Hormones (2019) 18:463–475
https://doi.org/10.1007/s42000-019-00137-1

ORIGINAL ARTICLE

Potential clinical variants detected in mitochondrial DNA D-loop hypervariable region I of patients with non-alcoholic steatohepatitis

Burcu Hasturk[1] · Yusuf Yilmaz[2,3] · Fatih Eren[3,4]

Received: 5 June 2019 / Accepted: 22 September 2019 / Published online: 26 October 2019
© Hellenic Endocrine Society 2019

Abstract

Purpose Non-alcoholic steatohepatitis (NASH) is a mitochondrial disease. However, the underlying role of mitochondrial genetics has not yet been completely elucidated. Evaluation of D-loop nucleotide variations with respect to statistical significance and clinical data distribution.

Methods Genomic DNAs were extracted from the peripheral blood samples of patients with biopsy-proven 150 NASH as well as from 150 healthy individuals to explore the functional D-loop region responsible for the replication and transcription of the mitochondrial genome. DNA sequencing by capillary electrophoresis analysis was performed for the D-loop region of mitochondrial DNA containing the hypervariable region I, and restriction fragment length polymorphism with *Mnl*I analysis was performed for the m.16189 T/C D-loop variant.

Results The m.A16318C variant was detected only in patients with NASH and approached significance level. Based on clinical data, six variants associated with histological subgroups of NASH and NASH-complicated diseases were identified. In patients with NASH, the m.16129 AA genotype was associated with advanced-stage fibrosis; the m.16249 CC genotype was associated with advanced lobular inflammation and advanced-stage histological steatosis; the m.16296 TT genotype was associated with hypothyroidism; the m.16163 GG and m.16294 TT genotypes were associated with metabolic syndrome; and the m.16256 TT+CT genotypes were associated with type II diabetes. In patients with NASH, microRNAs were estimated by targeting the significant variants identified in this study.

Conclusion These findings suggest that NASH may be associated with D-loop nucleotide variations and that microRNA-based in vitro and/or in vivo studies may be developed by targeting the D-loop variants.

Keywords D-loop · MicroRNA target prediction · Mitochondrial DNA · Non-alcoholic steatohepatitis · Non-alcoholic fatty liver disease · Polymorphism

Introduction

Non-alcoholic steatohepatitis (NASH) is defined as conditions such as liver inflammation secondary to simple fatty liver and liver cell damage accompanying fibrosis in non-alcoholic fatty liver disease (NAFLD). Several diseases such as insulin resistance (IR), type 2 diabetes mellitus (T2DM), obesity, and metabolic syndrome (MetS) are considered to be risk factors for NASH [1]; however, NASH can develop in individuals who do not present with these conditions.

NASH represents a critical point of the disease, after which it may progress to cause severe complications, such as cirrhosis and hepatocellular carcinoma [2]. It is, furthermore, a severe disease requiring treatment even in the absence of complications. Although developmental mechanisms underlying NASH and other NASH-complicated conditions in patients

Electronic supplementary material The online version of this article (https://doi.org/10.1007/s42000-019-00137-1) contains supplementary material, which is available to authorized users.

✉ Fatih Eren
fatiheren@marmara.edu.tr

[1] Department of Medical Biology and Genetics, Institute of Health Sciences, Marmara University, Istanbul, Turkey

[2] Department of Gastroenterology, School of Medicine, Marmara University, Istanbul, Turkey

[3] Institute of Gastroenterology, Marmara University, Istanbul, Turkey

[4] Department of Medical Biology, School of Medicine, Marmara University, Maltepe Basibuyuk Yolu Road No: 9/2, 34854 Istanbul, Turkey

with NAFLD have not to date been completely elucidated, it is thought to develop as a result of a combination of behavioral risk factors, such as lifestyle and diet, as well as epigenetic mechanisms and genetic predisposition [3]. In patients who develop IR due to sedentary lifestyles and intake of fat- and carbohydrate-rich diets, increased glucose and free fatty acids are transported to hepatocytes, which are insulin-responsive cells.

Mitochondria maintain cellular homeostasis by metabolizing glucose and lipids through oxidative phosphorylation and beta-oxidation, respectively [4]. However, reactive oxygen species generated as by-products lead to mutations in mitochondrial DNA (mtDNA), causing damage. Mitochondrial dysfunction triggers inflammation and fibrosis in the liver [5]. The non-coding D-loop control region of mtDNA harbors mutational domains [6]. Because the mtDNA D-loop region contains the replication origins of heavy chains and promoters for the transcription of light and heavy chains [7], nucleotide variations occurring in this region have frequently been investigated to determine their association with various diseases. In this study, we examined the association between NASH and nucleotide variations in the D-loop region containing the hypervariable region I (HVI).

Materials and methods

Subjects

This is a case-control study involving 300 subjects (150 consecutive patients with NASH and 150 age- and sex-matched healthy controls) aged 18–65 years, who were followed up in the Department of Gastroenterology at Marmara University School of Medicine and Clinic of the Institute of Gastroenterology, Istanbul, Turkey. The above patients were selected as healthy controls due to normal transaminase levels, absence of hepatic steatosis in liver ultrasonography, and absence of known chronical disease according to each individual's self-declaration. The exclusion criteria of patients are described in detail in our previous studies [8, 9]. Demographic characteristics of the study group are summarized in Table 1.

Liver histology

Liver biopsy conditions were as follows: (i) evidence of hepatic steatosis on ultrasonographic examination, (ii) abnormal liver enzymes or organomegaly, and (iii) exclusion of secondary etiologies of fat accumulation in the liver (e.g., significant alcohol consumption (> 30 g/day of alcohol for men and 20 g/day of alcohol for women) and previous history of steatogenic drugs use). Liver biopsy specimens having at least 11 portal tracts were accepted as adequate material and evaluated.

Table 1 Demographic and biochemical characteristics of patients with NASH according to clinical findings

Variables	Quantities
Age (years)*	45.08 ± 9.41 (n = 150)
Gender (female/male)	75/75 (n = 150)
BMI (kg/m^2)*	31.7 ± 5.69 (n = 150)
Type 2 diabetes mellitus (yes/no)	44/103 (n = 147)
Metabolic syndrome (yes/no)	94/56 (n = 150)
Hypothyroidism (yes/no)	23/127 (n = 150)
Hypertension (yes/no)	44/104 (n = 148)
Hyperlipidemia (yes/no)	79/67 (n = 146)
Gallstone disease (yes/no)	20/127 (n = 147)
HOMA-IR**	5.02 (5.02–32.95) (n = 137)
Systolic blood pressure (mmHg)*	129.22 ± 18.08 (n = 148)
Diastolic blood pressure (mmHg)*	82.72 ± 11.14 (n = 148)
AST (U/L)*	51.53 ± 37.1 (n = 150)
ALT (U/L)*	79.05 ± 44.92 (n = 150)
Total cholesterol (mg/dL)*	211.92 ± 43.94 (n = 149)
HDL cholesterol (mg/dL)*	44.62 ± 10.32 (n = 149)
LDL cholesterol (mg/dL)*	134.33 ± 40.13 (n = 147)
Triglycerides (mg/dL)*	192.44 ± 108.52 (n = 149)
CRP (mg/dL)*	2.39 ± 3.14 (n = 141)
Liver histology	Quantities
Histologic steatosis**	2 (1–3)
Lobular inflammation**	2 (1–2)
Ballooning**	2 (1–2)
Fibrosis**	1 (0–2)
NAS score**	5 (4–6)

*Mean and standard deviation values were calculated using descriptive statistics and compared using Student's t test

**Median values were expressed in interquartile range and compared with Mann-Whitney U test. *BMI*, body mass index; *NASH*, non-alcoholic steatohepatitis; *HOMA-IR*, homeostatic model of assessment-insulin resistance; *AST*, aspartate aminotransferase; *ALT*, alanine aminotransferase; *HDL*, high-density lipoprotein; *LDL*, low-density lipoprotein; *CRP*, C-reactive protein

The scoring and severity of all patients with NASH were classified according to the NASH-Clinical Research Network (CRN) criteria. The liver biopsy specimens were evaluated and scored according to the NAFLD activity (NAS) scoring system [10] and according to the steatosis, activity, and fibrosis (SAF) histological algorithm by an expert liver pathologist. SAF algorithm consists of three following parameters: grade of steatosis (0–3), activity (0–4) and fibrosis (0–4). The activity grade showed the summation of the scores for ballooning (0–2) and lobular inflammation (0–2). When ballooning and lobular inflammation was at least 1 with presence of hepatic steatosis, the specimen was considered as NASH and included in our study [11]. All liver biopsies were performed by a single

physician who is an expert in invasive radiology. The specimens were also evaluated by a single pathologist.

