RHEUMATOLOGY

Rheumatology 2020;59:2982–2991
doi:10.1093/rheumatology/keaa063
Advance Access publication 11 March 2020

Original article

Germline and somatic mtDNA mutation spectrum of rheumatoid arthritis patients in the Taizhou area, China

Juping Du[1], Sufei Yu[1], Donglian Wang[2], Shuaishuai Chen[1], Suyun Chen[1], Yufen Zheng[1], Na Wang[1], Shiyong Chen[1], Jun Li[1] and Bo Shen [1,2]

Abstract

Objective. Reactive oxygen species are believed to be involved in the onset of RA, and the association between nuclear-encoded mitochondrial respiratory chain-related variants and RA has recently been revealed. However, little is known about the landscape of mitochondrial DNA (mtDNA) variants in RA.

Methods. Next-generation sequencing was conducted to profile mtDNA germline and somatic variants in 124 RA patients and 123 age- and sex-matched healthy controls in the Taizhou area, China. Fisher's exact test, SKAT and SKAT-O were used for gene-burden tests to investigate RA-related variants of mitochondrial genes. Predictive tools were applied to evaluate the pathogenicity of mtDNA variants, and mtDNA haplogroups were assigned according to mtDNA mutations recorded in PhyloTree database. The frequency distribution of mtDNA haplogroups between the groups was compared using χ^2 analysis.

Results. We identified 467 RA-unique and 341 healthy control-unique mtDNA variants, with 443 common variants. Only *MT-ATP6* with a significant burden of variants was identified by Fisher's exact test, SKAT and SKAT-O, even after Bonferroni adjustment, and the enrichment variants in *MT-ATP6* was mainly driven by m.8830C>A, m.8833G>C and m.8843T>A variants. Besides, four frequently low-heteroplasmic variants including the three variants above and m.14135T>G of *MT-ND5* were detected in RA only; except for m.8830C>A, they are considered potential pathogenicity based on functional predictions. χ^2 analysis before Bonferroni adjustment revealed haplogroup F1/F1a to be negatively associated with RA ($P < 0.05$).

Conclusion. These results profiled the landscape of germline and somatic mtDNA variants in RA and supported the effects of mitochondrial genes on RA.

Key words: rheumatoid arthritis, mitochondrial DNA (mtDNA), reactive oxygen species (ROS), next-generation sequencing

CLINICAL SCIENCE

Rheumatology key messages

- Enrichment of variants in *MT-ATP6* was detected in RA patients.
- RA patients frequently carried four potential pathogenic variants in a low heteroplasmic state.
- Haplogroup F1 and its subhaplogroup F1a might be a protective factor for RA.

Introduction

RA is a multifactorial and chronic inflammatory autoimmune disease that is characterized by autoantibody production and persistent synovial inflammation, which leads to articular cartilage destruction and bone erosion. Although the mechanism involved in the development of RA is not completely understood, oxidative damage caused by overproduction of reactive oxygen species (ROS) has been believed to activate key signalling pathways, such as nuclear factor kappa B (NF-κB), mammalian target of rapamycin (mTOR) and antiapoptotic pathways, in RA onset [1–4].

Mitochondrial respiratory chain complexes have been considered as the major source of ROS production [5–8]. Recently, Mitsunaga [4] and colleagues revealed

[1]Department of Clinical Laboratory, Taizhou Hospital of Zhejiang Province, Taizhou Enze Medical Centre (Group), Linhai and [2]Department of Clinical Laboratory, Enze Hospital, Taizhou Enze Medical Centre (Group), Luqiao, Taizhou, Zhejiang Province, China

Submitted 12 August 2019; accepted 21 January 2020

Correspondence to: Bo Shen, Department of Clinical Laboratory, Taizhou Hospital of Zhejiang Province, Taizhou Enze Medical Centre (Group), 150 Ximen Road, Linhai, Zhejiang Province, China. E-mail: shenb@enzemed.com

that the aggregation of rare/low-frequency variants of the four nuclear-encoded mitochondrial respiratory chain-related proteins contribute to the risk of severe erosive RA, offering a genetic link between ROS and RA. Mitochondria also have their own genetic material, mitochondrial DNA (mtDNA), which is characterized by different haplogroups and the combination of a particular set of single nucleotide polymorphisms (SNPs) accumulated under selective forces of long periods of climatic and environmental changes. Distinct mtDNA haplogroups might be protective or risk factors in degenerative disorders, aging, diabetes and cancers [9–12]. However, few studies to date have reported mtDNA haplogroups or variants associated with RA [13, 14]. High mtDNA somatic mutations and oxidatively damaged mtDNA in the synovial tissue and fluid of RA patients have been reported [15–17]. Yu *et al.* found that the mtDNA G7778T variant in the murine mtAtp8 gene was able to promote ROS production, impairing mitochondrial function and increasing susceptibility to collagen-induced arthritis [18]. In addition, that group reported that the mtDNA G13708A variant was associated with an increased risk of RA [13]. Conversely, no associations between mtDNA haplogroups and RA and its manifestations were observed in a Denmark population [14]. As mtDNA haplogroups or variants present ethnic- and geographic-specific distributions, the relationship between mtDNA and RA might be essential for clarifying the situation in China. Moreover, all the studies above focused on either mtDNA mutation load or specific variants, whereas the general profiling of the entire mtDNA sequence has remained largely unclear. In this study, we gave a sight into the landscape of germline and somatic mtDNA mutations in RA and analysed the associations between mtDNA haplogroups and RA patients in China.

