

1
2
3 453 25. Devlin B, Roeder K (1999) Genomic control for association studies. *Biometrics* 55:
4
5
6 454 997–1004.
7
8
9
10 455 26. Barrett JC, Fry B, Maller J, Daly MJ (2005) Haploview: analysis and visualization of
11
12 LD and haplotype maps. *Bioinformatics* 21: 263 –265. doi:10.1093/bioinformatics/bth457.
13 456
14
15
16 457
17
18
19 458
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61 30
62
63
64
65

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65

459 **Legends to Figures**

460

461 **Figure 1. Manhattan plot of the GWA study.** Association p -values are calculated by exact
462 trend test and plotted along the chromosome in $-\log_{10}$ scale. The horizontal line indicates
463 Bonferroni-adjusted significance threshold ($p=1.03 \times 10^{-7}$).

464

465 **Figure 2. A schematic organization of the human *PNPLA3* locus at 22q13.31 with the**
466 **genome scan results.** P -values calculated by the exact trend test were plotted in $-\log_{10}$ scale.
467 Red and blue dots indicate the p -values of genotyped and imputed SNPs, respectively. Local
468 recombination rate obtained from HAPMAP release 22 is indicated by a red line plotted in
469 cm/Mb scale. The structure and orientation of four genes in the region are shown below the
470 plots with their transcriptional orientations according to NCBI Reference Sequence Build 36.3.
471 LD blocks were generated according to pairwise LD estimates of the SNPs located within the
472 region using the genome scan results. The LD block showing the strongest association is
473 highlighted with the triangle, and the corresponding chromosomal region is represented by the
474 dotted lines.

475

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65

Figure 3. Histogram of odds ratios for genotype distribution of rs738409 between
Matteoni types. Each box denotes the odds ratio (OR) comparing the corresponding Matteoni
types on the horizontal axes. N represents the number of samples. Odds ratios and
p-values are calculated for the higher Matteoni type per risk allele (G) on additive model by
multivariable logistic regression adjusted for age, sex and BMI, and are shown with 95% CI
above each box.

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49

205

Table 1. Clinical characteristics according to the histological classification

Phenotype	Matteoni classification of NAFLD				Control	<i>p</i> -value
	Type 1	Type 2	Type 3	Type 4		
Number of samples	100	73	29	327	932	
Sex (Male/Female)	59/41	47/26	13/16	130/197	471/461	0.0023‡
Age (year)	49.7 ± 15.3	51.5 ± 15.3	49.4 ± 14.0	57.6 ± 14.8	48.8 ± 16.3	<0.001
Physical measurement						
BMI	26.2 ± 4.3	27.7 ± 4.8	27.6 ± 3.5	27.7 ± 5.2	-	0.054
Amount of visceral fat (cm ²)	146.8 ± 65.3	154.3 ± 47.7	136.8 ± 53.8	151.7 ± 57.4	-	0.46
Abdominal circumscript (cm)	90.9 ± 9.9	94.1 ± 10.0	88.5 ± 10.2	94.1 ± 11.8	-	0.10
Biochemical trait						
AST (IU/L)	31.1 ± 14.6	36.4 ± 18.5	52.4 ± 35.1	57.7 ± 48.4	-	<0.001
ALT (IU/L)	48.6 ± 30.8	62.8 ± 47.6	81.5 ± 46.9	74.9 ± 48.4	-	<0.001
GGT (IU/L)	71.0 ± 62.5	67.1 ± 66.9	96.1 ± 91.3	76.6 ± 73.9	-	0.25
Albumin (g/dL)	4.5 ± 0.4	4.4 ± 0.3	4.5 ± 0.3	4.3 ± 0.4	-	<0.001
Total bilirubin (mg/dL)	0.9 ± 0.5	0.9 ± 0.5	0.9 ± 0.6	0.8 ± 0.4	-	0.063
Cholinesterase (unit)	389.1 ± 97.0	354.3 ± 97.2	371.1 ± 109.9	348.9 ± 93.2	-	<0.001
Type IV collagen 7S (ng/dL)	3.8 ± 0.7	3.9 ± 0.9	3.9 ± 0.8	5.1 ± 1.7	-	<0.001
Hyaluronic acid (ng/dL)	25.6 ± 22.5	33.6 ± 29.5	31.5 ± 24.0	80.9 ± 84.3	-	<0.001
Triglycerides (mg/dL)	151.9 ± 73.8	154.0 ± 92.1	166.1 ± 86.5	161.2 ± 85.7	-	0.23
Total cholesterol (mg/dL)	209.1 ± 32.8	194.0 ± 38.0	203.0 ± 39.9	200.3 ± 39.0	-	0.093
HbA1c (%)	6.1 ± 1.1	5.9 ± 1.2	6.5 ± 1.8	6.2 ± 1.3	-	0.13
IRI (µg/dL)	9.1 ± 5.4	11.4 ± 9.0	10.4 ± 6.3	14.9 ± 9.9	-	<0.001

