OPEN

Pediatric Non-Alcoholic Fatty Liver Disease Is Affected by Genetic Variants Involved in Lifespan/Healthspan

*Annalisa Crudele, †Serena Dato, ‡Oriana Lo Re, ‡§Andrea Maugeri, ‡Paola Sanna, ‡Sebastiano Giallongo, ‖Jude Oben, *Nadia Panera, †Francesco De Rango, ¶ Antonella Mosca, †Giuseppina Rose, †Giuseppe Passarino, *Anna Alisi, and ‡‖#Manlio Vinciguerra

See ''Identifying the Genetics Underlying Nonalcoholic Fatty Liver Disease: A Quest That is Far From Over'' by Koot and Jansen on page 139.

ABSTRACT

Objectives: Non-alcoholic fatty liver disease (NAFLD) is the leading cause of chronic liver disease in both adults and children. Along with obesity and metabolic syndrome, genetic predisposition influences the progression of NAFLD. Here, we investigated the effect of lifespan/healthspan-related single nucleotide polymorphisms (SNPs) on metabolically associated fatty liver disease in children.

Methods: We evaluated the impact of 10 SNPs involved in both human liver/metabolic diseases and healthspan (interleukin-6 [IL-6] rs1800795, antisense non coding RNA in the *INK4* locus (ANRIL) rs1556516, SH2B3/ATXN2 rs7137828, FURIN rs17514846, TP53 rs1042522, APOC3 rs2542052, KL rs9536314, KL rs9527025, SIRT6 rs107251, FOXO3 rs2802292) on NAFLD-related metabolic and liver features in 177 pediatric patients with biopsy-proven NAFLD, by comparing them to 146 healthy controls. We then applied a multidimensional reduction (MDR) case–control analysis of SNP–SNP interactions, to identify the joint effect of analyzed SNPs in predicting NAFLD and associated features.

Results: Discrete SNPs were significantly associated with individual metabolic NAFLD features, but none of them significantly associated with NAFLD diagnosis. By testing potential synergies using the MDR approach, the best combination to diagnose NAFLD ($P = 0.0011$) resulted in the one encompassing IL-6 rs1800795 and ANRIL rs1556516. Consistently, the risk combinations suggested by SNP–SNP analysis strongly associated with a higher level of fasting plasma blood glucose level ($P = 0.024$).

Conclusion: In conclusion, here we demonstrated a synergic interaction between IL-6 rs1800795 and ANRIL rs1556516 in the diagnosis of NAFLD, and NAFLD-associated hyperglycemia in children. Larger studies are required to confirm our findings and to elucidate mechanisms by which the genetic interaction between these two genes influences healthspan in pediatric NAFLD.

Key Words: children, glucose metabolism, lifespan, non-alcoholic fatty liver disease, single nucleotide polymorphism

What Is Known

- Non-alcoholic fatty liver disease (NAFLD) is the leading cause of chronic liver disease in both adults and children.
- Both longevity and NAFLD have a known genetic component, however, it is unknown whether genes involved in healthspan/lifespan are implicated in NAFLD progression in children.

What Is New

- Discrete single nucleotide polymorphisms (SNPs) were significantly associated with individual metabolic NAFLD features, but none of them significantly associated with NAFLD diagnosis.
- A synergic interaction between interleukin-6 rs1800795 and ANRIL rs1556516 in the diagnosis of NAFLD, and NAFLD-associated hyperglycemia in children.

One of the most common chronic non-communicable diseases is an obesity-associated non-alcoholic fatty liver disease (NAFLD) (1,2). The histologic spectrum of NAFLD ranges from steatosis through non-alcoholic steatohepatitis (NASH), fibrosis and hepatocellular carcinoma (HCC) (3,4). The recent increase in the prevalence of NAFLD in the pediatric population is particularly worrying for the impact on several chronic diseases later in life (5). NAFLD is a multifactorial disease where behavioral and genetic factors interact. Several genome-wide association studies (GWAS) have identified several single nucleotide polymorphisms (SNPs), such as *PNPLA3*, *TM6SF2*, and *Klotho*, associated with a higher NAFLD risk (6–13). However, other unidentified genetic variants may determine the complex heritable pathological pattern associated with NAFLD. In this respect, increased healthspan, the

major determinant of lifespan, has a strong genetic component (14–17). It has been advocated that nutritional programs based on activation of anti-aging signaling pathways, should be administered to young individuals to increase their healthspan later in life (18). It is currently unknown if gene variants predisposing to a longer lifespan are associated with protection from pediatric NAFLD.

Here, we evaluated the impact of 10 gene variants, including interleukin (IL)-6 rs1800795, antisense non-coding RNA in the INK4 locus (ANRIL) rs1556516, Ataxin-2 (ATXN2) rs7137828, FES Upstream Region (FURIN) rs17514846, TP53 rs1042522, apolipoprotein C3 (APOC3) rs2542052, Klotho (KL) rs9527025 and rs9536314, sirtuin-6 (SIRT6) rs107251, and longevity-associated Forkhead Box O3 (FOXO3) rs2802292)—which were carefully selected from manually curated literature for their reported involvement in lifespan/healthspan and liver diseases (19–31) on NAFLD-related pathological features in a cohort of pediatric patients with biopsy-proven NAFLD, by comparing them to healthy controls. These genes perform diverse functions: IL-6 is a multifunctional interleukin modulating inflammation; ANRIL is a long non-coding RNA implicated in multiple diseases (19); ATXN2 is a regulator of mRNA translation involved in neurological disorders; FURIN is a ubiquitous endoprotease (20); TP53 is a master tumor suppressor (21,23); apolipoprotein C3 is a regulator of triglyceride-rich lipoproteins; klotho is a type-I membrane protein that is related to β-glucuronidases (25,26); SIRT6 is a stress-responsive protein deacetylase and mono-ADP ribosyltransferase enzyme (27); FOXO3 belongs to the O subclass of the forkhead family of transcription factors which are characterized by a distinct forkhead DNA-binding domain, which regulates inflammation and aging (28). We then applied a bioinformatics analysis to identify the synergistic effect of SNPs related to lifespan/healthspan in predicting NAFLD-associated liver and metabolic features.