PCR amplification

Genomic DNA was extracted from peripheral blood samples of our 300 subjects according to the PureLink Genomic DNA Mini kit protocol (Thermo Fisher Scientific, USA). A 630-bp length target region (np 15911-16540) was amplified according to the i-StarTaq DNA Polymerase kit protocol (Intron Biotechnology, Korea). For polymerase chain reaction (PCR), 5 μL of 10× buffer (Mg^{++} free), 3 μL of 25 mM $MgCl_2$, 1 μL of 10 mM dNTP mix, 1 μL each of forward (5′-ACCAGTCTTGTAAACCGGAG-3′) and reverse (5′-GTGGGCTATTTAGGCTTTAT-3′) reference primers [12], 0.3 μL of DNA polymerase, and 5 μL of DNA (~ 150 ng) were added; the final volume was raised to 50 μL with DNase- and RNase-free H_2O. In a thermal cycler (Biometra TProfessional, Germany), 30 cycles of 45 s at 95 °C, 1 min at 55 °C, and 45 sec at 72 °C were performed. After amplification, PCR success rate and semi-quantitative analysis of amplicons were performed using 1% agarose gel electrophoresis (Vilber lourmat, France). PCR products were used for restriction fragment length polymorphism (RFLP) and DNA sequencing analyses.

m.16189 T/C genotyping

PCR-RFLP analysis for the m.16189 T/C genotype was performed in accordance with the Anza 48 *MnlI* (5 U/μL) kit protocol (Thermo Fisher Scientific, USA). Fragments with lengths of 286, 118, 84, 39 (2), and 28 bp in the presence of the T allele, and fragments with lengths of 321, 118, 84, 39 (2), and 28 bp in the presence of the C allele were excised from the CCTC(7/6)^ recognition site of PCR products in *MnlI* enzyme activation. Restriction fragment lengths of RFLP products (10 μL) were visualized using agarose gel electrophoresis (1% nusieve, 2% agarose). All PCR-RFLP results were confirmed by sequence analysis.

MtDNA D-loop genotyping

Sequence analysis of the target region (np 15911-16540; ~ 630 bp) in the D-loop control area containing the HVI region (np 16024-16383) was performed using Applied Biosystems 3500 Genetic Analyzer (Thermo Fisher Scientific, USA). PCR products were purified using the ExoSAP-IT PCR Clean-Up (Affymetrix, USA) reagent and purified PCR products were sequenced using the BigDye Terminator v3.1 Cycle Sequencing kit (Thermo Fisher Scientific, USA). Sequenced PCR products were purified using EDTA/NaOAc/Ethanol precipitation. Hi-Di formamide (Thermo Fisher Scientific, USA) was added, followed by elution, denaturation, and renaturation steps. Then, the samples were loaded into the Genetic Analyzer and mtDNA was sequenced. Forward and reverse primers were used for each sample to ensure bidirectional control of sequencing.

Haplogroup analysis

Raw data obtained from sequence analysis were processed using Chromas free version 2.6.5 software (Technelysium Pty Ltd, Australia, https://technelysium.com.au/wp/chromas/) and compared with a revised Cambridge reference sequence (rCRS, NC_012920) from the NCBI GenBank database using Nucleotide Basic Local Alignment Search Tool, (BLAST, https://blast.ncbi.nlm.nih.gov). MITOMAP allele search (https://www.mitomap.org/foswiki/bin/view////Main/SearchAllele, date of access, May 2017) was used to detect point mutations. In addition, haplogroups were identified using the mtDNA profiles of samples using HaploGroup 2.0 software (automatic mtDNA haplogroup classification using PhyloTree 17, http://haplogrep.uibk.ac.at, date of access, May 2017).

Prediction of microRNAs in target sequences

In this study, the prominent D-loop variants (m.A16318C; m.G16129A; m.A16163G; m.T16249C; m.C16256T; m.C16294T; and m.C16296T) were targeted, and for two alleles of each variant, the sequence was shortened, starting from a sequence with a length of approximately 100 nucleotides (50 bases before and 50 bases after the nucleotide position) and analyzed using the miRBase database (http://www.mirbase.org/search.shtml; date of access, October 2018). According to the default settings of the search sequences tool, human mature microRNAs (miRNAs) and stem-loop sequences were screened using the *E* value cutoff of 10. Thus, the extracted miRNAs were compared for two bases of the variant that showed significant frequency difference. miRNAs showing nucleotide base variation were recorded. Moreover, the literature was reviewed to determine the potential association of mature and hairpin miRNAs with NASH.

Statistical analysis

Descriptive statistics were used to describe study sample characteristics. Normality was assessed with the Shapiro-Wilk test and Kolmogorov-Smirnov test. The normally distributed quantitative variables were presented as mean ± standard deviation and compared with Student's *t* test, whereas the non-normally distributed quantitative variables were expressed as median (interquartile range) and compared with the Mann-Whitney *U* test. Qualitative variables were presented as absolute and relative (%) frequencies.

The statistical analysis was performed using GraphPad InStat software (a free statistical calculator) and IBM SPSS Statistics V25 statistical programs. The Hardy-Weinberg equilibrium (HWE) is not evaluated in allele distribution, since the mitochondrial genome is a homoplasmic structure. The chi-squared test was performed for categorical variables by using the GraphPad InStat software. Fischer's exact test (two-sided) was used to detect the association of genotypes with study groups. Odds ratios (ORs) with 95% confidence intervals (CIs) were calculated to assess the risk conferred by the genotypes. A two-sided P value < 0.05 was considered statistically significant.

Accession numbers

MtDNA D-loop sequences (300 samples) of data from this study have been reported to the European Molecular Biology Laboratory (EMBL) Nucleotide Sequence Database (EMBL via Webin, http://www.ebi.ac.uk/embl/Submission/webin.html) under accession numbers PRJEB25168.

Results

Clinical findings

Mean age (standard deviation) of the patients with NASH was 45.08 (9.41) years and that of the healthy individuals was 44.76 (9.14) years. Clinical findings of patients included in the study are presented in Table 1.

m.16189 T/C, CC genotype is not associated with NASH

The m.16189 T/C, CC variant was analyzed in patients with NASH and healthy individuals using PCR-RFLP with *Mnl*I analysis (Fig. 1) and confirmed by sequence analysis. The T allele was the major allele, while the C allele was the minor allele of the m.16189 T/C polymorphism: all patients were shown to have a homoplasmic polymorphism (m.16189 TT/CC), except for a male patient (EMBL via Webin, PRJEB25168; ERZ493062), who was shown to have a heteroplasmic polymorphism (m.16189 TC). Sequence analysis of the D-loop region revealed no statistically significant differences between patients with NASH and healthy individuals in terms of genotype and allele frequencies (P = 0.532). The TT, TC, and CC genotype frequencies are depicted in Table 2. Similarly, PCR-RFLP with *Mnl*I analysis of the m.16189 T/C variant revealed no significant differences between patients with NASH and healthy individuals (P = 0.740), but individuals with atypical RFLP band patterns were identified (Fig. 1c, d). Samples with an atypical *Mnl*I PCR-RFLP band pattern used to identify the m.16189 T/C variant precluded the determination of actual genotype prevalence at this position. After excluding individuals with an atypical pattern, there was a 100% concordance between the results of PCR-RFLP and DNA sequence analysis.