Methods

Study population

A total of 124 outpatients were consecutively randomly recruited at Taizhou Hospital of Zhejiang Province from October 2016 to December 2017. All cases fulfilled the classification criteria of the ACR and EULAR in 2010. Basic clinical characteristics and laboratory data were obtained from our medical records or free clinics. A total of 124 age- and sex-matched healthy controls (HC) without a personal or family history of autoimmune disease were selected from our Physical Examination Center; one patient was excluded due to poor DNA quality. Written informed consent to draw blood samples was obtained from each participant under protocols approved by the institutional review board of the Ethics Committee of Taizhou Hospital of Zhejiang Province.

mtDNA next-generation sequencing

Total DNA was extracted from blood samples using HiPure Blood DNA Midi Kit I (Magen, Beijing, China). Multiplexed DNA libraries were generated using MultipSeq® AI-MitoMulti Panel (iGeneTech, Beijing, China) following the manufacturer's instructions. The quality and concentration of the libraries were confirmed by agarose gel electrophoresis, Qubit 3.0 and the Agilent 2100 Bioanalyzer system and then sequenced on the Illumina HiSeq platform.

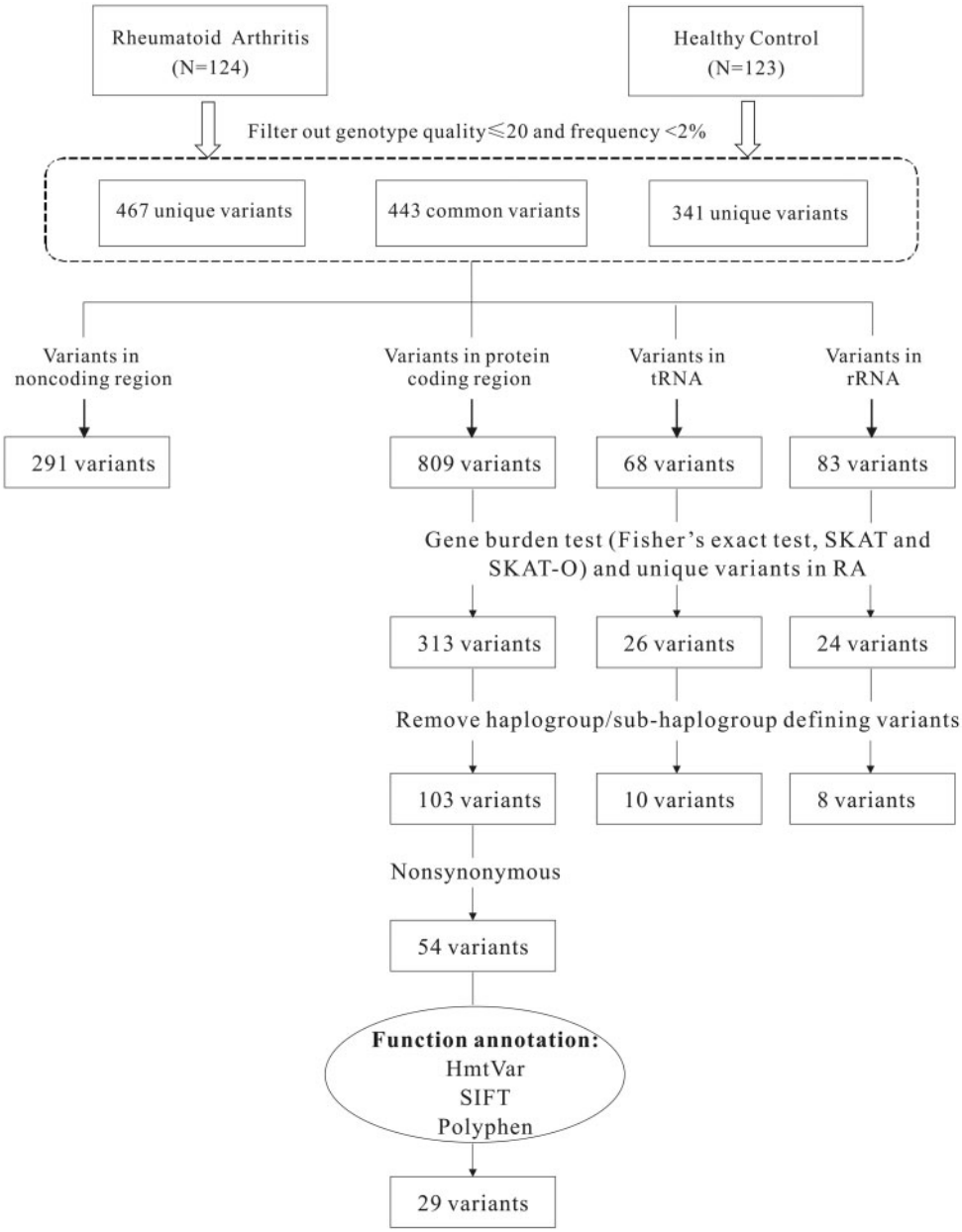
mtDNA variant detection and haplogroup assignment