METHODS

Study Participants

The study cohort included 177 Italian pediatric (mean age 13.7 years) patients with biopsy-proven NAFLD, evaluated at Bambino Gesù Children's Hospital between January 2018 and June 2019. Other causes of liver disease including viral and autoimmune hepatitis, hereditary hemochromatosis, alpha1-antitrypsin deficiency, and history, Wilson disease, and infection with hepatitis B or hepatitis C were excluded. Body mass index (BMI) and waist circumference were measured using standard procedures. Alanine aminotransferase (ALT), aspartate aminotransferase (AST), triglycerides (TGs), total cholesterol, high-density lipoprotein cholesterol (HDL-C), and low-density lipoprotein cholesterol (LDL-C) were measured by standard laboratory methods. As a control group, we included samples from 39 pediatric healthy children (mean age 10.7 years) who adhered to special programs of liver disease screening performed by our hospital each year. As an additional control group, 107 self-declared healthy Italians (Tuscans) individuals from the 1000 Genomes project for whom the genotypes of interest were available (*https://www.internationalgenome.org/data-portal/population/TSI*) were included. Samples and data were collected and used after the institutional review board approval ethics committee (protocol numbers 734_OPBG_

2014 and 1956_OPBG_2019). Written informed consent was obtained from the parents of each child enrolled in the study.

Additional Laboratory Tests on Non-Alcoholic Fatty Liver Disease Patients

Further laboratory tests were routinely performed in patients with NAFLD. Venous blood samples were collected in the morning after an overnight fast of at least 8 hours and immediately processed. Gamma-glutamyltransferase (GGT), glucose, insulin, bilirubin, uric acid, transferrin, and ferritin levels were measured using standard laboratory procedures. Homeostasis model assessment (HOMA-IR) score was used for estimating insulin resistance according to the equation: fasting insulin (μU/mL) × fasting glucose (mg/dL)/405.

Liver Histology

Liver histology was evaluated by an experienced histopathologist unaware of clinical and genetic data as already described (11). The main histological features, commonly described in NAFLD, including steatosis, lobular inflammation, hepatocyte ballooning, and fibrosis were scored according to the Scoring System for Non-Alcoholic Fatty Liver Disease developed by the NIH-sponsored NASH Clinical Research Network (CRN) (32). Disease activity was assessed according to the NAFLD Activity Score (NAS).

Genotyping

Analyses were limited to 10 SNPs found in nine genes. We selected our genes of interest firstly based on their association with longevity, preferably in two or more cohorts, and based on the association with liver/metabolic disorders. For our study, 7 of 26 genes were selected on the base of longevity-related genes reported on the online source SNPedia (*https://www.snpedia.com/index.php/Longevity*). Namely: IL6, ANRIL, ATXN2, FURIN, TP53, APOC3, and FOXO3. We also selected two SNPs on KLOTHO gene (KL), termed "KL-VS" variant, which were previously associated with longevity in multiple studies and in particular, in an Italian population. Moreover, we added another variant on KLB gene (β-Klotho), recently associated with liver damage in young NAFLD patients. Finally, we selected SIRT6 because of its association with longevity in two different cohorts, with glucose metabolism and because of its interaction with FOXO3. Genetic association studies of longevity were examined both manually, through researches on available bibliography, and by using specific browsers such as the already mentioned SNPedia, LongevityMap (*https://genomics.senescence.info/longevity/*), and GWAS catalog (*https://www.ebi.ac.uk/gwas/*). The gene variants of interest were selected from Ensembl Database with a frequency of >0.1 in the European population. The SNPs (Table 1, Supplemental Digital Content, *http://links.lww.com/MPG/C294*) were genotyped by TaqMan SNP genotyping assays (Thermo Fisher Scientific, USA). Briefly, genomic DNA was isolated from venous blood using a Blood & Tissue DNA Extraction Kit (Qiagen, Valencia, CA, USA). The absorbance ratio at 260/280 nm of all the samples ranged from 1.8 to 2 indicating that they were all free from contaminants. Real-time polymerase chain reaction (PCR) was

Drs Annalisa Crudele, Serena Dato, Oriana Lo Re, Andrea Maugeri, and Paola Sanna contributed equally to this study

performed using StepOnePlus Real-Time PCR System (Thermo Fisher Scientific, USA). Negative controls (NTC) were included on each reaction plate. Analyses were performed through Thermo Fisher Connect TM online tool for genotyping (n = 216).

Multifactor Dimensionality Reduction Analysis of Epistatic Interactions

The presence of epistatic interactions between pairs of SNPs in relation to NAFLD diagnosis was investigated by applying an multifactor dimensionality reduction (MDR) analysis (33,34). Through this approach, it is possible to estimate high-order interactions among variants with respect to a given phenotype and discover multilocus genotype combinations associated with high/low disease risk. In brief, MDR applies an entropy-based clustering algorithm, which starts with a contingency table for *k* SNPs and calculates case–control ratios for each of the possible multilocus genotypes. The MDR interaction model finally resulting reports in a network describing the percentage of entropy (information gain or IG) by each or two-way interaction. Graphical visualization is made through connections among the markers: values in the nodes indicate independent main effect, while values in the connectors the combined effect. The network helps to interpret additive and non-additive interactions effects on phenotype: positive values of entropy indicate synergistic or non-additive interactions, while negative entropy values indicate redundancy or lack of any synergistic interaction between the markers. Consistently, red and orange connections indicate epistatic interaction, in green and brown independence or additivity and blue indicates redundancy. For significance, permutation testing was applied, dividing the dataset into 10 portions, and using nine portions as a training data set, and the remaining as a testing data set. Missing genotypes were imputed directly with the MDR data tool software (version 0.4.3) from the existing data set. MDR analyses were implemented in the open-source MDR software package version 3.0.2 (available on *https://sourceforge.net/projects/mdr/*).

Statistical Analyses

All statistical analyses were conducted using GraphPad Prism (version 6.0, GraphPad Software, USA) or SPSS Statistics software (version 22.0, IBM Corporation, USA). The Kolmogorov-Smirnov test was first used to test the normality of continuous variables before further analyses. Results are expressed as median and minimum and maximum ranges or as the median and inter-quartile range (IQR) for skewed variables. Clinical parameters were compared across genotypes by Mann-Whitney *U* test for skewed variables. Association of the phenotypic trait (diagnosis of NAFLD vs healthy control) with genetic variants was first analyzed by Chi-squared test and eventually by fitting logistic regression models. The interaction between IL-6 rs1800795 and ANRIL rs1556516 variants on the association with clinical parameters was analyzed by generalized linear regression, adjusted for sex and age. All tests were two-sided and *P* values <0.05 were considered statistically significant, unless otherwise stated. The Bonferroni correction was applied for multiple comparison testing.

