46	1.06
	Multiplicative	[C]	240	95	68%	Reference			
rs1044498	Genotypes	A/A	171	46	66%	Reference			
		A/C	26	22	31%	0.0006	3.15	1.63	6.06
ENPP1	Trend	Increasing [C]				0.3280	2.48	0.32	15.38
	Multiplicative	[A]	368	114	81%	Reference			
rs1801278	Genotypes	G/G	178	61	87%	Reference			
		A/G	22	9	13%	0.6752	1.19	0.50	2.66
IRS1	Trend	Increasing [A]				0.6752	1.19	0.50	2.66
	Multiplicative	[G]	378	131	94%	Reference			
rs12979860	Genotypes	C/C	88	28	40%	Reference			
		C/T	81	28	40%	0.7881	1.09	0.59	1.99
IFNL4	Trend	Increasing [T]				0.3998	1.17	0.81	1.70
	Multiplicative	[A]	22	9	6%	0.6847	1.18	0.50	2.55

(Continues)

TABLE 3 (Continued)

SNP	Genetic model	Genotype	Controls	Cases	LR test	Wald test P-value	OR	95% CI	
								L	U
rs12137855	Multiplicative	[C]	257	64%	60%	Reference	1.20	0.80	1.78
		[T]	143	36%	40%	0.3699			
LYPLAL1	Genotypes	C/C	130	65%	64%	Reference	0.88	0.47	1.59
		C/T	66	33%	29%	0.6661			
	Trend	T/T	4	2%	7%	0.0638	3.61	0.92	15.15
		Increasing [T]				0.4446	1.21	0.74	1.95
	Multiplicative	[C]	326	82%	79%	Reference	1.20	0.74	1.92
		[T]	74	19%	21%	0.4499			

L and U are the lower and upper bounds of the 95% CI.

Crude significant P-values are printed bold, and FDR-adjusted significant P-values are printed bold and red.

risk factors were tested using normal logistic, multinomial logistic and ordinal logistic regression. Histology-SNP associations were additionally analysed adjusted for age, gender, BMI z-score and HOMA-IR. A detailed description is given in the Methods S1.

2.6 | Computational modelling

We simulated stimulus implications of GCKR, UCP2 and PNPLA3 polymorphisms on hepatic glucose and fat metabolism with a metabolic model of central hepatic metabolism and hepatic lipid droplet metabolism.^{19,20} Please refer to Methods S1 for full details. For all simulations, we used MATLAB release 2012a (The MathWorks, Inc, Natick, MA).

3 | RESULTS

3.1 | Clinical study cohort characteristics

The recently established Berlin Adolescent NAFLD cohort consists of 70 patients in the age range of 10-17 years. Clinical and laboratory characteristics are summarized in Table 2. The mean age of the study group was 14.1 ± 2.2 years, and 77% (n = 54) were boys. The prevalence of extreme obesity was 70% (n = 49) with a mean BMI of 34.63 ± 6.11 kg/m² (mean BMI z-score: 2.76 ± 0.58). All patients displayed severe insulin resistance (mean HOMA-IR: 7.04 ± 3.73). Alanine aminotransferase levels were roughly 2.5-fold above the upper limit of normal (mean ALT: 108 ± 68 U/L). All patients were diagnosed with NAFLD upon histological evaluation of clinically indicated liver biopsy.

3.2 | Genetic variants associated with paediatric NAFLD in a case-control design

In an explorative analysis, six of the 14 analysed SNPs were associated with paediatric NAFLD with P-values smaller than 0.05 in the unadjusted likelihood ratio (LR) test or in the unadjusted trend test (Table 3A). Remarkably high odds ratios (OR) were obtained, 10.2 (95% CI: 4.3-25.4; $P = 2.4 \times 10^{-7}$) for the G/G genotype of rs738409 and 3.2 (95% CI: 1.6-6.1; $P = 6 \times 10^{-4}$) for the A/C variant of rs1044498 (Table 3B, see also Table S3). Haplotype analysis of PNPLA3 SNP rs738409 and a second SNP in this gene rs2294918 showed that the quite frequent high-risk group rs738409 G/G contained only rs2294918 G/G genotypes but not vice versa (Table S4). The odds ratio of this SNP combination was 12.1 (95% CI: 3.6-49.5; $P = 1 \times 10^{-4}$) and so higher as of rs738409 alone. However, adjusting effects of rs738409 to rs2294918 and vice versa showed that only rs738409 was associated with NAFLD (Table S5).

Backward selection including all SNPs in a logistic regression model indicated that four SNPs significantly differ between paediatric NAFLD patients and the normal population (Table S6). Weighted risk allele scores indicated that genetic risks add up to ORs of more than 10 for persons carrying most risk alleles compared to those with few risk alleles (Table S7).