1						
2						
3						
4						
5						
6	FPG (mg/dL)	112.9 ± 33.7	107.3 ± 27.4	109.9 ± 27.7	114.8 ± 33.8	- 0.14
7	HOMA-IR	2.4 ± 1.5	2.9 ± 2.4	3.0 ± 2.1	4.2 ± 3.0	- <0.001
8	hs-CRP (mg/dL)	1078.9 ± 1407	1048.3 ± 1185.0	865.8 ± 658.4	1579.2 ± 2377.9	- 0.027
9	Adiponectin (µg/mL)	7.4 ± 4.4	8.5 ± 6.6	6.6 ± 2.6	6.9 ± 4.3	- 0.24
10	Leptin (ng/mL)	9.9 ± 7.4	9.1 ± 6.2	11.3 ± 9.4	12.4 ± 7.9	- <0.001
11	Ferritin (ng/mL)	145.8 ± 101.1	176.5 ± 134.0	271.2 ± 307.0	208.3 ± 180.3	- 0.027
12	Uric acid (mg/dL)	5.9 ± 1.5	5.7 ± 1.2	5.4 ± 1.9	5.7 ± 1.6	- 0.77
13	PLT (x10 ⁴ /µL)	23.0 ± 5.9	22.9 ± 4.9	21.9 ± 6.7	20.2 ± 6.4	- <0.001
14	ANA (0/1/2/3/4)	42/17/4/0/0	31/8/4/1/2	15/6/2/0/0	147/76/31/8/12	- 0.015
15	Clinical history					
16	Diabetes (NGT/IGT/DM)	36/11/34	24/7/27	12/8/7	103/35/119	- 0.45*
17	Hyperlipidemia (+/-)	31/68	31/42	9/20	120/206	- 0.60‡
18	Hypertension (+/-)	64/35	33/40	19/10	155/172	- 0.013‡
19	Liver biopsy feature					
20	Brunt grade (1/2/3)	-	-	19/3/2	149/133/44	- <0.001‡
21	Brunt stage (1/2/3/4)	-	-	-	123/74/105/24	- -
22	Fat droplet (1/2/3/4)	38/32/19/11	14/29/18/7	7/3/10/4	51/99/104/52	- <0.001
23	Iron deposition (0/1/2/3/4)	30/14/21/10/1	24/9/12/2/1	10/5/2/2/0	132/56/29/29/11	- 0.16

Measurements are shown as mean ± standard deviation. Categorical values are shown by the count number. *P*-values are calculated by Jonckheere-Terpstra test unless otherwise stated; ‡Chochran-Armitage trend test, *Kruskal-Wallis test. Abbreviations used for each trait are summarized in materials and methods.