RESULTS

Association Between Pediatric Non-Alcoholic Fatty Liver Disease and Associated Hepato-Metabolic Damage with "Longevity" Variants

The study population was composed of 177 Italian children with biopsy-proven NAFLD and 39 healthy Italian children (CTL).

TABLE 1. Anthropometric and biochemical characteristics of healthy (CTL) and NAFLD children

	CTL (39)	NAFLD (177)	*P* value
Age (y)	12.1 (3.3)	13.7 (3.1)	–
Sex (female/male)	12/27	71/106	–
BMI (kg/m^2)	18 (12–30)	25.9 (14.3–66.1)	<0.001
Total cholesterol (mg/dL)	152 (58–244)	154 (51–226)	–
Triglycerides (mg/dL)	68 (29–195)	82 (20–322)	<0.01
HDL-C (mg/dL)	55 (28–72)	45 (21–105)	<0.01
LDL-C (mg/dL)	87 (46–129)	94 (14–166)	–
AST (UI/L)	24 (12–41)	31 (13–256)	<0.001
ALT (UI/L)	22 (13–35)	30 (6–216)	<0.01
GGT (UI/L)	–	16 (1–121)	–
Glucose (mg/dL)	–	95 (39–125)	–
Insulin (mg/dL)	–	50 (15.3–211)	–
HOMA-IR	–	3 (0.8–10)	–
Bilirubin (mg/dL)	–	0.63 (0.2–3.5)	–
Uric acid (mg/dL)	–	5.1 (2.4–10.4)	–
Transferrin (mg/dL)	–	296 (155–406)	–
Ferritin (ng/dL)	–	49 (10–140)	–

Data are reported as median (ranges). Unpaired Student *t*-test was used for comparisons between two groups in quantitative data. Statistically significant values are at $P < 0.05$ versus CTL.

ALT = alanine aminotransferase; AST = aspartate aminotransferase; BMI = body mass index; CTL = Italian children; GGT = gamma-glutamyltransferase; HDL-C = high-density lipoprotein cholesterol; HOMA-IR = homeostasis model assessment of insulin resistance; LDL-C = low-density lipoprotein cholesterol; NAFLD = non-alcoholic fatty liver disease.

Anthropometric and biochemical characteristics of the two groups (CTL vs NAFLD) are reported in Table 1. Median values of BMI, triglycerides, AST and ALT were significantly higher in NAFLD children than in controls; while HDL-C levels in NAFLD children were lower when compared to controls.

As previously indicated in methods, the sample size of controls was increased by considering 107 genotypes of healthy subjects (Caucasians, Tuscans) retrieved by 1000 Genomes. Since we found no difference in the allele frequency distribution between the young and adult control group, for genetic analyses the two subsamples were merged in a unique control group of 146 subjects. Minor allele frequency (MAF) analysis showed that FOXO3 rs2802292-G and SH2B3/ATXN2 rs7137828-T tended to be slightly over-represented in children with NAFLD when compared with controls, while IL6 rs1800795-C was under-represented. However, these differences were not statistically significant (Fig. 1).

Differences in clinical parameters based on the presence of each genetic variant in children with NAFLD are shown in Table 2, Supplemental Digital Content, *http://links.lww.com/MPG/C295*. The comparison between carriers of each genetic variant and homozygous carriers of the corresponding major allele revealed that some variants might be associated with metabolic derangement. Carrying the G allele of ANRIL rs1556516 tended to be associated with increased levels of glucose and insulin, and corresponding HOMA-IR. Carriers of the G allele of TP53 rs1042522 exhibited lower albumin levels and higher HOMA-IR value. Moreover, subjects with C allele for IL-6 rs1800795 exhibited decreased HOMA-IR, with allele A for FURIN rs17514846 showed decreased insulin levels, with allele A for APOC3 rs2542052 decreased total bilirubin levels, and with allele G for FOXO3 rs2802292 increased ferritin levels; however, none of these comparisons was statistically significant after the Bonferroni correction.
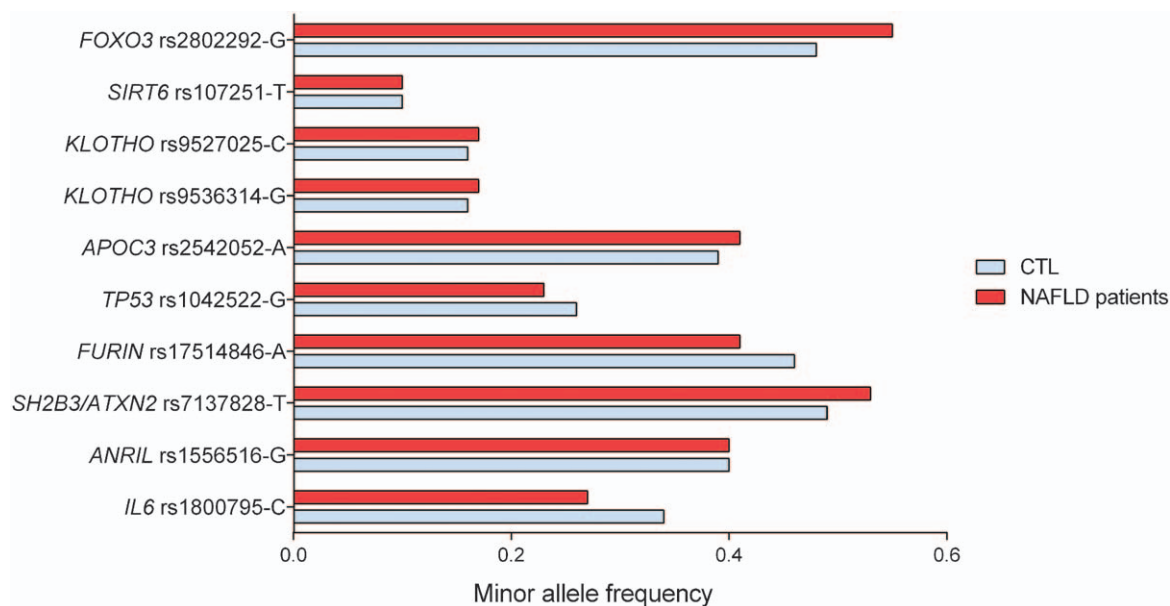
FIGURE 1. Allele frequencies distribution of healthspan variants. The histogram reports the frequency distribution of minor allele in children with NAFLD (n = 177) and healthy (n = 146: pediatric healthy controls [n = 39] + healthy Italian individuals from the 1000 genomes project [n = 107]). For each SNP, the minor allele was defined as the one with the lowest frequency in the control group. None of the differences observed was statistically significant. CTL = Italian children; NAFLD = non-alcoholic fatty liver disease; SNP = single nucleotide polymorphism.