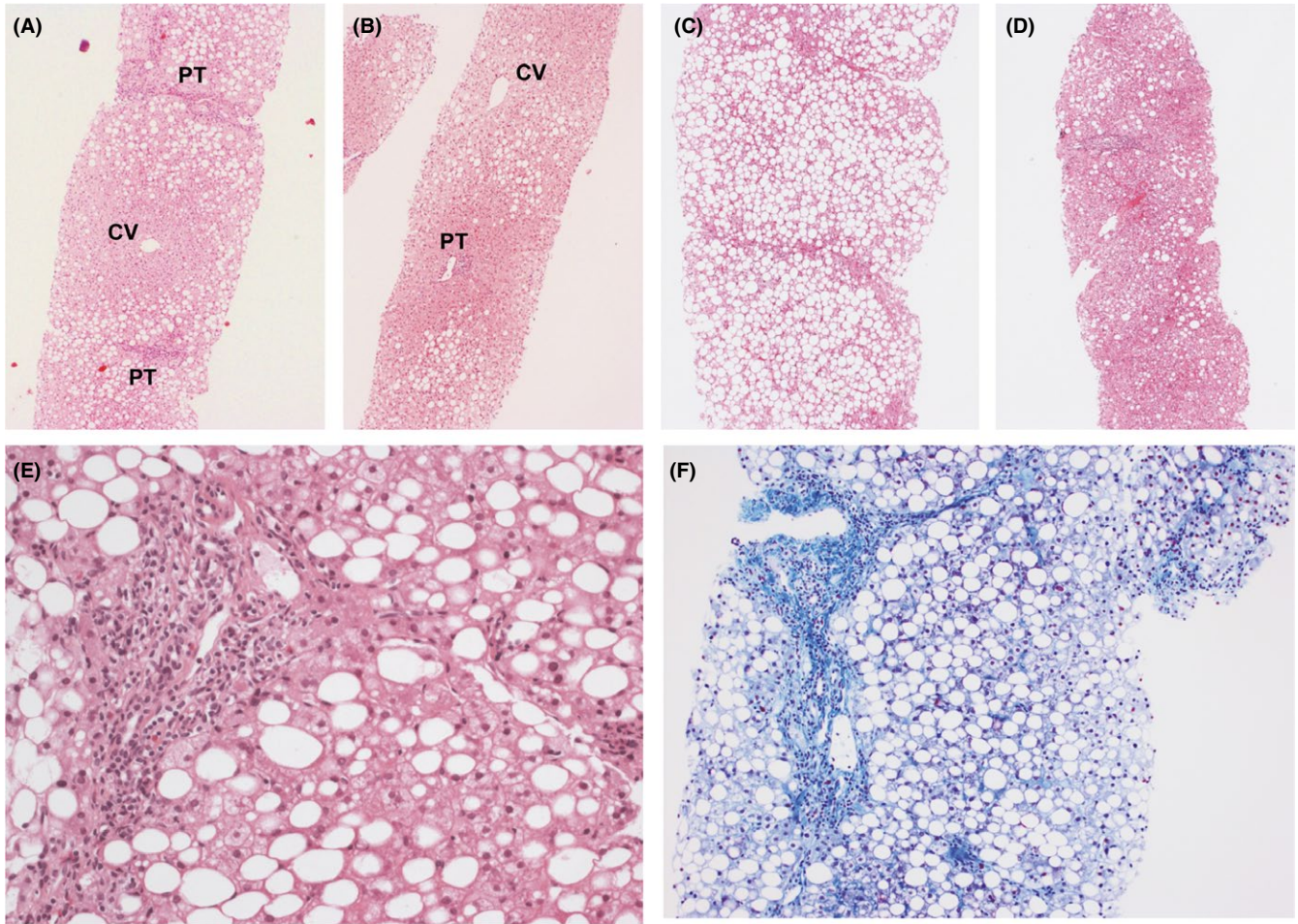


FIGURE 1 Differential steatosis zonation may present with zone 1-predominant steatosis (A), zone 3-predominant steatosis (B), panlobular steatosis (C) or azonal steatosis (D). (E) shows an enlarged portal triad with a moderately dense, mainly lymphocytic, inflammatory infiltrate, and (F) depicts fibrous periportal septa infiltrating into the parenchyma. Micrographs A-D: haematoxylin and eosin (H&E) staining, 40× magnification, E, H&E, 200× magnification, and F, chromotrope aniline blue trichrome (CAB), 100× magnification. PT, portal tract; CV, central vein

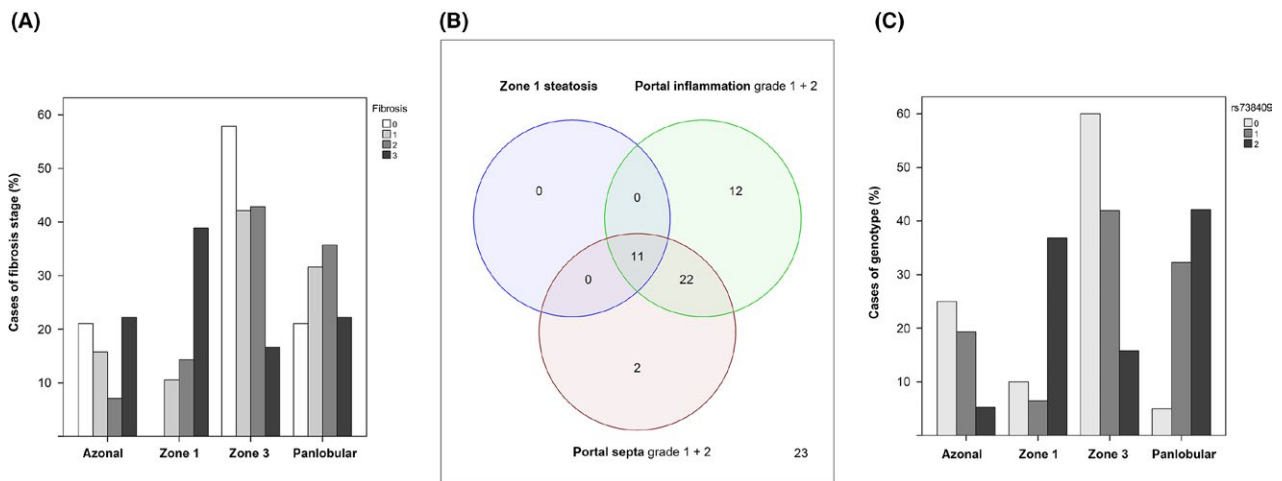


FIGURE 2 Zone 1-predominant steatosis is associated with the severity of portal inflammation, periportal septa and stage of fibrosis. A, Stratification of fibrosis stages according to steatosis zonation. B, Venn diagram for the distribution of zone 1-predominant steatosis, portal inflammation and periportal septa. C, Stratification of PNPLA3 genotype according to steatosis zonation