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49

207

Table 2. List of the SNP markers in the *PNPLA3* locus showing genome wide significance

dbSNPID	Chr	Allele		Genotyping result						Statistics		
		A1	A2	Control	Total	Type 1	Type 2	Type 3	Type 4	NAFLD vs. Control	Matteoni	
										<i>p</i> -value†	OR (95%CI)	<i>p</i> -value‡
rs738407	22	T	C*	124/447/361 (0.627)	46/200/283 (0.724)	12/51/37 (0.625)	10/28/35 (0.671)	4/14/11 (0.621)	20/107/200 (0.775)	1.0 x 10 ⁻⁷	1.56 (1.32-1.83)	3.4 x 10 ⁻⁵
rs738409	22	C	G*	247/468/217 (0.484)	88/236/203 (0.609)	20/59/21 (0.505)	21/30/22 (0.507)	8/11/9 (0.518)	39/136/151 (0.672)	1.4 x 10 ⁻¹⁰	1.66 (1.43-1.94)	3.6 x 10 ⁻⁶
rs2076211	22	C	T*	248/473/211 (0.480)	92/242/195 (0.597)	21/58/21 (0.500)	21/30/22 (0.507)	8/11/10 (0.534)	42/143/142 (0.653)	1.4 x 10 ⁻⁹	1.61 (1.38-1.87)	3.2 x 10 ⁻⁵
rs2896019	22	T	G*	246/473/213 (0.482)	91/234/204 (0.607)	20/57/23 (0.515)	22/29/22 (0.500)	7/12/10 (0.552)	42/136/149 (0.664)	1.5 x 10 ⁻¹⁰	1.66 (1.42-1.93)	2.6 x 10 ⁻⁵
rs1010023	22	T	C*	249/473/210 (0.479)	94/239/196 (0.596)	21/57/22 (0.505)	22/29/22 (0.500)	7/12/10 (0.552)	44/141/142 (0.650)	1.5 x 10 ⁻⁹	1.61 (1.38-1.87)	6.5 x 10 ⁻⁵
rs926633	22	G	A*	247/474/211 (0.481)	93/237/199 (0.600)	21/56/23 (0.510)	22/29/22 (0.500)	7/12/10 (0.552)	43/140/144 (0.654)	7.5 x 10 ⁻¹⁰	1.62 (1.39-1.89)	5.8 x 10 ⁻⁵
rs3810622	22	T*	C	330/445/157 (0.407)	263/208/58 (0.306)	40/48/12 (0.360)	28/29/16 (0.418)	14/12/3 (0.310)	181/119/27 (0.265)	1.0 x 10 ⁻⁷	0.64 (0.55-0.75)	0.0017

Reference (A1) and non-reference (A2) alleles refer to NCBI Reference Sequence Build 36.3 with the effective allele marked by an asterisk. Genotyping results are shown by genotype count of A1A1/A1A2/A2A2 with allele frequency of A2 in parenthesis. †*P*-values are calculated by exact trend test with odds ratios (OR) calculated for A2 with 95% confidence interval (CI). ‡*P*-values are calculated by Jonckheere-Terpstra test in NAFLD patients for Matteoni type and additive model of genotype. SNPs are ordered by chromosomal location.