The PCR-RFLP with the *Mnl*I method cannot accurately detect the m.16189 T/C genotype

The results of PCR-RFLP with *Mnl*I analysis that was used to identify m.16189 T/C nucleotide variations were compared with the results of mtDNA sequence analysis. Indeed, PCR-RFLP with *Mnl*I analysis had failed to identify the genotype in 7% of the study samples owing to the presence of atypical patterns. Furthermore, it had yielded incorrect results in the genotyped samples. For instance, a sample from a male patient (EMBL via Webin, PRJEB25168; ERZ493071) that showed normal band pattern in RFLP analysis did not match the results of sequence analysis. In an individual harboring the combination of m.C16187T-m.T16189C variants, a new *Mnl*I enzyme recognition sequence emerged from relevant nucleotide variations and resulted in cleavage. RFLP analysis identified TT genotype for the m.16189 T/C in an individual who harbored CC genotype. These discrepancies indicated that PCR-RFLP with *Mnl*I analysis incorrectly identified the genotypes. Nucleotide variations among individuals showing atypical RFLP band patterns in PCR-RFLP with *Mnl*I analysis performed to identify the m.16189 T/C polymorphism were examined. The results of sequence analysis corroborated the results of RFLP analysis, and 29 samples with atypical patterns (AP) were divided into six groups (AP1a (10/29, 34.48%); AP1c, AP1d, AP1f, AP2, AP3, AP5b (1/29, 3.44%); AP1b, AP1e, AP5a, AP6 (2/29, 6.89%); and AP4 (5/29, 17.24%)) (Supplementary Fig. S1).

Number of mtDNA OriB variants in groups

In patients with NASH and healthy individuals, oriB variants were found to be wild type, rare, and np 16184-16193 poly-cytosine tract (poly-C), as mentioned in Supplementary Fig. S2.

m.A16318C is the candidate variant of NASH disease

In our study group, the m.16318 A>C, CC genotype detected only in patients with NASH at a frequency close to statistical significance was determined to be a candidate variant for predicting the development of NASH (NASH, 5/150; control, 0/150; P = 0.060). In addition, there was a significant association between the mtDNA D-loop variants and histological subgroups of NASH (fibrosis, lobular inflammation, and

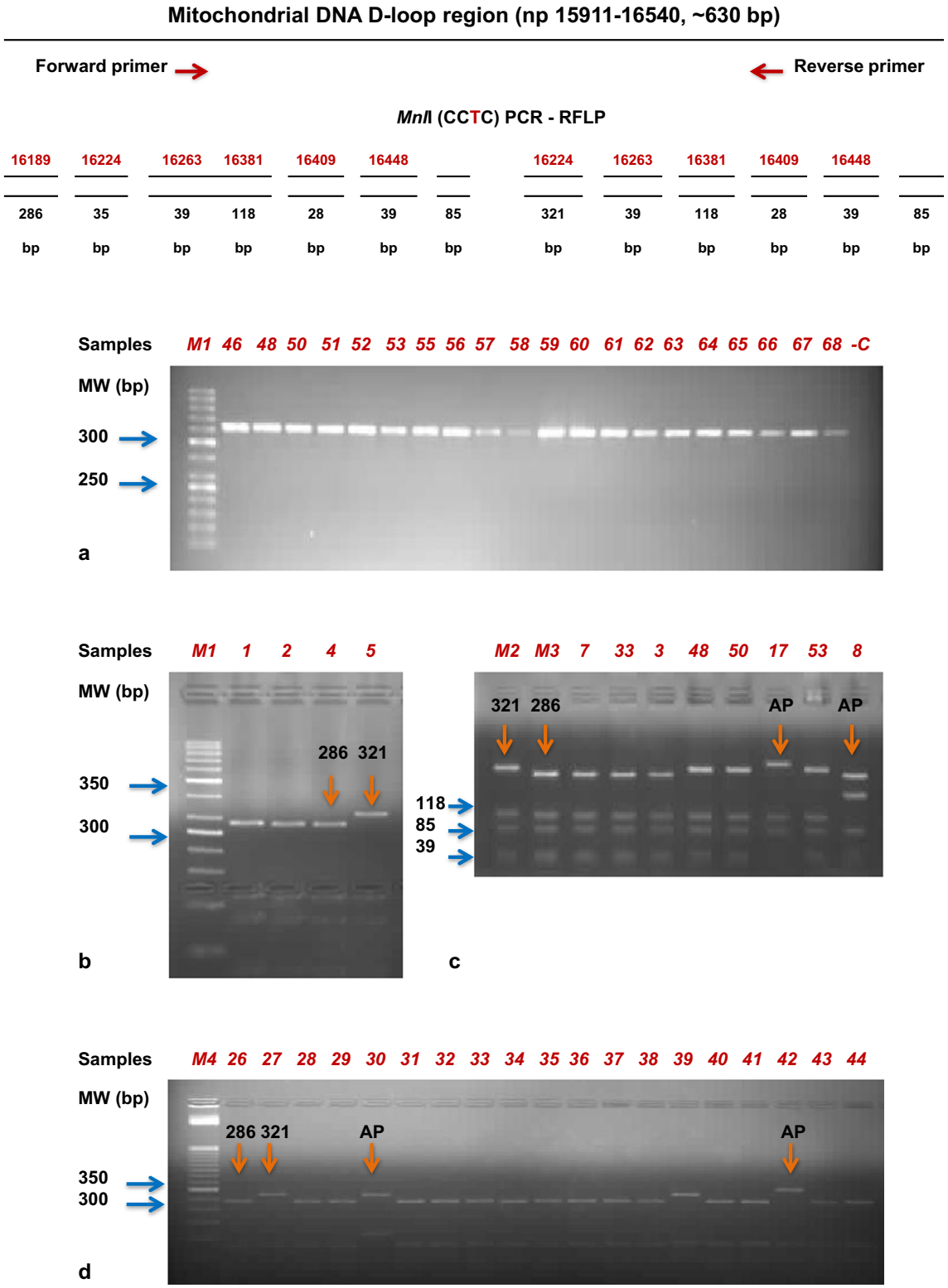


Fig. 1 *Mnl*I PCR-RFLP gel profiles. **a** PCR amplicons. **b**, **c**, **d** Normal and atypical bands of *Mnl*I-cleavage. Fragments with lengths of 286, 118, 84, 39 (2), and 28 bp indicate T allele; fragments with lengths of 321, 118, 84, 39 (2), and 28 bp indicate C allele. AP, atypical pattern; bp, base pair; −C, negative control; MW, molecular weight. M1 is 50 bp DNA ladder (Thermo Scientific, US); M2 and M3 are (np 321 and np 286, respectively) internal markers; M4 is 50 bp DNA ladder (Invitrogen, US)

Table 2 Genotype, allele, and atypical pattern frequency distribution of m.16189 T/C