TABLE 4 Results of the trend test assuming an increasing or decreasing effect of the risk alleles. Italic print indicates FDR-adjusted significance for associations of SNPs with the histological features given in the topmost line. Correspondingly, *P*-values that reached significance only without FDR adjustment are printed bold (black). For *P*-values that did not reach any significance (black, not bold), neither the OR nor the 95% CIs are given. Italic print in the columns "Gene" and "SNP" indicates an FDR-adjusted significant difference between cases and controls. Correspondingly, *P*-values for comparisons between cases and controls that reached significance only with FDR adjustment are printed bold (black). *P*-values of the multiplicative model are given in Tables S10B,C

<i>P</i> -value (unadjusted) OR (95% CI)		Inflammation			NAS	
Gene	SNP	Portal	Lobular	Ballooning	All grades	<5 vs ≥5
<i>PNPLA3</i>	<i>rs738409</i>	0.0605	0.2010	0.9164	0.0262 OR = 1.90 (1.09-3.38)	0.1472
<i>PNPLA3</i>	<i>rs2294918^a</i>	0.1720	0.0361 OR = 0.44 (0.20-0.94)	0.5218	0.0193 OR = 0.46 (0.24-0.88)	0.0836
<i>GCKR</i>	<i>rs1260326</i>	0.5619	0.6376	0.2841	0.7287	0.1493
<i>GCKR</i>	<i>rs780094</i>	0.5619	0.6376	0.2841	0.7287	0.1493
<i>TM6SF2</i>	<i>rs58542926</i>	0.6832	0.6611	0.6056	0.9762	0.4601
<i>LPIN1</i>	<i>rs13412852</i>	0.8205	0.5486	0.1643	0.0466 OR = 0.54 (0.29-0.98)	0.1184
<i>UCP2</i>	<i>rs659366</i>	0.0498 OR = 2.03 (1.01-4.21)	0.1512	0.0214 OR = 2.40 (1.15-5.14)	0.0583	0.0377 OR = 2.30 (1.07-5.26)
<i>ENPP1</i>	<i>rs1044498</i>	0.5939	0.4948	0.2277	0.5541	0.1562
<i>LYPLAL1</i>	<i>rs12137855</i>	0.0217 OR = 2.51 (1.16-5.64)	0.5123	0.2683	0.8139	0.6040

Trend tests include depending on the statistical model, logistic regression, ordinal logistic regression, trend over increasing progression scores and increasing number of minor alleles.

After FDR adjustment steatosis, *P*-values of rs738409 remain significant:

^aOR and 95% CI for increasing number of minor allele [A].

^bTrend over number of risk alleles but not over zonation results in rs738409.

^cFDR-adjusted *P* = 0.0070.

^dFDR-adjusted *P* = 0.0042.

^eFDR-adjusted *P* = 0.0070.

3.3 | Histological features of disease progression

Histological features of the BaNA cohort represent a wide spectrum of NAFLD ranging from simple steatosis to progressive NASH (Table 2). Nearly half of the patients presented with severe steatosis (*n* = 33, 47.1%) and roughly a quarter had advanced fibrosis (*n* = 18, 25.7%). The risk of progression to fibrosis was

not significantly associated with BMI (*P* = 0.2814), BMI z-score (*P* = 0.3431) or sex (*P* = 0.2195; Table S8). However, an association of fibrosis with younger age could be observed (*P* = 0.0048). After multiple linear regression analysis of histological features (steatosis, lobular inflammation, portal inflammation, hepatocellular ballooning), only portal inflammation remained independently associated with fibrosis stage. While the severity of

Steatosis			Fibrosis			
All grades	Severe = S3	Zonation 1, 3, pan, azonal ^b	All stages	Any ≥ F1	Progressive ≥ F2	Periportal Septa
0.0005^c OR = 3.25 (1.72-6.52)	0.0003^d OR = 4.58 (2.14-11.32)	0.0005^e OR _{1 vs azonal} = 5.56 (1.35-25.00), OR _{1 vs 3} = 5.26 (1.61-16.67), OR _{1 vs pan} = 1.23 (0.22-6.67)	0.0888	0.3285	0.1556	0.0106 OR = 2.26 (1.23-4.35)
0.0210 OR = 0.43 (0.21-0.87)	0.0178 OR = 0.36 (0.15-0.81)	0.0442 OR _{1 vs azonal} = 0.51 (3.7 × 10 ⁻²⁹ -7.1 × 10 ²⁷), OR _{1 vs 3} = 0.31 (0.10-0.93), OR _{1 vs pan} = 1.20 (0.35-4.17)	0.0966	0.3024	0.1461	0.0239 OR = 0.42 (0.19-0.86)
0.2653	0.2364	0.1089	0.0845	0.0275 OR = 2.59 (1.15-6.36)	0.1100	0.2702
0.2653	0.2364	0.1089	0.0845	0.0275 OR = 2.59 (1.15-6.36)	0.1100	0.2702
0.9481	0.6768	0.8797	0.4438	0.3777	0.7574	0.2930
0.0489 OR=0.52 (0.27-0.99)	0.0618	0.1292	0.2608	0.2547	0.4868	0.3382
0.7795	0.7742	0.5845	0.0449 OR = 2.04 (1.02-4.15)	0.3666	0.0226 OR = 2.48 (1.17-5.64)	0.0691
0.4921	0.5775	0.6490	0.8804	0.3101	0.9594	0.5576
0.7195	0.7419	0.6609	0.6949	0.0382 OR = 3.74 (1.26-16.42)	0.5104	0.4316

steatosis was not correlated with fibrosis, a significant effect of the zonality of steatosis was evident (Figure 1). All patients with zone 1-predominant steatosis displayed also portal inflammation and any degree of fibrosis (Figure 2A,B). Moreover, the severity of steatosis, portal inflammation, periportal septa and fibrosis are distinctively aggravated in zone 1-predominant steatosis (Table S9).