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49

Table 3. Association of rs738409 with clinical traits

A. Biochemical traits

Phenotype	Statistical calculation 1		Statistical calculation 2	
	Coef. (S.E.)	<i>p</i> -value	Coef. (S.E.)	<i>p</i> -value
AST (IU/L)	0.22 (0.056)	1.2 x 10⁻⁴	0.11 (0.052)	0.038
ALT (IU/L)	0.19 (0.058)	0.0016	0.093 (0.056)	0.098
GGT (IU/L)	-0.056 (0.061)	0.37	-0.088 (0.062)	0.16
Albumin (g/dL) *	0.015 (0.051)	0.77	-0.012 (0.052)	0.81
Total bilirubin (mg/dL)	-0.011 (0.063)	0.86	0.0059 (0.064)	0.93
Cholinesterase (unit) *	0.062 (0.040)	0.12	0.069 (0.041)	0.092
Type IV collagen 7S (ng/dL) *	-0.19 (0.064)	0.0025	-0.11 (0.062)	0.069
Hyaluronic acid (ng/dL)	0.30 (0.065)	4.9 x 10⁻⁶	0.22 (0.063)	4.6 x 10⁻⁴
Triglycerides (mg/dL)	-0.10 (0.058)	0.072	-0.15 (0.059)	0.013
Total cholesterol (mg/dL)	-0.066 (0.060)	0.27	-0.057 (0.061)	0.34
HbA1c (%)	-0.17 (0.053)	0.0012	-0.18 (0.054)	0.0011
IRI (μg/dL)	0.16 (0.063)	0.012	0.086 (0.061)	0.16
FPG (mg/dL)	-0.14 (0.049)	0.0047	-0.15 (0.05)	0.0035
HOMA-IR	0.084 (0.064)	0.19	0.0092 (0.062)	0.88
Hs-CRP (mg/dL)	-0.013 (0.048)	0.79	-0.031 (0.049)	0.52
Adiponectin (μg/mL)	0.048 (0.066)	0.47	0.12 (0.066)	0.072
Leptin (ng/mL)	0.11 (0.068)	0.11	0.10 (0.069)	0.15
Ferritin (ng/mL)	0.031 (0.047)	0.51	-0.0042 (0.048)	0.93
Uric acid (mg/dL)	-0.097 (0.061)	0.11	-0.11 (0.062)	0.067
PLT (x10 ⁴ /μL)	-0.056 (0.020)	0.0052	-0.045 (0.020)	0.028

B. Immunological and histological traits

Phenotype	Statistical calculation 1		Statistical calculation 2	
	OR (95%CI)	<i>p</i> -value	OR (95%CI)	<i>p</i> -value
ANA (0/1/2/3/4)	0.92 (0.70-1.21)	0.56	0.86 (0.65-1.15)	0.31
Brunt grade (1/2/3)	1.42 (1.06-1.92)	0.021	1.38 (1.02-1.87)	0.036
Brunt stage (1/2/3/4)	1.28 (0.95-1.72)	0.11		
Fat deposition (1/2/3/4)	1.44 (1.15-1.81)	0.0019	1.24 (0.98-1.56)	0.76
Iron deposition (0/1/2/3/4)	0.61 (0.47-0.80)	3.0 x 10⁻⁴	0.62 (0.47-0.81)	5.6 x 10⁻⁴

Associations between distribution of rs738409 genotypes and clinical traits are calculated by multivariable regression. Statistical calculation 1 is adjusted for age, sex and BMI, while the Matteoni types are additionally included as covariate in statistical calculation 2.

A) Statistics are calculated for biochemical traits by multivariable linear regression.

B) Statistics are calculated for immunological and histological traits by multivariable ordinal logistic regression.

Coefficients and odds ratios are calculated for the increase of each trait per risk allele (G). The *p*-values showing significance after Bonferroni's correction for multiple testing ($p=0.0021$) was shown in bold. *Reciprocal numbers are used for normalization and a negative coefficient implicates an increase in value according to the increase of the risk allele.

Table 4. Replication study of previously reported SNPs

dbSNPID	Chr	Allele		Gene	Genotyping result						Statistics		
		A1	A2		Control	NAFLD					NAFLD vs. Control		Matteoni
						Total	Type1	Type2	Type3	Type4	<i>p</i> -value†	OR (95%CI)	<i>p</i> -value‡
rs12137855	1	C*	T	<i>LYPLAL1</i>	828/102/2 (0.056)	475/54/0 (0.051)	90/10/0 (0.050)	67/6/0 (0.041)	24/5/0 (0.086)	294/33/0 (0.050)	0.55	0.89 (0.64-1.25)	0.98
rs780094	2	T*	C	<i>GCKR</i>	321/433/178 (0.423)	212/238/79 (0.374)	34/54/12 (0.390)	28/34/11 (0.383)	17/11/1 (0.224)	133/139/55 (0.381)	0.011	0.82 (0.70-0.95)	0.92
rs4240624	8	G	A	<i>PPP1R3B</i>	-	-	-	-	-	-	-	-	-
rs2228603	19	C	T*	<i>NCAN</i>	842/88/2 (0.049)	478/47/4 (0.051)	93/7/0 (0.035)	65/8/0 (0.054)	28/1/0 (0.017)	292/31/4 (0.059)	0.80	1.05 (0.75-1.48)	0.58