Item	Patient groups (n = 150)	Control groups (n = 150)	P value
*Mnl*I PCR-RFLP Analysis			
Genotype frequency			
Genotype of the major allele (TT)	119 (85.61%)	117 (83.57%)	0.741
Genotype of the minor allele (CC)	20 (14.38%)	23 (16.42%)	
Atypical pattern frequency	11 (7.33%)	10 (6.66%)	1.000
Total	150	150	
Allele frequency			
T allele	238	234	1.000
C allele	40	46	
Total	278	280	
Sequence Analysis			
Genotype frequency			
Genotype of the major allele (TT)	115 (76.66%)	112 (74.66%)	0.532
Genotype of the minor allele (CC)	34 (22.66%)	38 (25.33%)	
Genotype of the heteroplasmy (TC)	1 (0.66%)	0 (0%)	
Total	150	150	
Atypical pattern frequency	14 (9.33%)	15 (10%)	1.000
Allele frequency			
T allele	231	224	0.567
C allele	69	76	
Total	300	300	

Data are presented as count and percent. Values derived from Fisher's Exact test by GraphPad QuickCalcs

histological steatosis) and NASH-complicated diseases (MetS, T2DM, and hypothyroidism) (Table 3).

m.16129 G>A, AA genotype is associated with fibrosis in NASH patients

In patients with NASH, there was a statistically significant association between the m.16129 G>A, AA genotype and advanced-stage fibrosis (F3–F4) (P = 0.046; OR = 5.436; 95% CI = 1.081–27.347).

m.16249 T>C, CC genotype is associated with lobular inflammation and histological steatosis in NASH patients

There was a significant association between the m.16249 T>C, CC genotype, and advanced-stage lobular inflammation (LI2–LI3) in patients with NASH (P = 0.021; OR = 13.199; 95% CI = 0.7469–233.24). There was a statistically significant association between the m.16249 CC genotype and advanced-stage histological steatosis (HS3) in patients with NASH (P = 0.028; OR = 13; 95% CI = 0.7082–238.63). Among patients with NASH, those carrying the m.16249 CC genotype had a 13.199 times increased risk of developing advanced-stage lobular inflammation and a 13 times increased risk of developing advanced-stage histological steatosis.

m.16296 C>T, TT genotype is associated with hypothyroidism in NASH patients

The presence of the m.16296 C>T, TT genotype among patients with NASH was compared between those with and without hypothyroidism. The frequency of the m.16296 TT genotype was significantly higher in patients with hypothyroidism than in patients without it (P = 0.030; OR = 4.132; 95% CI = 1.217–14.033).

m.16163 A>G, GG genotype is associated with MetS in NASH patients

In patients with NASH, the presence of the m.16163 A>G, GG genotype was compared between those with and without MetS. The genotype frequency of this variant was significantly higher in patients with MetS than in patients without it. Therefore, this polymorphism could be regarded as a risk factor for developing MetS in patients with NASH (P = 0.046; OR = 9.686; 95% CI = 0.5421–173.05).

m.16294 C>T, TT genotype is associated with MetS in NASH patients

In patients with NASH, the frequency of the m.16294 C>T, TT genotype was compared between those with and without

Table 3 Results of mtDNA D-loop sequence analysis: relationship between NASH group and mtDNA variants

Variables	Genotypes		Values		
m.16318 A>C	AA	CC	*P* value	OR	95% CI
NASH	145	5	0.060	-	-
Control	150	0			
m.16129 G>A	GG	AA	*P* value	OR	95% CI
Fibrosis F0	53	3	0.047	5.436	1.081–27.347
Fibrosis F3–F4	13	4			
m.16163 A>G	AA	GG	*P* value	OR	95% CI
Metabolic syndrome (no)	56	0	0.046	9.686	0.5421–173.05
Metabolic syndrome (yes)	87	7			
m.16249 T>C	TT	CC	*P* value	OR	95% CI
Lobular inflammation LI1	62	0	0.021	13.199	0.7469–233.24
Lobular inflammation LI2–LI3	80	8			
Histologic steatosis HS1	40	0	0.028	13	0.7082–238.63
Histologic steatosis HS3	40	6			
m.16256 C>T	CC	TT/CT	*P* value	OR	95% CI
Type II Diabetes (no)	100	3	0.021	5.263	1.252–22.119
Type II Diabetes (yes)	38	6			
m.16294 C>T	CC	TT	*P* value	OR	95% CI
Metabolic syndrome (no)	51	5	0.029	3.117	1.107–8.778
Metabolic syndrome (yes)	72	22			
m.16296 C>T	CC	TT	*P* value	OR	95% CI
Hypothyroidism (no)	119	8	0.030	4.132	1.217–14.033
Hypothyroidism (yes)	18	5			

Values derived from Fisher's Exact test by GraphPad QuickCalcs. *NASH*, non-alcoholic steatohepatitis; *F*, fibrosis; *LI*, lobular inflammation; *HS*, histologic steatosis; *OR*, odds ratio; *CI*, confidence interval

MetS. This genotype was a risk factor for developing MetS in patients with NASH (P = 0.028; OR = 3.117; 95% CI = 1.107–8.778).

m.16256 C>T, TT+CT genotype is associated with T2DM in NASH patients

In patients with NASH, the frequency of the m.16256 C>T, TT+CT genotypes were compared between those with and without T2DM, and the frequency of the m.16256 TT+CT genotypes were significantly higher in patients with T2DM (P = 0.021; OR = 5.263; 95% CI = 1.252–22.119).

Other nucleotide alterations detected in the mtDNA D-loop

The D-loop sequence analysis of the mtDNA control region covering the 300 subjects of this study identified 140 nucleotide variations (Table 4). Among these, the prevalence of 19 nucleotide variations was > 5%. The literature was reviewed using the Allele Search Tool on the MITOMAP website for results of the sequence analysis obtained from our study samples. One new mutation (m.C16053G, ERZ493061) in a healthy female (Fig. 2) and three new heteroplasmic nucleotide variations (m.16129 G-GA, ERZ496726; m.16234 C-CT, ERZ493063; and m.16256 C-CT, ERZ493064) were novel (Fig. 3a–c) and were introduced into the literature. In addition, two previously reported heteroplasmic variations (m.16399 A-AG, ERZ493068; and m.16189 T-TC, ERZ493062) were identified (Fig. 3d and e). Interestingly, three trinucleotide variations (m.C16187-m.C16188-m.T16189, TAC, ERZ493066; m.C16187-m.C16188-m.T16189, TGC, ERZ493067; and m.C16294-m.C16295-m.C16296, TTT, ERZ493065) in three different females (Fig. 4a–c) and four dinucleotide variations (m.A16182C-m.A16183C, ERZ651039; m.A16194C-m.T16195C, ERZ493128; m.C16294T-m.C16295T, ERZ651043; and A16318T-G16319A, ERZ496697) in different individuals (Supplementary Table S1) were identified.

Haplogroup findings of partial mtDNA D-loop sequencing

Nucleotide variations in the mtDNA D-loop region determined in this study conducted in the Turkish population were confirmed by haplotype analyses using HaploGrep 2. All 300 subjects belonging to the R macro-haplogroup were found to be in the haplogroup H2a2a1 and subhaplogroup H2a2a1d-T16172C (four individuals); H2a2a1g-T16189C (15

Table 4 Results of mtDNA D-loop sequence analysis: common and rare nucleotide alterations