The NAS in the whole population was 4.25 ± 1.3 on average. Among the 177 pediatric patients with biopsy-proven NAFLD, 141 children (79.7%) had NASH, 139 children (78.5%) had fibrosis. The 12.9% (n = 23) of patients had a severe degree of F3 fibrosis and as many as 34.5% (n = 61) had a degree F2 (see Table 3, Supplemental Digital Content, *http://links.lww.com/MPG/C296*). When carriers of each genetic variant were compared to homozygous carriers of the corresponding ancestral allele, emerged that NAFLD children with steatosis >1 were less frequent in those carrying the G allele for ANRIL rs1556516, while subjects with histologically definite NASH and with fibrosis >1 were less frequent in those with the T allele for SIRT6 (Table 2); however, these potential associations were not statistically significant after adjusting for multiple comparisons.

Analysis of Single Nucleotide Polymorphism–Single Nucleotide Polymorphism Interactions

Next, we analyzed whether gene–gene interactions among the 10 analyzed loci could be a predictor of pediatric NAFLD. For this purpose, we used the MDR method, which has been previously shown to be effective for detecting and characterizing gene to gene interactions in case-control studies on dyslipidemias, with relatively small samples (35,36), and successfully applied in the study of the genetic component of human longevity (37).

MDR analysis was performed by considering the NAFLD cases versus the whole control group. A significant epistatic interaction was found between the SNPs IL-6 rs1800795 and ANRIL rs1556516 (odds ratio: 2.34 [1.34, 3.93]; training X^2: 10.67 [$P = 0.0011$]). As shown by the red line in the interaction graph (A) and dendrogram (B) of Figure 2, the combination of these two SNPs gives a positive IG (2.01% explained entropy of the system), evidencing an increased contribution to the phenotype with respect to the single variants (0.97% and 0.07%, respectively). Moreover, Figure 2A and B shows a weaker but not significant interaction ($P > 0.05$ by permutation test) between the ANRIL rs1556516 and the ATXN2 rs7137828 variants (1.60% of entropy for the combination with respect to the single IG values, 0.07% and 0.61%, respectively).

Panel C of Figure 2 shows the risk/protective combinations of IL-6 rs1800795 and ANRIL rs1556516 genotypes with respect to the phenotype. A careful analysis of these distributions, which results globally significant ($P = 0.0002$) also after Bonferroni correction ($P < 0.005$), suggests that the genotype GG at IL-6 rs1800795 is an absolute risk genotype; and the combination of the two SNPs defines four risky combinations for NAFLD (IL-6-GG/ANRIL-CC; IL-6-GG/ANRIL-G carriers; IL-6-CG/ANRIL-CC; IL-6-CC/ANRIL-G carriers) and two protective ones (IL-6-CG/ANRIL-G carriers and IL-6-CC/ANRIL-CC), although less relevant because represented by a fewer number of samples.

L-6/ANRIL Single Nucleotide Polymorphism–Single Nucleotide Polymorphism Interaction is Associated with Blood Glucose in Children with Non-Alcoholic Fatty Liver Disease

We further evaluated whether the interaction between the SNPs IL-6 rs1800795 and ANRIL rs1556516 affected clinical parameters in patients with NAFLD. To this purpose, we applied generalized linear models to investigate both genotype-genotype interactions and risky/protective combinations for NAFLD after adjusting for age and sex. Except for mean blood glucose, no interaction was evident for the other clinical parameters. The genotype–genotype interaction on the mean glucose level, indeed, was statistically significant ($P = 0.042$), resulting in a different effect of the genotype GG at IL-6 rs1800795 across ANRIL rs1556516 genotypes (Fig. 3A). Specifically, the lowest glucose levels were found in subjects with the protective disease risk IL-6-CC/ANRIL-CC combination; however, among patients

TABLE 2. Comparison of histopathologic features by genetic variants in children with NAFLD

Genotypes	Steatosis > 1	NASH	Fibrosis > 1	Lobular inflammation > 1	Ballooning > 1	NAS > 4
IL-6 rs1800795						
CC + CG	75.6%	76.9%	46.2%	32.1%	28.2%	47.4%
GG	78.4%	81.4%	49.0%	38.1%	41.2%	49.5%
P-value	0.671	0.462	0.713	0.402	0.073	0.788
ANRIL rs1556516						
GG + GC	72.3%	82.1%	50.9%	40.2%	36.6%	47.3%
CC	85.7%	74.6%	41.9%	27.0%	33.3%	50.8%
P-value	0.043	0.236	0.257	0.080	0.664	0.659
SH2B3/ATX2 rs7137828						
TT + TC	74.0%	78.9%	43.9%	33.3%	31.7%	45.5%
CC	84.6%	80.8%	56.9%	40.4%	44.2%	55.8%
P-value	0.126	0.775	0.119	0.373	0.113	0.215
FURIN rs17514846						
AA + AC	74.3%	78.0%	44.4%	34.9%	34.9%	49.5%
CC	81.8%	81.8%	53.0%	36.4%	36.4%	47.0%
P-value	0.252	0.543	0.271	0.841	0.841	0.741
TP53 rs1042522						
GG + GC	80.3%	81.7%	50.0%	33.8%	31.0%	42.3%
CC	75.0%	77.9%	46.2%	36.5%	38.5%	52.9%
P-value	0.414	0.541	0.618	0.710	0.310	0.167
APOC3 rs2542052						
AA + AC	76.4%	79.2%	47.6%	36.8%	36.8%	48.6%
CC	80.6%	80.6%	48.4%	29.0%	29.0%	48.4%
P-value	0.609	0.853	0.933	0.412	0.412	0.982
KLOTHO rs9527025						
GG + GT	80.4%	74.5%	56.0%	35.3%	35.3%	49.0%
TT	75.8%	81.5%	44.4%	35.5%	35.5%	48.4%
P-value	0.512	0.302	0.164	0.981	0.981	0.939
KLOTHO rs9536314						
CC + CG	80.4%	74.5%	56.0%	35.3%	35.3%	49.0%
GG	75.8%	81.5%	44.4%	35.5%	35.5%	48.4%
P-value	0.512	0.302	0.164	0.981	0.981	0.939
SIRT6 rs107251						
TT + TC	78.1%	65.6%	31.3%	28.1%	28.1%	40.6%
CC	76.9%	82.5%	51.4%	37.1%	37.1%	50.3%
P-value	0.884	0.033	0.039	0.339	0.339	0.320
FOXO3 rs2802292						
GG + GT	83.3%	80.6%	48.6%	34.5%	35.3%	46.8%
TT	75.5%	75.0%	44.4%	38.9%	36.1%	55.6%
P-value	0.321	0.461	0.660	0.626	0.923	0.347