3.4 | PNPLA3 modifies the zonal pattern and severity of steatosis

Two of the SNPs, rs738409 (missense of PNPLA3) and rs13412852 (intron of LPIN1), were associated with the severity of steatosis (Table 4). PNPLA3 rs738409[G] increased the risk of severe steatosis by a factor of 4.58 ($P = 0.0003$) and was also strongly associated

with steatosis zonation, increasing the odds of zone 1-predominant steatosis (FDR-adjusted $P < 0.01$; per allele OR ≥ 3 ; Table 4, Figure 2C).

3.5 | PNPLA3, GCKR and UCP2 modify the severity of fibrosis

The increase in PNPLA3 rs738409[G] risk alleles was significantly associated with advanced fibrosis (OR = 2.2; 95% CI: 1.0-5.0; $P = 0.0445$; multiplicative model, Table S10C) and the severity of periportal septa (OR = 2.3; 95% CI: 1.2-4.4; $P = 0.0106$, additive model, Table 4). The GCKR SNPs rs780094 and rs1260326, which are in linkage disequilibrium and therefore are perfectly correlated in cases, increased the risk of any fibrosis by a factor of 2.6 ($P = 0.0275$). Rs659366 (UCP2) showed an association with fibrosis as well as with portal inflammation (OR = 2.0; 95% CI: 1.0-4.2; $P = 0.0498$) and ballooning (OR = 2.4; 95% CI: 1.2-5.1; $P = 0.0214$). Moreover, rs12137855 (LYPLAL1), rs641738 (MBOAT7, TMC4) and rs1801278 (IRS1) belong to the category of inflammation/fibrosis SNPs. However, the results should be interpreted with caution, because the P -values are only slightly below 0.05 or amount to significance only for single scores (Table 4, Table S10A-C). Odds ratios for the prediction of progressive fibrosis are further increased, when a risk score compiled from a simple addition of the three SNPs with significant fibrosis modification (rs738409 PNPLA3, rs780094 GCKR and rs659366 UCP2) is applied (progressive fibrosis: per allele OR = 2.0; 95% CI: 1.3-3.5; $P = 0.0036$; Table S11).

3.6 | GCKR genotype directly affects hepatic GCKR protein levels

Proteomic analysis of liver tissue identified the direct protein products of four of the analysed SNPs, GCKR, SOD2, ENPP1 and LYPLAL1, while protein levels for the other studied genes were not detected (Figure S1). The genotype of GCKR was significantly associated with reduced hepatic protein levels (Figure 3A). Patients homozygous for the GCKR polymorphism showed a distinct decrease (fold change 6.78, $P < 0.001$) of the GCKR protein (Figure 3B). GCKR protein level was significantly associated with any ($P = 0.0379$, Welch t test), progressive ($P = 0.0134$) or advanced ($P = 0.0483$) fibrosis, as well as the severity of periportal septa ($P = 0.0269$, Welch ANOVA). To further investigate the impact of the GCKR polymorphism, a pathway enrichment analysis was performed. GSEA revealed three significantly upregulated and four significantly downregulated KEGG and reactome pathways between rs780094 T/T and C/C ($P < 0.05$ and FDR < 0.25 ; Table S12A,B). These included the increased pathway "lipid digestion mobilization and transport" and the decreased pathways "drug metabolism cytochrome P450," "regulation of glucokinase by glucokinase regulatory protein" and "glucose transport." Thus, the pathway analysis indicated a metabolic change towards hepatic insulin resistance.

3.7 | PNPLA3 variant rs738409[G] I148M is associated with altered hepatic retinol metabolism

The GSEA analysis revealed several significantly over-represented KEGG pathways, of which retinol metabolism may be of particular interest, since it represents a well-established liver function (Table S12D). Retinol metabolism was shown to be a significantly downregulated pathway in homozygous PNPLA3 148M carriers (nominal $P < 0.0001$; FDR q -value = 0.0315). Hepatic protein levels of retinol dehydrogenase 16 (RDH16) were significantly lower in patients with PNPLA3 rs738409 G/G compared to C/C carriers (mean: C/C: 1.09×10^{09} , G/G: 8.17×10^{08} , $P = 0.0226$) (Figure 3C). Hepatic RDH16 protein levels decreased with increasing fibrosis ($P = 0.0252$) and were significantly lower in patients with any fibrosis ($P = 0.0139$), progressive fibrosis ($P = 0.0055$) or advanced fibrosis ($P = 0.0229$) respectively. Hepatic levels of RBP4 were significantly increased in rs738409 G/G carriers compared to C/C genotypes (mean: C/C: 8.56×10^{08} , G/G: 1.17×10^{09} ; $P = 0.0165$; Figure 3D). While hepatic RBP4 levels were significantly increased in liver tissue of any ($P = 0.0367$) and progressive fibrosis ($P = 0.0181$), the opposite was found with regard to serum RBP4. Here, decreased serum RBP4 was significantly associated with severity of periportal septa ($P = 0.0177$) and was found to be lower in patients with any ($P = 0.0467$) or progressive fibrosis ($P = 0.0181$). Serum RBP4 levels showed a trend for a lower abundance in rs738409 G/G carriers compared to C/C genotypes without reaching statistical significance (Figure 3E).