19 common polymorphisms (> 5%) (n = 300)			121 rare nucleotide variations (< 5%) (n = 300)				
mtSNPs	Patient	Control	mtSNPs in patient and control groups				
C16069T	15 (10%)	12 (8%)	A16051G	T16172C	T16217C	C16261T	G16319A
T16093C	4 (0.6%)	8 (5.33%)	*C16053G**	C16173T	C16218T	C16262T	C16320T
T16126C	42 (28%)	32 (21.33%)	A16058C	C16174T	A16219G	C16264G	T16324C
G16129A	14 (9.33%)	15 (10%)	A16066G	C16176G	A16220C	C16264T	T16325C
G16145A	10 (6.66%)	7 (4.66%)	C16067T	C16176T	C16221T	A16265G	C16327T
A16183C	14 (9.33%)	19 (12.66%)	A16070G	T16178C	C16222T	C16266T	C16328A
T16189C	34 (22.66%)	38 (25.33%)	C16071T	C16179T	A16227G	C16270T	C16344T
C16223T	23 (15.33%)	29 (19.33%)	T16075C	A16182C	A16230G	T16271C	T16352C
T16224C	15 (10%)	14 (9.33%)	A16078G	C16184T	T16231C	A16272G	C16353A
T16249C	8 (5.33%)	6 (4%)	T16086C	C16186T	C16232A	G16274A	C16355T
T16256C	8 (5.33%)	9 (6%)	T16092C	C16187T	C16234T	C16287T	T16356C
C16278T	5 (3.33%)	10 (6.66%)	T16102C	C16188G	C16234G	T16288C	T16357C
C16292T	8 (5.33%)	8 (5.33%)	C16111T	C16188A	*C16234CT**	A16289G	T16359C
C16294T	27 (18%)	20 (13.33%)	C16114A	*T16189TC**	A16235G	C16290T	T16368C
C16296T	13 (8.66%)	9 (6%)	T16124C	C16192T	C16239T	C16291T	G16390A
T16304C	9 (6%)	11 (7.33%)	*G16129GA**	C16193T	A16240G	A16293G	G16391A
T16311C	30 (20%)	29 (19.33%)	T16136C	A16194C	C16242T	C16295T	A16399G
A16343G	10 (6.66%)	13 (8.66%)	C16147A	T16195C	T16243C	T16298C	*A16399AG**
T16362C	9 (6%)	14 (9.33%)	C16148T	C16201T	G16244A	A16299G	T16443C
			C16150T	A16203G	C16248T	A16300G	C16465T
			G16153A	A16206C	C16250T	C16301T	A16482G
			A16162G	T16209C	*C16256CT**	A16309G	
			A16163G	G16213A	C16257A	A16316G	
			C16168T	C16214T	A16258C	A16318T	
			C16169T	A16215G	C16260T	A16318C	

*Novel variant and heteroplasmic variants are indicated as italic. Data are presented as count, and frequencies (%). *mtSNPs*, mitochondrial DNA single-nucleotide polymorphisms; *n*, number of individuals with the respective polymorphism

individuals); and H2a2a1h-C16320T (six individuals) (Supplementary Table S2).

Predicted microRNA findings based on target sequences

Supplementary Table S3 presents mature and hairpin miRNAs obtained from the miRBase database for both alleles of the variant of interest by targeting the D-loop variants (m.A16318C; m.G16129A; m.A16163G; m.T16249C; m.C16256T; and m.C16294T; m.C16296T) associated with NASH.

Discussion

Mitochondrial dysfunction plays an important role in the development of NASH and in the appearance of clinical findings [13, 14]. Because the D-loop region is the functional unit responsible for mitochondrial replication and transcription, the effect of nucleotide variations in this region on mitochondrial dysfunction has been the focus of investigation. Our findings suggest that nucleotide variations in the mtDNA D-loop region in patients with NASH are associated with histopathological patterns of NASH and NASH-complicated diseases. In addition, miRNAs targeting the D-loop region were identified through the variants of mtDNA associated with the clinical data of patients with NASH; these warrant further research to investigate their effects on NASH.

A study of the association between the m.16189 T/C, a common variation in the mtDNA D-loop region, and the components of NASH, including IR, T2DM, obesity, and MetS, is currently underway. In previous studies, PCR-RFLP analysis has been widely employed to identify variations using *Mnl*I restriction endonuclease enzymes, which are specific to the reference sequence. The m.C16189T variant that interrupts the poly-C homopolymeric region in the HVI region (np 16184-16193) of the D-loop region forms [15] the target for

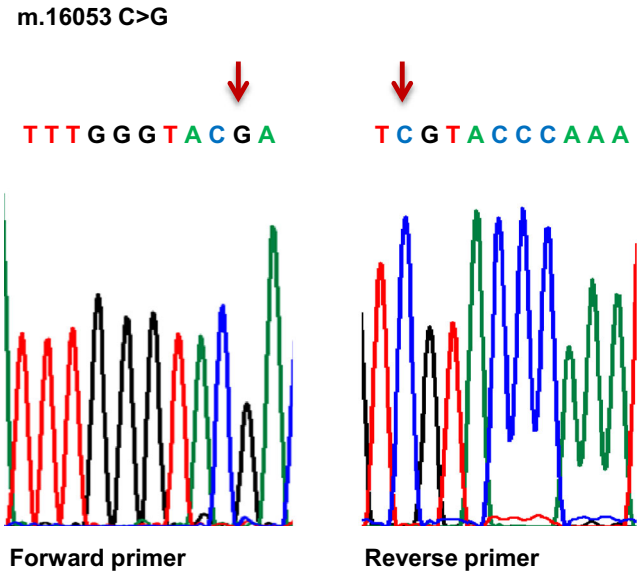


Fig. 2 Chromatogram images of novel mutation m.16053 C>G. The m.16053 C>G (ERZ493061) reported to MITOMAP

its *Mnl*I enzyme and provides a specific RFLP band pattern for T/C alleles.

The present study demonstrated that *Mnl*I PCR-RFLP analysis may not always identify the m.16189 T/C variant due to increases or decreases in *Mnl*I enzyme recognition sites induced by various nucleotide variations occurring in the HVI of the D-loop region. The mtDNA control region is not suitable for PCR-RFLP analysis of single-nucleotide polymorphisms (SNPs). DNA sequence analysis is the gold standard method, the use of which is one of the strengths of the present study. Therefore, the results of previous studies of the m.16189 T/C polymorphism using *Mnl*I PCR-RFLP analysis should be reviewed for their accuracy. As a result, studies using methods other than sequence analysis to identify SNPs in the control region are weak and their results are controversial, as mtDNA is a hotspot for alterations through mutation.

Genotyping studies using *Mnl*I PCR-RFLP analysis to identify the m.16189 T/C polymorphism have reported findings similar to our results. Gibson et al. have reported that approximately 10% of the samples showed different band patterns due to the gain or loss of the *Mnl*I areas in the mtDNA HV region of the m.16189 T/C variant [16]. Tang et al. have analyzed the samples using sequence analysis, which were observed to contain "abnormal products" for the m.16189 T/C variant in *Mnl*I PCR-RFLP analysis [17]; meanwhile, Park et al. have reported that the *Mnl*I restriction enzyme was not specific to the m.16189 T/C variant and tended to identify other T areas at np 16184-16193 of the poly-C region, resulting in incorrect classifications [18]. In an *Mnl*I PCR-RFLP analysis of the D-loop region (np 16106-16437), Ye et al. have shown five common RFLP patterns to be "somatic *Mnl*I site mutations" [19]. As a contribution to the studies in the literature, we emphasized the drawbacks of *Mnl*I PCR-RFLP analysis by testing the hypothesis and the classified atypical *Mnl*I PCR-RFLP samples in our study. Contrary to the significant association reported with T2DM in a very recent meta-analysis evaluating the association of the m.T16189C polymorphism with cancer and T2DM [20], the present study reported no statistically significant association in either the patient group or the control group, or subgroups of patients with NASH. In the Turkish population, this variant is not distinctive for NASH and its risk factors. However, in another recent in vivo study using cell lines, samples containing short sequences of the OriB variant (np 16184-16193 poly-C tract), independent of nuclear DNA, exhibited resistance to cisplatin, a drug used in cancer therapy [21]. Therefore, we believe that although the m.T16189C variant does not play a diagnostic role in NASH in the Turkish population, differences in the nucleotide length of the mtDNA poly-C region and other D-loop nucleotide variations might be important in terms of developing individualized treatments for NASH. Furthermore, in contrast to publications investigating the relationship between oriB variants and MetS [22] and T2DM [15], different variants other than wild type and T^2 were detected in this study.