Results are reported as percentage and compared by the Chi-squared test. In this analysis, the Bonferroni-adjusted significance level is 0.0008.
NAS = non-alcoholic fatty liver disease activity score; NASH = non-alcoholic steatohepatitis.

with the genotype CC at ANRIL, mean glucose level increased with the presence of the allele G at IL-6. A different scenario was observed among children who were G carriers for ANRIL rs1556516, where a progressive decrease of mean glucose levels was observed. The analysis of the risky/protective combinations uncovered by the MDR method yielded similar findings ($P = 0.026$; Fig. 3B). Indeed, three risky combinations for NAFLD (IL-6-GG/ANRIL-CC; IL-6-CG/ANRIL-CC; IL-6-CC/ANRIL-CG+GG) were also associated with higher mean glucose levels, while the protective IL-6-CC/ANRIL-CC combination led to the lowest value. However, the IL-6-GG/ANRIL-GG+CG combination, which constituted a risk for NAFLD, seemed not to be associated with high glucose levels, as well as the IL-6-CG/ANRIL-GG+CG combination, which appears to confer a protective effect for NAFLD, seemed not to be associated with low glucose levels.

DISCUSSION

The present study has identified genetic variants affecting lifespan/healthspan associated with metabolic and liver features in pediatric NAFLD.

NAFLD reflects increased rates of obesity and its metabolic consequences; however genetic factors clearly determine how individuals respond to environmental factors (ie, diet). Among the discovered genetic variants, the rs738409 in patatin-like phospholipase domain-containing protein 3 (PNPLA-3) gene has emerged as a major common determinant of NAFLD, but also others relevant mutations have been discovered to exert a small effect on the risk of the disease (38).

Moreover, NAFLD is considered a nexus of metabolic and liver disease both in adults and children (5). Insulin resistance is crucial for pediatric NAFLD progression and common
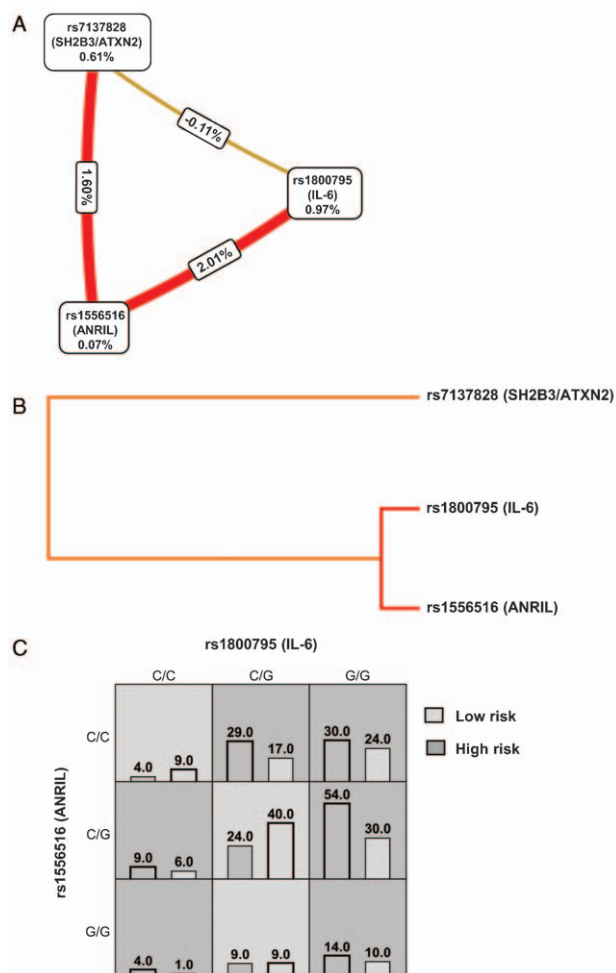
FIGURE 2. Interaction in NAFLD data set resulting from MDR analysis. (A) Interaction graph for NAFLD data set. For each SNP is reported in percent the value of information gain (IG) and numbers in the connections indicate the entropy-based IG for the SNP pairs. (B) Interaction dendrogram for NAFLD data set. (A and B) Red bar and orange bar indicate the high-level synergies on the phenotype, while the brown indicates a medium-level interaction between the markers. (C) Histograms report the distributions of cases (left bars) and controls (right bars) for two-locus genotype combinations of IL-6 and ANRIL SNPs, globally significant (*P* = 0.0002) and holding Bonferroni correction (*P* < 0.005). Dark-shaded cells are considered ''high risk,'' whereas light-shaded cells are ''low risk'' combination for NAFLD. White cells indicate the absence of subjects with a given genotype combination in the dataset. IG = information gain; IL-6 = interleukin-6; MDR = multidimensional reduction; NAFLD = non-alcoholic fatty liver disease; SNP = single nucleotide polymorphism.

polymorphisms of genes involved in the insulin-signaling pathway influence the susceptibility to T2D. Since 2007, genome-wide association studies (GWAS) have identified approximately 88 loci associated with the risk of developing T2D, with most of them primarily associated with insulin sensitivity, such as glucokinase regulator (GCKR) gene locus (39).