3.8 | Metabolic implications of GCKR, UCP2 and PNPLA3 polymorphisms

As GSEA hinted to metabolic changes, we used mathematical modelling to better understand the possible metabolic implications of GCKR, UCP2 and PNPLA3 polymorphisms on hepatic glucose and fat metabolism. Using the model of Berndt et al.,²⁰ we simulated the diurnal dynamics of hepatic glucose and fat metabolism for increased glucokinase (GK) and decreased UCP2 activity (one magnitude of order, respectively, for details please refer to Methods S1). Figure 4 shows that compared to normal conditions (black line), increased GK activity (blue line) in response to decreased GCKR abundance leads to an increase in hepatic fat accumulation (Figure 4A), an increase in hepatic de novo synthesis of free fatty acids (Figure 4B) and an increase in glucose uptake (Figure 4C). While UCP2 downregulation (red line) has a similar effect on hepatic triglyceride storage (Figure 4A) and de novo lipid synthesis (Figure 4B), the hepatic glucose response remains intact (Figure 4C). To investigate the effects of altered PNPLA3 activity on hepatic lipid metabolism, we used a recent model of hepatic lipid droplet metabolism.¹⁹ As it was suggested recently that PNPLA3 binds competitively with ATGL to CGI58 on lipid droplets,²¹ we simulated PNPLA3 polymorphism by decreasing ATGL binding affinity to CGI58 to 10%. Again, simulating a diurnal cycle of lipid accumulation, Figure 4D shows that compared to normal conditions (black line), an increased PNPLA3

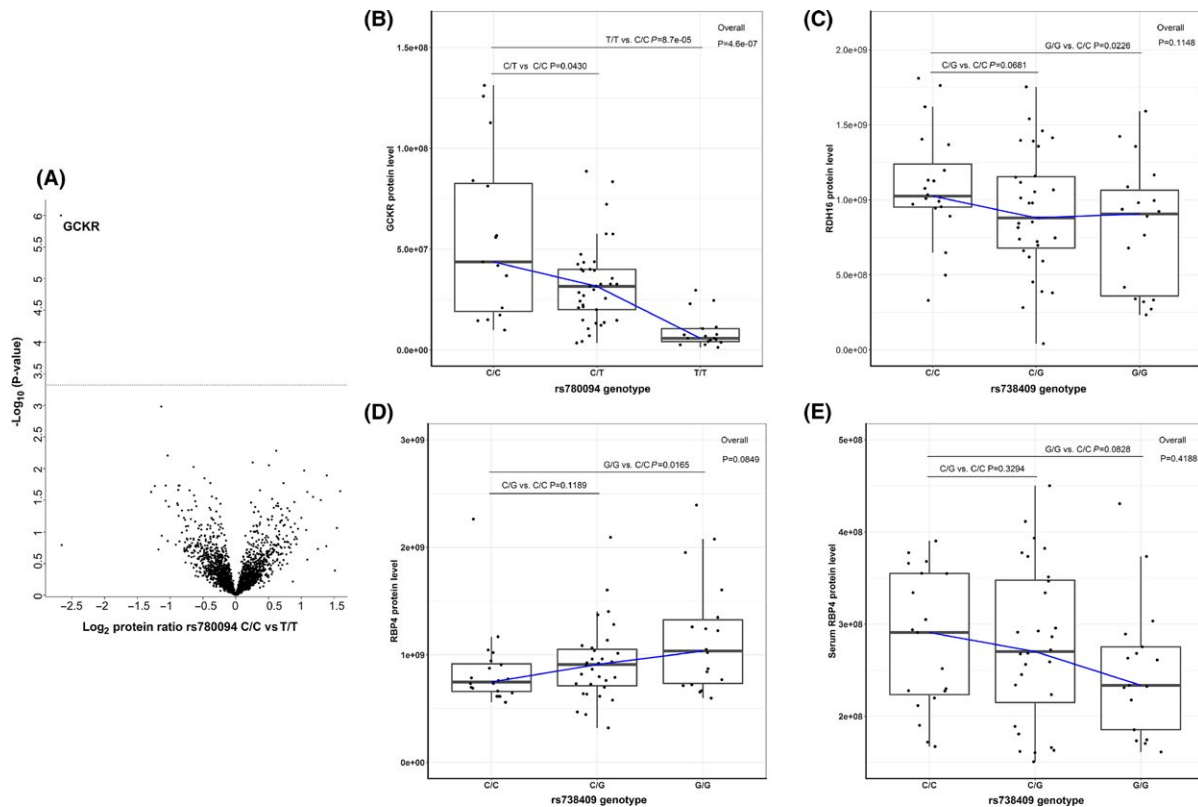


FIGURE 3 Association of GCKR genotype with hepatic protein abundance of GCKR (A, B) and association of PNPLA3 rs738409 genotype and retinol metabolism (C-E). A, Volcano plot of \log_2 protein abundance ratios of GCKR rs780094 G/G (homozygous mutant allele) vs A/A (homozygous wild-type allele) of liver tissues against the $-\log_{10}$ (P -value). Indicated is the downregulated GCKR protein. The fold change of the GCKR homozygous mutant allele vs the homozygous wild-type allele is 6.78, $P < 0.001$. Dotted line: FDR-corrected P -value of 0.05. B, GCKR protein levels depending on GCKR rs1260326 genotypes and P -values of the Welch ANOVA and pairwise Welch t tests accounting for heterogeneity of the variances. C, Hepatic RDH16 expression was decreased in rs738409 [G] carriers. P -values of the ANOVA and one-sided t tests are given. D, Hepatic RBP4 expression increased with increasing number of rs738409 [G] alleles. P -values of the Kruskal-Wallis test and one-sided Wilcoxon tests are given. E, Serum RBP4 expression was decreased in rs738409 G/G genotypes. P -values of the ANOVA and one-sided t tests are given

activity leads to severely increased hepatic triglyceride accumulation underpinning the metabolic importance of the genetic polymorphisms identified in this study.