Based on the partial mtDNA D-loop profile, the R macrohaplogroup of the Turkish population falls into the H2a2a1 haplogroup. We identified one novel mutation (m.C16053G) and three novel heteroplasmic polymorphisms (m.16129 G-GA, m.16234 C-CT, and m.16256 C-CT) in the mtDNA control region, which finding is a significant contribution to the current knowledge.

In the 300 samples of our study, we determined specific sequences of the Turkish population with haplotypic nucleotides showing 29 atypical RFLP band patterns, which are different from the reference sequence. In a study by Malyarchuk and Derenko, the m.T16126C and m.C16294T variants were identified in the control regions, which are specific to the Caucasian race and located outside the *Mnl*I enzyme recognition sites [23]. Similarly, we found a high prevalence of these variants (total, 41/300, 13.66%; NASH, 24/150, 16%; control, 17/150, 11.33%).

Di- or trinucleotide variations are rare in nDNA. Although the prevalence of SNPs is higher in mtDNA than in nDNA, di- and trinucleotide variations occur less frequently and are clinically significant in terms of their possible association with diseases. We detected three trinucleotide variations in three individuals and four dinucleotide variations in eight individuals, which were previously reported in populations included in genetic studies of mtDNA [24–26]. Details regarding the function and location of mtDNAs of these variants are presented in Supplementary Table S1.

To the best of our knowledge, the present pilot study is the first to identify six variants (m.G16129A; m.A16163G; m.T16249C; m.C16256T; m.C16294T; and m.C16296T) that

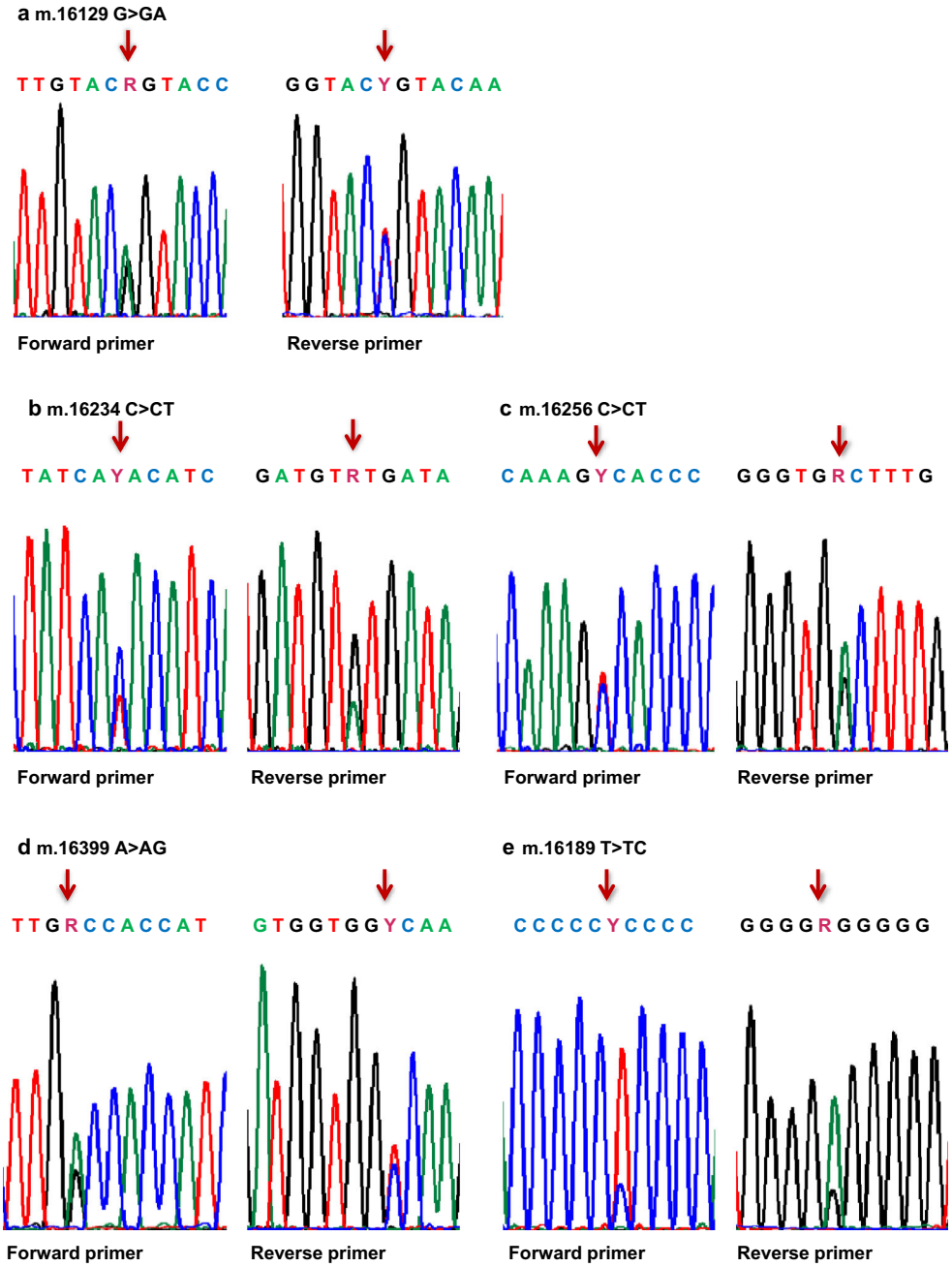


Fig. 3 Chromatogram images of three novel heteroplasmy and other two heteroplasmy. Three novel heteroplasmy, **a** m.16129 G>GA, ERZ496726. **b** m.16234 C>CT, ERZ493063. **c** m.16256 C>CT, ERZ493064 (reported to MITOMAP). Other two heteroplasmy, **d** m.A16399 A>AG, ERZ493068; **e** m.T16189 T>TC, ERZ493062

were significantly associated with the subgroups of NASH and NASH-complicated diseases. The m.16318 A>C, CC (P = 0.060) variant identified only in the NASH group at a level approaching statistical significance were able to predict NASH; however, the pathogenesis underlying this remains unknown. Presumably, low prevalence of the m.A16318C variant in the patient and control groups (< 5%) may have affected statistical significance; thus, increasing the sample size may yield results approaching statistical significance. In conclusion, the effects of the m.A16318C variant should be investigated in a larger population, as well as in other ethnic groups.

In previous studies, the m.G16129A variant functioning in mtDNA loci HRV-I as MT-TAS2 was associated with headache and cyclic vomiting syndrome [27], while the m.T16249C variant was associated with tongue cancer [28] and presbycusis [29].