A genetic variation in GCKR (rs1260326), encoding for the P446L protein variant affects the ability of the protein to negatively regulate the glucokinase in response to fructose-6-phosphate, thus determining constitutive activation of hepatic glucose uptake and

contributing to NAFLD development (40). The combined effects of PNPLA3 rs738409 and GCKR rs1260326 variants may explain up to one-third of variability in liver fat content in obese children (41).

However, several other rare genetic variants may presumably explain the risk of disease in subjects that exhibit specific phenotypes in only one ethnic or few populations. Currently, it is unknown if rare gene variants predisposing to longer lifespan are associated with protection from NAFLD and its metabolic disturbances. Hence, in this study, we carefully selected from the literature ten SNPs for their established involvement in both lifespan/healthspan and liver/metabolic diseases in humans (19–28). None of these SNPs was significantly associated with the risk of NAFLD, even if ANRIL rs1556516-G variant was less frequent in NAFLD children with steatosis, and SIRT6 rs107251-T variant was less frequent in subjects with NASH and fibrosis. These data represent the first evidence that couples these gene variants with pediatric NAFLD.

We found that ANRIL rs1556516 (CC) and TP53 rs1042522 (CC) tended to be associated with reduced, while IL-6 rs1800795 (GG) associated with increased, insulin resistance assessed by HOMA-IR. ANRIL rs1556516 (CC) was also related to lower levels of glucose and insulin.

Despite the lack of significance after correction of multiple testing, these results are however in agreement with the already existing evidence of a link between IL-6 and ANRIL genes, and insulin resistance. Indeed, IL-6 is one of several pro-inflammatory cytokines associated with insulin resistance and T2D. Plasma levels of IL-6 are two- to three-fold higher in patients with obesity and T2D than in lean control subjects. This elevation is highly related to increased blood glucose, decreased glucose tolerance, and decreased insulin sensitivity (42). It has been reported that IL-6 can inhibit insulin receptor (IR) signal transduction and insulin action in both primary mouse hepatocytes and the human HepG2 hepatocarcinoma cell line (43). Furthermore, individuals with G/G genotype of the IL-6 gene have a polymorphism that leads to increased IL-6, increased circulating insulin, and higher blood glucose than those individuals who have the C/C genotype (43). Whereas, ANRIL has been confirmed to be associated with the onset of diabetic retinopathy, a severe diabetic complication. Overexpression of ANRIL may be a result of the activation of the renin–angiotensin system (RAS) and nuclear factor kappa B (NF-κB) pathway (44,45). Anu Alice Thomas et al revealed that ANRIL expression can be upregulated in vascular endothelial cells through PRC2, miR200b, and p300 and that ANRIL can mediate the proliferation of retinal vascular endothelial cells and participate in the pathogenesis of diabetic retinopathy (46).

MDR, which we used in our study to identify the genotype–genotype interactions and risky combinations for NAFLD, is a well-known multivariate technique, often used in genetics for studying population substructure. If individuals come from populations with different allele frequencies, then an MDR of the genetic marker data typically separates the individuals of the different populations. Besides its use in studies focusing on hyperlipidemia, MDR allowed the identification of SAMM50 gene variants in the susceptibility to NAFLD (35,36).

In the present study, MDR revealed, for the first time, epistatic interaction between IL6 and ANRIL highlighting risk and protective genotype combinations for the disease. The evidence that ANRIL, which is a long non-coding RNA, acts as an upstream regulator of IL-6 expression, particularly in stress conditions (45), provides a biologically plausible mechanism for the reported genetic interactions.

Interestingly, our analysis showed a protective combination for NAFLD associated with lower glucose levels, as well as risk combinations associated with higher glucose levels, suggesting that
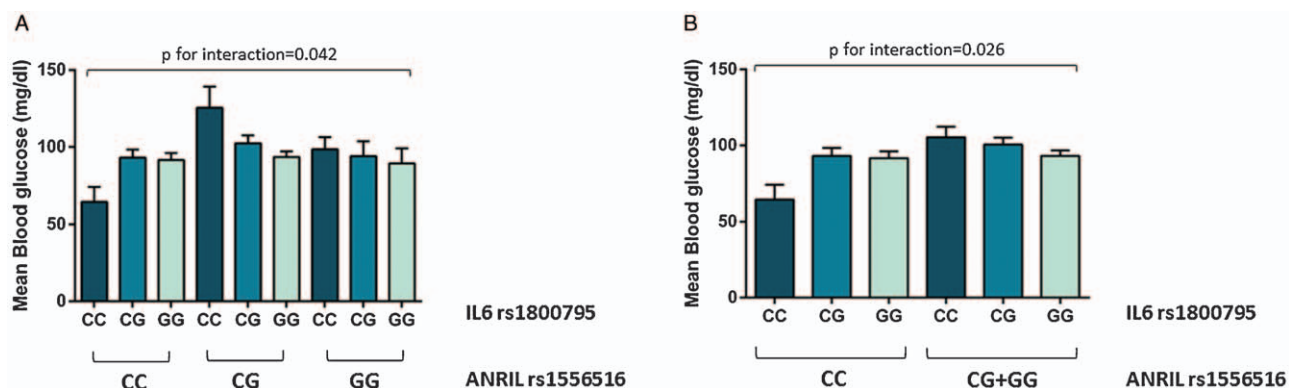
FIGURE 3. Interaction of mean blood glucose levels in children with NAFLD stratified for IL-6 rs1800795 and ANRIL rs1556516 genotypes. (A) The bar graph shows mean blood glucose levels for each IL-6 rs1800795-ANRIL rs1556516 genotype-genotype interaction. (B) The bar graph shows mean blood glucose level for risky/protective combinations uncovered by the MDR method. *P*-values are based on test for interaction using generalized linear models adjusting for age and sex. IL-6 = interleukin-6; MDR = multidimensional reduction; NAFLD = non-alcoholic fatty liver disease.

genetic interactions between specific combinations of *IL-6* and *ANRIL* alleles may influence the risk of disease by influencing hepatocellular glucose turnover.