Reference (A1) and non-reference (A2) alleles refer to NCBI Reference Sequence Build 36.3 with the effective allele marked by an asterisk. Genotyping results are shown by genotype count of A1A1/A1A2/A2A2 with allele frequency of A2 in parenthesis. †*P*-values are calculated by exact trend test with odds ratios (OR) calculated for A2 with 95% confidence interval (CI). ‡*P*-values are calculated by Jonckheere-Terpstra test in NAFLD patients for Matteoni type and additive model of genotype.

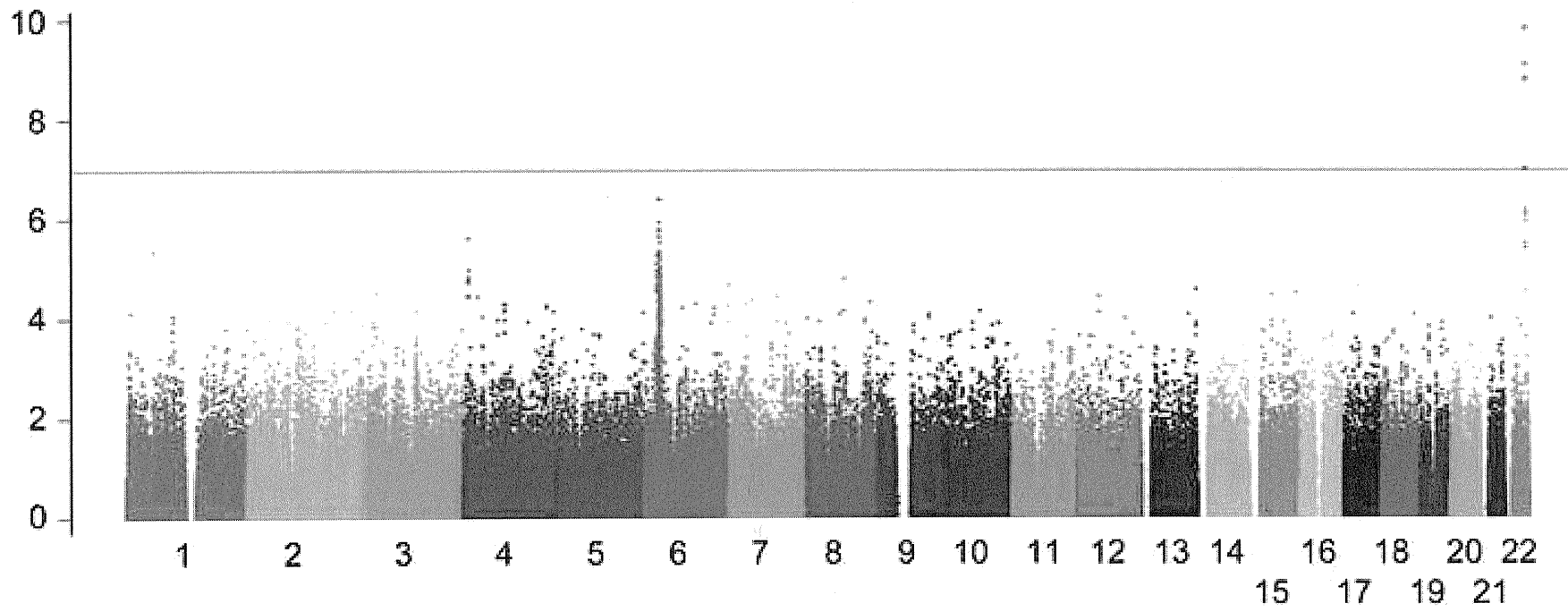


Figure
Click here to download high resolution image