Based on our promising results, mature and hairpin miRNAs, selected by considering those with an E value of < 10 to be significant, were obtained using the miRNAs target prediction database for the mtDNA D-loop variants detected. Two studies on hepatocellular carcinoma have suggested that hsa-miR-608 confers anticancer effects. In this study [30, 31], we matched hsa-mir-608 with the m.C16256T, CC genotype. In the presence of m.C16256T, TT genotype, which is a risk factor for patients

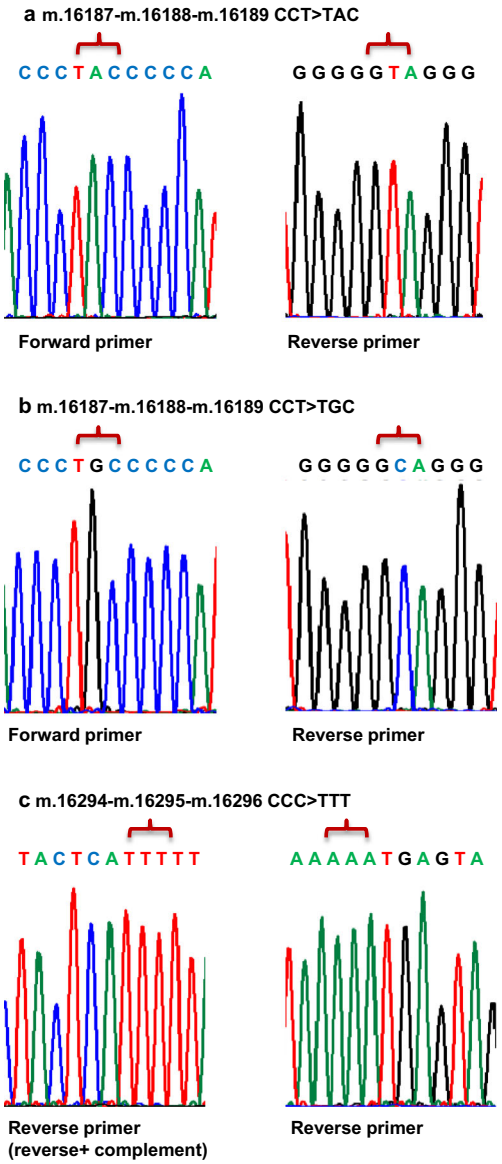


Fig. 4 Chromatogram images of three trinucleotide variations. Three trinucleotide variations, **a** m.16187-m.16188-m.16189 CCT>TAC, ERZ493066; **b** m.16187-m.16188-m.16189 CCT>TGC, ERZ493067; **c** m.16294-m.16295-m.16296 CCC>TTT, ERZ493065

with NASH with T2DM, hsa-mir-608 was eliminated. Therefore, hsa-mir-608 may confer protective effects in patients with NASH with the m.C16256T, TT genotype. Unfortunately, no studies have reported any association of other detected miRNAs with the histological subgroups of NASH. Thus, in vitro and/or in vivo studies aimed at the entire D-loop sequence, accompanied by advanced bioinformatics analyses, are warranted.

The present study should be interpreted in the context of several limitations. First, our study is observational in nature and lacks a validation cohort. Consequently, no cause-effect inferences can be made. Second, we were unable to perform liver biopsies to rule out the presence of NAFLD in control subjects for ethical reasons. Third, we acknowledge the lack of basic demographic and laboratory data of the control group as an important caveat inherent in our study. Although transaminase levels were determined in all controls and invariably found to be within the normal reference range, we did not measure other laboratory parameters in the control group. For this reason, we did not tabulate laboratory data of the control subjects. Fourth, it can be argued that NAFLD histology was assessed using two distinct scoring systems. However, the two scores have complementary features. Specifically, NAS should not be considered a diagnostic tool, but rather a continuous scale for activity assessment once the diagnosis of NASH has been histologically established. SAF has been introduced more recently to capture the heterogeneity of NAFLD. This simple scoring system which has been specifically designed to dissociate the grade of steatosis, the grade of activity, and the stage of fibrosis may thus be more relevant than simply dichotomizing cases according to the presence or absence of NASH. By separately describing the main histopathologic features in a practical manner, the SAF score also allows an easier comparison between biopsies and histological changes observed in paired biopsies during clinical trials. Finally, our sample included subjects of Turkish nationality, so that results cannot be extrapolated to populations with different ethnic backgrounds. In addition, because our clinic is located within a tertiary care center, it is possible that our study patients do not entirely reflect the general NAFLD population.

In conclusion, the present study emphasizes the importance of control region nucleotide variations in addition to the variants of mtDNA D-loop known to be associated with various diseases. Furthermore, the importance of nucleotide variations has been emphasized on the basis of the future verification of potential miRNAs determined by us through the accumulated significant study results. Analysis of nucleotide variations in the mtDNA control region demonstrated significant results, which can contribute to the elucidation of the pathogenesis of NASH. In the future, we intend to develop miRNA-based anti-miRNA therapies for preventing the development and progression of NASH.

Acknowledgments We thank all volunteers who agreed to participate in the study. The authors are grateful to Dr. Ruslan Asadov for performing liver biopsies and Prof. Dr. Cigdem Ataizi Celikel for the histological evaluation of liver specimens.

Funding information This study was financially supported by the Scientific Research Project Coordination Unit to Marmara University, Istanbul, Turkey (Grant No. SAG-C-YLP-131016-0439; Grant No. SAG-D-090517-0263).

Compliance with ethical standards

Conflict of interest The authors declare that they have no conflict of interest.

Ethical approval The study protocol was approved by the local ethics committee (Marmara University School of Medicine, Istanbul, Turkey; protocol no. 09.2016.097; approval date, 29 January 2016).

Informed consent Informed consent was obtained from all individual participants included in the study.