During NAFLD pathogenesis, insulin resistance results in increased concentrations of circulating free fatty acids, which inhibits glucose uptake by muscle cells and increases glucose production by the liver. Impaired glucose metabolism leads to impaired glucose tolerance that is an intermediate stage in the natural history of T2D and predicts the risk of CVD (47,48). Thus, the association of healthspan/lifespan genetics with the hyperglycemia that we found here could be also partially explain the predisposition to T2D of children with NAFLD. Glucose metabolism and its association with NAFLD were, in fact, investigated in several cohorts of obese children of different ethnic origins (49–51); however, a recent study reported that children with NAFLD exhibited a higher prevalence of abnormal glucose tolerance than subjects without the disease, and that impaired glucose metabolism also increased predisposition to NASH in this pediatric setting (52). Thus, in this vicious circle glucose metabolism may impact on NAFLD and vice versa, and combined interaction between IL-6 rs1800795, ANRIL rs1556516 and hyperglycemia could increase the risk of children with NAFLD in developing T2D strongly affecting their healthspan and lifespan.

It should however also be noted that our data indicate that the relationship among IL-6/ANRIL genotypic combinations, glucose levels, and disease risk is not linear, reflecting the complex genetic architecture of NAFLD, which is affected by genetic variant combinations potentially modified by non-genetic factors. We could then speculate that SNPs located in *IL-6* and *ANRIL* synergistically cooperate, ultimately affecting NAFLD risk in a manner that is dependent on allele dosage for each SNP and by the pleiotropic effects of these genes on different clinicopathological aspects of the disease.

In conclusion, in the present study, we have demonstrated a synergic interaction between IL-6 rs1800795 and ANRIL rs1556516 in NAFLD diagnosis, and NAFLD-associated hyperglycemia in a pediatric setting. Quality longitudinal data of the natural history of pediatric NAFLD is limited, and rigorous efforts for structured follow-up are a priority to better develop the understanding of life-long outcomes of children with NAFLD. Larger studies are required to confirm our findings and to elucidate the mechanisms through which the genetic interaction between IL-6 and ANRIL influences healthspan in subjects with NAFLD.

REFERENCES

1. Eslam M, Sanyal AJ, George J, et al. MAFLD: a consensus-driven proposed nomenclature for metabolic associated fatty liver disease. *Gastroenterology* 2020;158:1999.e1–2014.e1.
2. Younossi ZM. Non-alcoholic fatty liver disease—a global public health perspective. *J Hepatol* 2019;70:531–44.
3. Yeh MM, Brunt EM. Pathological features of fatty liver disease. *Gastroenterology* 2014;147:754–64.
4. Goh GB, McCullough AJ. Natural history of nonalcoholic fatty liver disease. *Dig Dis Sci* 2016;61:1226–33.
5. Nobili V, Alisi A, Newton KP, et al. Comparison of the phenotype and approach to pediatric vs adult patients with nonalcoholic fatty liver disease. *Gastroenterology* 2016;150:1798–810.
6. Romeo S, Kozlitina J, Xing C, et al. Genetic variation in PNPLA3 confers susceptibility to nonalcoholic fatty liver disease. *Nat Genet* 2008;40:1461–5.
7. 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.
8. Valenti L, Alisi A, Galmozzi E, et al. I148M patatin-like phospholipase domain-containing 3 gene variant and severity of pediatric nonalcoholic fatty liver disease. *Hepatology* 2010;52:1274–80.
9. Dongiovanni P, Petta S, Maglio C, et al. Transmembrane 6 superfamily member 2 gene variant disentangles nonalcoholic steatohepatitis from cardiovascular disease. *Hepatology* 2015;61:506–14.
10. Nobili V, Svegliati-Baroni G, Alisi A, et al. A 360-degree overview of paediatric NAFLD: recent insights. *J Hepatol* 2013;58:1218–29.
11. Dongiovanni P, Crudele A, Panera N, et al. beta-Klotho gene variation is associated with liver damage in children with NAFLD. *J Hepatol* 2020;72:411–9.
12. Dongiovanni P, Valenti L. Genetics of nonalcoholic fatty liver disease. *Metabolism* 2016;65:1026–37.
13. Bomba L, Walter K, Soranzo N. The impact of rare and low-frequency genetic variants in common disease. *Genome Biol* 2017;18:77. doi: 10.1186/s13059-017-1212-4.
14. Melzer D, Pilling LC, Ferrucci L. The genetics of human ageing. *Nat Rev Genet* 2020;21:88–101.
15. Franceschi C, Garagnani P, Olivieri F, et al. The contextualized genetics of human longevity: JACC focus seminar. *J Am Coll Cardiol* 2020;75:968–79.
16. Longo VD, Antebi A, Bartke A, et al. Interventions to slow aging in humans: are we ready? *Aging Cell* 2015;14:497–510.

17. Ruby JG, Wright KM, Rand KA, et al. Estimates of the heritability of human longevity are substantially inflated due to assortative mating. *Genetics* 2018;210:1109–24.

18. Longo VD. Programmed longevity, youthspan, and juventology. *Aging Cell* 2019;18:e12843. doi: 10.1111/acel.12843.

19. Huang Y, Xiang B, Liu Y, et al. LncRNA CDKN2B-AS1 promotes tumor growth and metastasis of human hepatocellular carcinoma by targeting let-7c-5p/NAP1L1 axis. *Cancer Lett* 2018;437:56–66.

20. Ueyama C, Horibe H, Yamase Y, et al. Association of FURIN and ZPR1 polymorphisms with metabolic syndrome. *Biomed Rep* 2015;3:641–7.

21. Di Pietro F, Dato S, Carpi FM, et al. TP53*P72 allele influences negatively female life expectancy in a population of central Italy: cross-sectional study and genetic-demographic approach analysis. *J Gerontol A Biol Sci Med Sci* 2013;68:539–45.

22. Gross S, Immel UD, Klintschar M, et al. Germline genetics of the p53 pathway affect longevity in a gender specific manner. *Curr Aging Sci* 2014;7:91–100.

23. Rebbani K, Marchio A, Ezzikouri S, et al. TP53 R72P polymorphism modulates DNA methylation in hepatocellular carcinoma. *Mol Cancer* 2015;14:74. doi: 10.1186/s12943-015-0340-2.

24. Atzmon G, Rincon M, Schechter CB, et al. Lipoprotein genotype and conserved pathway for exceptional longevity in humans. *PLoS Biol* 2006;4:e113. doi: 10.1371/journal.pbio.0040113.