4 | DISCUSSION

The impact of genetic determinants on the development and progression of NAFLD has become increasingly recognized throughout the last decade, identifying and validating numerous candidate gene polymorphisms with significant effect size in adult disease.^{22,23} Yet, in paediatric subjects, the number of studies analysing multiple genes in a single cohort with biopsy-proven NAFLD is limited.²⁴ The unfortunate rise of paediatric NAFLD has also shed light on some issues of adult disease. Most importantly, it has become evident that epidemiologic differences seen in adult disease, such as differences in sex and ethnical background, cannot be solely attributed to the effect of lifestyle, for example, higher alcohol consumption in males.² Consequently, investigating genetic determinants of NAFLD in children and adolescents constitutes the opportunity to elucidate the unadulterated

effects of SNPs on mechanisms of disease progression. In this study, we analysed a set of 14 previously described NAFLD SNPs in a population of paediatric patients with a broad spectrum of NAFLD disease severity upon histological characterization.

A remarkable characteristic of our study population was the high frequency of heterozygous and homozygous carriers of the PNPLA3 rs738409[G] allele. Besides being the strongest determinant for the presence of NAFLD as compared to healthy controls, PNPLA3 rs738409[G] also conferred the highest risk of severity of steatosis and modified the presence of portal fibrosis. This finding is of particular interest, since homozygous PNPLA3 rs738409[G] was also associated with a zone 1-predominant pattern of steatosis and increased amount of portal inflammation; consequently, we report a linkage of PNPLA3 rs738409[G] to portal/periportal disease localization in paediatric NAFLD. Of interest, recently, hepatic progenitor cell activation and portal/periportal histological pattern were described in adult PNPLA3 rs738409[G] carriers presenting with NASH.²⁵

PNPLA3 has been attributed a retinyl-palmitate lipase activity in hepatic stellate cells (HSC),²⁶ and stable overexpression of PNPLA3

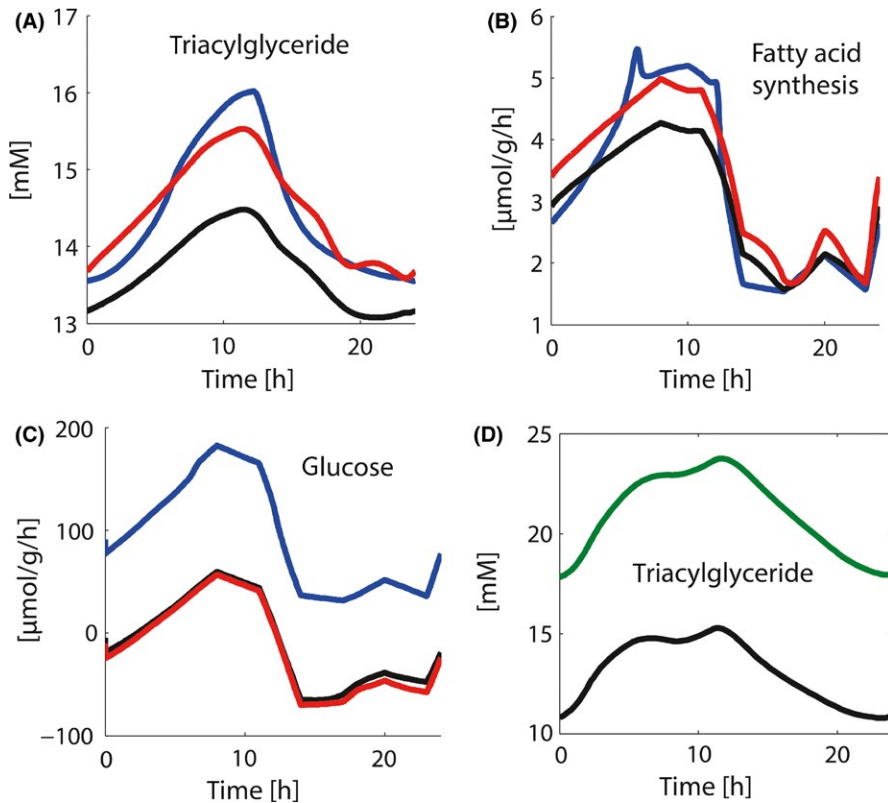


FIGURE 4 Metabolic implication of GCKR, UCP2 (A-C, Model: Ref. ²⁰) and PNPLA3 (D, Model: Ref. ¹⁹) polymorphism on hepatic metabolism. A, Hepatic triglyceride content for a diurnal cycle (GCKR blue line, UCP2 red line, wt black line). B, Hepatic de novo fatty acid synthesis for a diurnal cycle (GCKR blue line, UCP2 red line, wt black line). C, Hepatic glucose exchange for a diurnal cycle (GCKR blue line, UCP2 red line, wt black line). D, Hepatic triglyceride content for a diurnal cycle (PNPLA3 green line, wt black line)

rs738409[G] was associated with decreased retinol release from human HSC *ex vivo* and concomitantly impaired reduction in fibrogenic factors.²⁷ Recently, the rs738409[G] variant has been shown to disrupt ubiquitination and proteasomal degradation of PNPLA3, resulting in accumulation of PNPLA3-148M and impaired mobilization.²⁸ Interestingly, hepatic levels of retinol dehydrogenase 16 (RDH16), a critical microsomal regulator in the generation of all-trans-retinoic acid, were significantly lower in homozygous PNPLA3 rs738409[G] carriers than in patients with homozygous wild-type allele and directly associated with stage of fibrosis. Thus, downregulated retinol metabolism may in part be due to activation of fibrogenic HSCs.