References

1. Younossi ZM, Koenig AB, Abdelatif D, Fazel Y, Henry L, Wymer M (2016) Global epidemiology of nonalcoholic fatty liver disease-meta-analytic assessment of prevalence, incidence, and outcomes. Hepatology 64(1):73–84. https://doi.org/10.1002/hep.28431
2. Spengler EK, Loomba R (2015) Recommendations for diagnosis, referral for liver biopsy, and treatment of nonalcoholic fatty liver disease and nonalcoholic steatohepatitis. Mayo Clin Proc 90(9): 1233–1246. https://doi.org/10.1016/j.mayocp.2015.06.013
3. Della Pepa G, Vetrani C, Lombardi G, Bozzetto L, Annuzzi G, Rivellese AA (2017) Isocaloric dietary changes and non-alcoholic fatty liver disease in high cardiometabolic risk individuals. Nutrients 9(10). https://doi.org/10.3390/nu9101065
4. Oseini AM, Sanyal AJ (2017) Therapies in non-alcoholic steatohepatitis (NASH). Liver Int 37(Suppl 1):97–103. https://doi.org/10.1111/liv.13302
5. Paradies G, Paradies V, Ruggiero FM, Petrosillo G (2014) Oxidative stress, cardiolipin and mitochondrial dysfunction in non-alcoholic fatty liver disease. World J Gastroenterol 20(39):14205–14218. https://doi.org/10.3748/wjg.v20.i39.14205
6. Galtier N, Enard D, Radondy Y, Bazin E, Belkhir K (2006) Mutation hot spots in mammalian mitochondrial DNA. Genome Res 16(2):215–222. https://doi.org/10.1101/gr.4305906
7. Jemt E, Persson Ö, Shi Y, Mehmedovic M, Uhler JP, Dávila López M, Freyer C, Gustafsson CM, Samuelsson T, Falkenberg M (2015) Regulation of DNA replication at the end of the mitochondrial D-loop involves the helicase TWINKLE and a conserved sequence element. Nucleic Acids Res 43(19):9262–9275. https://doi.org/10.1093/nar/gkv804
8. Yilmaz Y, Eren F, Colak Y, Senates E, Celikel CA, Imeryuz N (2012) Hepatic expression and serum levels of syndecan 1 (CD138) in patients with nonalcoholic fatty liver disease. Scand J Gastroenterol 47(12):1488–1493. https://doi.org/10.3109/00365521.2012.725093
9. Yilmaz Y, Eren F (2019) Serum biomarkers of fibrosis and extracellular matrix remodeling in patients with nonalcoholic fatty liver disease: association with liver histology. Eur J Gastroenterol Hepatol 31(1):43–46. https://doi.org/10.1097/MEG.0000000000001240
10. Kleiner DE, Brunt EM, Van Natta M, Behling C, Contos MJ, Cummings OW, Ferrell LD, Liu YC, Torbenson MS, Unalp-Arida A, Yeh M, McCullough AJ, Sanyal AJ (2005) Nonalcoholic steatohepatitis clinical research network. Design and validation of a histological scoring system for nonalcoholic fatty liver disease. Hepatology 41(6):1313–1321. https://doi.org/10.1002/hep.20701
11. Bedossa P, Poitou C, Veyrie N, Bouillot JL, Basdevant A, Paradis V, Tordjman J, Clement K (2012) Histopathological algorithm and scoring system for evaluation of liver lesions inmorbidly obese patients. Hepatology 56(5):1751–1759. https://doi.org/10.1002/hep.25889
12. Aral C, Akkiprik M, Caglayan S, Atabey Z, Ozişik G, Bekiroglu N, Ozer A (2011) Investigation of relationship of the mitochondrial DNA 16189 T>C polymorphism with metabolic syndrome and its associated clinical parameters in Turkish patients. Hormones (Athens) 10(4):298–303. https://doi.org/10.14310/horm.2002.1321
13. Begriche K, Massart J, Robin MA, Bonnet F, Fromenty B (2013) Mitochondrial adaptations and dysfunctions in nonalcoholic fatty liver disease. Hepatology 58(4):1497–1507. https://doi.org/10.1002/hep.26226
14. Verbeek J, Lannoo M, Pirinen E, Ryu D, Spincemaille P, Vander Elst I, Windmolders P, Thevissen K, Cammue BP, van Pelt J, Fransis S, Van Eyken P, Ceuterick-De Groote C, Van Veldhoven PP, Bedossa P, Nevens F, Auwerx J, Cassiman D (2015) Roux-en-y gastric bypass attenuates hepatic mitochondrial dysfunction in mice with non-alcoholic steatohepatitis. Gut 64(4):673–683. https://doi.org/10.1136/gutjnl-2014-306748
15. Chinnery PF, Elliott HR, Patel S, Lambert C, Keers SM, Durham SE, McCarthy MI, Hitman GA, Hattersley AT, Walker M (2005) Role of the mitochondrial DNA 16184-16193 poly-C tract in type 2 diabetes. Lancet 366(9497):1650–1651. https://doi.org/10.1016/S0140-6736(05)67492-2
16. Gibson AM, Edwardson JA, Turnbull DM, McKeith IG, Morris CM, Chinnery PF (2004) No evidence of an association between the T16189C mtDNA variant and late onset dementia. J Med Genet 41(1):e7. https://doi.org/10.1136/jmg.2003.010983
17. Tang DL, Zhou X, Li X, Zhao L, Liu F (2006) Variation of mitochondrial gene and the association with type 2 diabetes mellitus in a Chinese population. Diabetes Res Clin Pract 73(1):77–82. https://doi.org/10.1016/j.diabres.2005.12.001
18. Park KS, Chan JC, Chuang LM, Suzuki S, Araki E, Nanjo K, Ji L, Ng M, Nishi M, Furuta H, Shirotani T, Ahn BY, Chung SS, Min HK, Lee SW, Kim JH, Cho YM, Lee HK (2008) Study Group of Molecular Diabetology in Asia. A mitochondrial DNA variant at position 16189 is associated with type 2 diabetes mellitus in Asians. Diabetologia 51(4):602–608. https://doi.org/10.1007/s00125-008-0933-z
19. Ye C, Shu XO, Pierce L, Wen W, Courtney R, Gao YT, Zheng W, Cai Q (2010) Mutations in the mitochondrial DNA D-loop region and breast cancer risk. Breast Cancer Res Treat 119(2):431–436. https://doi.org/10.1007/s10549-009-0397-y
20. Kumari T, Vachher M, Bansal S, Bamezai RNK, Kumar B (2018) Meta-analysis of mitochondrial T16189C polymorphism for cancer and type 2 diabetes risk. Clin Chim Acta 482:136–143. https://doi.org/10.1016/j.cca.2018.03.041
21. Amo T, Kamimura N, Asano H, Asoh S, Ohta S (2017) Cisplatin selects short forms of the mitochondrial DNA OriB variant (16184-16193 poly-cytosine tract), which confer resistance to cisplatin. Sci Rep 7:46240. https://doi.org/10.1038/srep46240
22. Skuratovskaia DA, Sofronova JK, Zatolokin PA, Vasilenko MA, Litvinova LS, Mazunin IO (2017) The association of the mitochondrial DNA oriB variants with metabolic syndrome. Biomed Khim 63(6):533–538. https://doi.org/10.18097/PBMC20176306533
23. Maliarchuk BA, Derenko MV (1997) Analysis of nucleotide combinations in types of the control region of human mitochondrial DNA. Genetika 33(3):387–392
24. Watkins WS, Bamshad M, Dixon ME, Bhaskara Rao B, Naidu JM, Reddy PG, Prasad BV, Das PK, Reddy PC, Gai PB, Bhanu A, Kusuma YS, Lum JK, Fischer P, Jorde LB (1999) Multiple origins of the mtDNA 9-bp deletion in populations of South India. Am J Phys Anthropol 109(2):147–158. https://doi.org/10.1002/(SICI)1096-8644(199906)109:2<147::AID-AJPA1>3.0.CO;2-C
25. Monsalve MV, Hagelberg E (1997) Mitochondrial DNA polymorphisms in Carib people of Belize. Proc Biol Sci 264(1385):1217–1224. https://doi.org/10.1098/rspb.1997.0168
26. Quintana-Murci L, Chaix R, Wells RS, Behar DM, Sayar H, Scozzari R, Rengo C, Al-Zahery N, Semino O, Santachiara-Benerecetti AS, Coppa A, Ayub Q, Mohyuddin A, Tyler-Smith

C, Qasim Mehdi S, Torroni A, McElreavey K (2004) Where west meets east: the complex mtDNA landscape of the southwest and Central Asian corridor. Am J Hum Genet 74(5):827–845. https://doi.org/10.1086/383236

27. Wang Q, Ito M, Adams K, Li BU, Klopstock T, Maslim A, Higashimoto T, Herzog J, Boles RG (2004) Mitochondrial DNA control region sequence variation in migraine headache and cyclic vomiting syndrome. Am J Med Genet A 131(1):50–58. https://doi.org/10.1002/ajmg.a.30323
28. Zhou HY, Shu HY, Dai J, Li HC, Tang L, Wang HW, Ni B (2018) Maternal genetic backgrounds contribute to the genetic susceptibility of tongue cancer patients in Hunan, central of China. Mitochondrial DNA A DNA Mapp Seq Anal 29(3):347–352. https://doi.org/10.1080/24701394.2016.1278539
29. Falah M, Farhadi M, Kamrava SK, Mahmoudian S, Daneshi A, Balali M, Asghari A, Houshmand M (2017) Association of genetic variations in the mitochondrial DNA control region with presbycusis. Clin Interv Aging 12:459–465. https://doi.org/10.2147/CIA.S123278
30. Ma XP, Yu G, Chen X, Xiao Q, Shi Z, Zhang LY, Chen H, Zhang P, Ding DL, Huang HX, Saiyin H, Chen TY, Lu PX, Wang NJ, Yu H, Conran C, Sun J, Zheng SL, Xu J, Yu L, Jiang DK (2016) MiR-608 rs4919510 is associated with prognosis of hepatocellular carcinoma. Tumour Biol 37(7):9931–9942. https://doi.org/10.1007/s13277-016-4897-1
31. Wang K, Liang Q, Wei L, Zhang W, Zhu P (2016) MicroRNA-608 acts as a prognostic marker and inhibits the cell proliferation in hepatocellular carcinoma by macrophage migration inhibitory factor. Tumour Biol 37(3):3823–3830. https://doi.org/10.1007/s13277-015-4213-5

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.