25. Xie B, Zhou J, Yuan L, et al. Epigenetic silencing of Klotho expression correlates with poor prognosis of human hepatocellular carcinoma. *Hum Pathol* 2013;44:795–801.

26. Poh W, Wong W, Ong H, et al. Klotho-beta overexpression as a novel target for suppressing proliferation and fibroblast growth factor receptor-4 signaling in hepatocellular carcinoma. *Mol Cancer* 2012;11:14. doi: 10.1186/1476-4598-11-14.

27. TenNapel MJ, Lynch CF, Burns TL, et al. SIRT6 minor allele genotype is associated with >5-year decrease in lifespan in an aged cohort. *PLoS One* 2014;9:e115616.

28. Banasik K, Ribel-Madsen R, Gjesing AP, et al. The FOXO3A rs2802292 G-allele associates with improved peripheral and hepatic insulin sensitivity and increased skeletal muscle-FOXO3A mRNA expression in twins. *J Clin Endocrinol Metab* 2011;96:E119–24.

29. Hurme M, Lehtimaki T, Jylha M, et al. Interleukin-6-174G/C polymorphism and longevity: a follow-up study. *Mech Ageing Dev* 2005;126:417–8.

30. Wang X, Yan Z, Ye Q. Interleukin-6 gene polymorphisms and susceptibility to liver diseases: a meta-analysis. *Medicine (Baltimore)* 2019;98:e18408.

31. Kong Y, Hsieh CH, Alonso LC. ANRIL: a lncRNA at the CDKN2A/B locus with roles in cancer and metabolic disease. *Front Endocrinol (Lausanne)* 2018;9:405.

32. Kleiner DE, Brunt EM, Van Natta M, et al. Design and validation of a histological scoring system for nonalcoholic fatty liver disease. *Hepatology* 2005;41:1313–21.

33. Ritchie MD, Hahn LW, Roodi N, et al. Multifactor-dimensionality reduction reveals high-order interactions among estrogen-metabolism genes in sporadic breast cancer. *Am J Hum Genet* 2001;69:138–47.

34. Moore JH. Computational analysis of gene-gene interactions using multifactor dimensionality reduction. *Expert Rev Mol Diagn* 2004;4:795–803.

35. Miao L, Yin RX, Pan SL, et al. BCL3-PVRL2-TOMM40 SNPs, gene–gene and gene–environment interactions on dyslipidemia. *Sci Rep* 2018;8:6189. doi: 10.1038/s41598-018-24432-w.

36. Chen L, Lin Z, Jiang M, et al. Genetic variants in the SAMM50 gene create susceptibility to nonalcoholic fatty liver disease in a Chinese Han population. *Hepat Mon* 2015;15:e31076. doi: 10.5812/hepatmon.31076.

37. Dato S, Soerensen M, De Rango F, et al. The genetic component of human longevity: new insights from the analysis of pathway-based SNP–SNP interactions. *Aging Cell* 2018;17:e12755. doi: 10.1111/acel.12755.

38. Eslam M, Valenti L, Romeo S. Genetics and epigenetics of NAFLD and NASH: clinical impact. *J Hepatol* 2018;68:268–79.

39. Mohlke KL, Boehnke M. Recent advances in understanding the genetic architecture of type 2 diabetes. *Hum Mol Genet* 2015;24 (R1):R85–92.

40. Beer NL, Tribble ND, McCulloch LJ, et al. The P446L variant in GCKR associated with fasting plasma glucose and triglyceride levels exerts its effect through increased glucokinase activity in liver. *Hum Mol Genet* 2009;18:4081–8.

41. Valenti L, Alisi A, Nobili V. Unraveling the genetics of fatty liver in obese children: additive effect of P446L GCKR and I148M PNPLA3 polymorphisms. *Hepatology* 2012;55:661–3.

42. Kern PA, Ranganathan S, Li C, et al. Adipose tissue tumor necrosis factor and interleukin-6 expression in human obesity and insulin resistance. *Am J Physiol Endocrinol Metab* 2001;280:E745–51.

43. Fernandez-Real JM, Broch M, Vendrell J, et al. Interleukin-6 gene polymorphism and insulin sensitivity. *Diabetes* 2000;49:517–20.

44. Zhang B, Wang D, Ji TF, et al. Overexpression of lncRNA ANRIL up-regulates VEGF expression and promotes angiogenesis of diabetes mellitus combined with cerebral infarction by activating NF-kappaB signaling pathway in a rat model. *Oncotarget* 2017;8:17347–59.

45. Zhou X, Han X, Wittfeldt A, et al. Long non-coding RNA ANRIL regulates inflammatory responses as a novel component of NF-kappaB pathway. *RNA Biol* 2016;13:98–108.

46. Thomas AA, Feng B, Chakrabarti S. ANRIL: a regulator of VEGF in diabetic retinopathy. *Invest Ophthalmol Vis Sci* 2017;58:470–80.

47. Polonsky KS, Sturis J, Bell GI seminars in medicine of the Beth Israel Hospital, Boston. Non-insulin-dependent diabetes mellitus—a genetically programmed failure of the beta cell to compensate for insulin resistance. *N Engl J Med* 1996;334:777–83.

48. Haffner SM, Stern MP, Hazuda HP, et al. Cardiovascular risk factors in confirmed prediabetic individuals. Does the clock for coronary heart disease start ticking before the onset of clinical diabetes? *JAMA* 1990;263:2893–8.

49. Newton KP, Hou J, Crimmins NA, et al. Prevalence of prediabetes and type 2 diabetes in children with nonalcoholic fatty liver disease. *JAMA Pediatr* 2016;170:e161971.

50. Bedogni G, Gastaldelli A, Manco M, et al. Relationship between fatty liver and glucose metabolism: a cross-sectional study in 571 obese children. *Nutr Metab Cardiovasc Dis* 2012;22:120–6.

51. Cali AM, De Oliveira AM, Kim H, et al. Glucose dysregulation and hepatic steatosis in obese adolescents: is there a link? *Hepatology* 2009;49:1896–903.

52. Nobili V, Mantovani A, Cianfarani S, et al. Prevalence of prediabetes and diabetes in children and adolescents with biopsy-proven non-alcoholic fatty liver disease. *J Hepatol* 2019;71:802–10.