In adult patients with NASH, high liver tissue RBP4 levels were independently associated with moderate fibrosis,²⁹ while serum RBP4 was negatively associated with fibrosis in adults and paediatric patients.^{30,31} Our findings are generally in line with the reported studies, establishing for the first time serum and hepatic association of RBP4 with fibrosis in the same cohort.

The GCKR risk genotype, attributed with distinct effects on glucose and lipid metabolism,³² has been associated with progressive fibrosis in adult NAFLD patients.³³ While imaging studies in paediatric cohorts indicated an association of NAFLD with GCKR,^{34,35} no data based on histology were available so far. Here, we show that GCKR is indeed associated with paediatric NAFLD as compared to a healthy control group. Importantly, on a histological level, GCKR affects the amount of fibrosis, while it does not account for differences in steatosis severity. Moreover, proteomic data indicate that hepatic levels of GCKR are directly dependent on GCKR genotype.

We simulated the effect of activity changes in the GCKR, UCP2 and PNPLA3 on hepatic glucose exchange flux and hepatic triglyceride accumulation. The effect of genetic polymorphisms on metabolic functions is difficult to assess for multiple reasons. Firstly, the changes on the molecular levels associated with a specific polymorphism (eg, activity changes, affinity changes, changes in the regulatory properties and changes in abundance) are largely unknown. Secondly, activity changes in one component of a molecular network can be compensated by adaptive changes in other contributors of the network masking the primary effect. Thirdly, unless the changes are massive, they will be challenging or even impossible to be assessed with experimental methods. Modelling is a useful tool to circumvent these problems as the changes in enzyme activity can be directly transferred to changes in the metabolic state. Interestingly, our simulations show that all three candidate genes are important for the regulation of fat accumulation in the liver. Furthermore, GCKR is important in the regulation of the hepatic glucose exchange flux. These findings support the hypothesis that the indicated polymorphism supports a change in hepatic metabolism towards the metabolic syndrome.

A limitation of our study is the relatively small sample size. However, the full range of histological disease severity (besides cirrhosis) could be observed in our study population, displaying a nearly even distribution of fibrosis stages. While this high-risk cohort enabled the identification of important genetic determinants of childhood NAFLD despite the limited case numbers, less strong genetic effects may not have amounted to statistical significance.

For example, the TM6SF2 variant is a well-established modifier of adult NAFLD,³⁶ which has repeatedly been shown to be

associated with liver fat and risk of fibrotic progression in large cohorts (Table S1). While we found TM6SF2 to be significantly (unadjusted) associated with paediatric NAFLD as compared to adult controls, no association with individual histological features was noted in our cohort. Yet, it remains elusive, whether this was due to our limited sample size (eg, a low frequency of the minor allele) or attributed to differing underlying mechanisms in paediatric disease.

A particular strength of our study was the employment of multiple methodologies upon the same clinically well-characterized patient cohort samples, ranging from detailed histopathological characterization to liver proteomic analysis.

Taken together, different SNPs modulate the progression of steatosis and fibrosis in paediatric NAFLD. The PNPLA3 SNP rs738409 indicates the highest risk of steatosis, while the GCKR SNPs rs1260326/rs780094 and the UCP2 SNP rs659366 represent the strongest determinants of fibrosis severity. PNPLA3 rs738409 confers periportal preference of paediatric NAFLD and is therefore of particular clinical relevance for the identification of patients at risk of early progression of disease.

CONFLICT OF INTEREST

All authors declare no competing interests and have approved the final article.

AUTHOR CONTRIBUTIONS

CAH, JH and SW conceived and designed the study. CAH and JH drafted the manuscript. SS performed statistical analysis and contributed to the manuscript. KG and JR performed and analysed SNP profiles. BR, HB and CL analysed and scored histology. RT and DM performed and analysed proteomic data and contributed to the manuscript. NB and HH performed computational studies and contributed to the manuscript. SH, CC, PB and PLMJ critically revised the manuscript.

ORCID

Christian A. Hudert  <https://orcid.org/0000-0002-3642-0838>

REFERENCES

- Nobili V, Alisi A, Newton KP, Schwimmer JB. Comparison of the phenotype and approach to pediatric vs adult patients with nonalcoholic fatty liver disease. *Gastroenterology*. 2016;150(8):1798-1810.
- Schwimmer JB, Deutsch R, Kahen T, Lavine JE, Stanley C, Behling C. Prevalence of fatty liver in children and adolescents. *Pediatrics*. 2006;118(4):1388-1393.
- Browning JD, Szczepaniak LS, Dobbins R, et al. Prevalence of hepatic steatosis in an urban population in the United States: impact of ethnicity. *Hepatology*. 2004;40(6):1387-1395.
- Makkonen J, Pietilainen KH, Rissanen A, Kaprio J, Yki-Jarvinen H. Genetic factors contribute to variation in serum alanine aminotransferase activity independent of obesity and alcohol: a study in monozygotic and dizygotic twins. *J Hepatol*. 2009;50(5):1035-1042.
- Loomba R, Schork N, Chen CH, et al. Heritability of hepatic fibrosis and steatosis based on a prospective twin study. *Gastroenterology*. 2015;149(7):1784-1793.
- Anstee QM, Seth D, Day CP. Genetic factors that affect risk of alcoholic and nonalcoholic fatty liver disease. *Gastroenterology*. 2016;150(8):1728-1744.e7.
- Diehl AM, Day C. Cause, pathogenesis, and treatment of nonalcoholic steatohepatitis. *N Engl J Med*. 2017;377(21):2063-2072.
- Crespo M, Lappe S, Feldstein AE, Alkhoury N. Similarities and differences between pediatric and adult nonalcoholic fatty liver disease. *Metabolism*. 2016;65(8):1161-1171.
- Schwimmer JB, Behling C, Newbury R, et al. Histopathology of pediatric nonalcoholic fatty liver disease. *Hepatology*. 2005;42(3):641-649.
- Africa JA, Behling CA, Brunt EM, et al. In children with nonalcoholic fatty liver disease, zone 1 steatosis is associated with advanced fibrosis. *Clin Gastroenterol Hepatol*. 2018;16:438-446.e2.
- Mann JP, De Vito R, Mosca A, et al. Portal inflammation is independently associated with fibrosis and metabolic syndrome in pediatric nonalcoholic fatty liver disease. *Hepatology*. 2016;63(3):745-753.
- Marzuillo P, Del Giudice EM, Santoro N. Pediatric non-alcoholic fatty liver disease: new insights and future directions. *World J Hepatol*. 2014;6(4):217-225.
- Selinski S, Blaszkewicz M, Ickstadt K, Hengstler JG, Golka K. Refinement of the prediction of N-acetyltransferase 2 (NAT2) phenotypes with respect to enzyme activity and urinary bladder cancer risk. *Arch Toxicol*. 2013;87(12):2129-2139.
- Kleiner DE, Brunt EM, Van Natta M, et al. Design and validation of a histological scoring system for nonalcoholic fatty liver disease. *Hepatology*. 2005;41(6):1313-1321.
- Saravana Devi S, Vinayagamoorthy N, Agrawal M, et al. Distribution of detoxifying genes polymorphism in Maharastrian population of central India. *Chemosphere*. 2008;70(10):1835-1839.
- Gielisch I, Meierhofer D. Metabolome and proteome profiling of complex I deficiency induced by rotenone. *J Proteome Res*. 2015;14(1):224-235.
- Cox J, Mann M. MaxQuant enables high peptide identification rates, individualized p.p.b.-range mass accuracies and proteome-wide protein quantification. *Nat Biotechnol*. 2008;26(12):1367-1372.
- Subramanian A, Tamayo P, Mootha VK, et al. Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *Proc Natl Acad Sci U S A*. 2005;102(43):15545-15550.
- Wallstab C, Eleftheriadou D, Schulz T, et al. A unifying mathematical model of lipid droplet metabolism reveals key molecular players in the development of hepatic steatosis. *FEBS J*. 2017;284(19):3245-3261.
- Berndt N, Bulik S, Wallach I, et al. HEPATOKIN1 is a biochemistry-based model of liver metabolism for applications in medicine and pharmacology. *Nat Commun*. 2018;9(1):2386.
- Wu JW, Yang H, Mitchell GA. Potential mechanism underlying the PNPLA3(I) (148) (M) -Hepatic steatosis connection. *Hepatology*. 2016;63(2):676-677.
- Anstee QM, Day CP. The genetics of nonalcoholic fatty liver disease: spotlight on PNPLA3 and TM6SF2. *Semin Liver Dis*. 2015;35(3):270-290.
- Eslam M, Valenti L, Romeo S. Genetics and epigenetics of NAFLD and NASH: clinical impact. *J Hepatol*. 2018;68(2):268-279.
- Goyal NP, Schwimmer JB. The genetics of pediatric nonalcoholic fatty liver disease. *Clin Liver Dis*. 2018;22(1):59-71.

25. Carpino G, Pastori D, Baratta F, et al. PNPLA3 variant and portal/periportal histological pattern in patients with biopsy-proven non-alcoholic fatty liver disease: a possible role for oxidative stress. *Sci Rep*. 2017;7(1):15756.
26. Pirazzi C, Valenti L, Motta BM, et al. PNPLA3 has retinyl-palmitate lipase activity in human hepatic stellate cells. *Hum Mol Genet*. 2014;23(15):4077–4085.
27. Pingitore P, Dongiovanni P, Motta BM, et al. PNPLA3 overexpression results in reduction of proteins predisposing to fibrosis. *Hum Mol Genet*. 2016;25(23):5212–5222.
28. 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(4):1111–1124.
29. Petta S, Tripodo C, Grimaudo S, et al. High liver RBP4 protein content is associated with histological features in patients with genotype 1 chronic hepatitis C and with nonalcoholic steatohepatitis. *Dig Liver Dis*. 2011;43(5):404–410.
30. Alkhouri N, Lopez R, Berk M, Feldstein AE. Serum retinol-binding protein 4 levels in patients with nonalcoholic fatty liver disease. *J Clin Gastroenterol*. 2009;43(10):985–989.
31. Nobili V, Alkhouri N, Alisi A, et al. Retinol-binding protein 4: a promising circulating marker of liver damage in pediatric nonalcoholic fatty liver disease. *Clin Gastroenterol Hepatol*. 2009;7(5):575–579.
32. Raimondo A, Rees MG, Gloyn AL. Glucokinase regulatory protein: complexity at the crossroads of triglyceride and glucose metabolism. *Curr Opin Lipidol*. 2015;26(2):88–95.
33. 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(2):e87523.
34. Santoro N, Zhang CK, Zhao H, et al. Variant in the glucokinase regulatory protein (GCKR) gene is associated with fatty liver in obese children and adolescents. *Hepatology*. 2012;55(3):781–789.
35. Lin YC, Chang PF, Chang MH, Ni YH. Genetic variants in GCKR and PNPLA3 confer susceptibility to nonalcoholic fatty liver disease in obese individuals. *Am J Clin Nutr*. 2014;99(4):869–874.
36. 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(4):352–356.

SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section at the end of the article.

How to cite this article: Hudert CA, Selinski S, Rudolph B, et al. Genetic determinants of steatosis and fibrosis progression in paediatric non-alcoholic fatty liver disease. *Liver Int*. 2019;39:540–556. <https://doi.org/10.1111/liv.14006>