Raw reads were filtered to remove low-quality reads by using Fastp. Clean reads were aligned to the revised Cambridge Reference Sequence (rCRS) with Burrows-Wheeler Aligner (BWA). Duplications were removed and sequence variants and small indels were called with the Genome Analysis Toolkit (GATK) Unified Genotyper and GATK Indel Genotyper V2. After a series of filter conditions, including total sequencing coverage ($\geq$100), quality score (each base quality $>$20) and the minimum allele frequency (MAF $\geq$2%), mtDNA variants were obtained for further analysis. The heteroplasmic fraction (HF) was defined as the percentage of variant reads among the total reads for the same nucleotide position.

To assess the associations between mtDNA variants and RA, we evaluated the potential functional impacts of these variants (Fig. 1). First, we performed a gene-burden test for protein-coding and tRNA variants based on the frequency of variants of each targeted gene in the RA or control group. Fisher's exact test was used to identify RA-associated genes. Cumulative effects of variants in the gene were applied using the sequence kernel association test (SKAT) and SKAT-optimal unified test (SKAT-O), which combined burden and kernel tests with small sample size adjustment [4, 19]. Bonferroni multiple comparison was conducted for the gene-based test ($P < 2.94 \times 10^{-3}$, 0.05/17 genes) and single variant test ($P < 4.00 \times 10^{-5}$, 0.05/1251 variants). Then, we focused on unique variants carried by RA patients. For variants in protein-coding genes, the synonymous and haplogroup/subhaplogroup defined variants were filtered out. A combined prediction system with HmtVar, SIFT and PolyPhen 2 was applied. Variants were considered potentially pathogenic when all three tools supported at least moderate evidence of pathogenicity. For variants located in tRNA, we determined the potential structure and functional alteration by using HmtVar, Mamit-tRNA, MitoTIP and PON-mt-tRNA. The global frequency data and evolutionary conservation index are available from the MITOMASTER database, which contains 48 882 full-length human mtDNA genomes and reference genomes for 45 species (Homo sapiens, 39 mammals and five non-mammalian species).

The haplogroup status was determined according to PhyloTree for global human mtDNA (www.phylotree.org) [20] and the phylogenetic tree for East Asian mtDNA and further checked by MitoTool (http://www.mitotool.org/) [21].

FIG. 1 The workflow for analysing mtDNA variants in rheumatoid arthritis ($n = 124$) and healthy controls ($n = 123$)



Statistical analysis

Statistical analyses were performed using SPSS 23.0 software (SPSS Inc., Chicago, IL, USA), and statistical significance was established at $P \leq 0.05$ or adjusted for independent tests (17 genes, 1251 variants or 34 haplogroups) using Bonferroni correction. Continuous or categorical variables were compared using Student's *t* test or Fisher's exact test. Binary logistic regression analysis was carried out to determine the contributions of each haplogroup by calculating odds ratios (OR) and 95% CI between the RA and HC groups.

Results

Study participants and mitochondrial genome sequencing

Supplementary Table S1, available at *Rheumatology* online depicted the study population characteristics, including 124 RA patients [female 73.4%, mean age 56.98 (11.14) years] and 123 age- and sex-matched HC [female 74.8%, mean age 56.24 (10.98) years]. Next-generation sequencing data across the entire mtDNA genome were generated for all subjects with a mean coverage depth of 1492 (665) in RA patients and 1203

(684) in HC. The sequence read coverage of the RA and HC was 100.00±0.00% and 99.99±0.02%, respectively.

Landscape of mtDNA variants in RA patients

A total of 910 mtDNA variants were identified in RA patients compared with 784 variants in HC. There were 443 common variants shared by RA patients and HC individuals (Supplementary Fig. S1 and Supplementary Table S2, available at *Rheumatology* online). There were 467 and 341 unique mtDNA variants detected only in the RA and HC groups, respectively (Supplementary Tables S3–4, available at *Rheumatology* online). The variant density (numbers of variants/region length) of the control region was much higher than that of the non-control region in both the RA and HC groups (data not shown). Except for the control region, the *MT-ND5* region carried the most variants (107 variants) in RA patients, followed by *MT-CO1* (72 variants) and *MT-CYB* (70 variants) (Supplementary Table S5, available at *Rheumatology* online). In HC, the *MT-ND5* region showed the most variants, at 87, followed by *MT-CYB* (72 variants) and *MT-CO1* (56 variants). The frequency of variants per mtDNA genome was defined by dividing the variant numbers by both region length and sample size. As shown in Fig. 2, higher variant density in *MT-ATP6* was observed in RA patients.

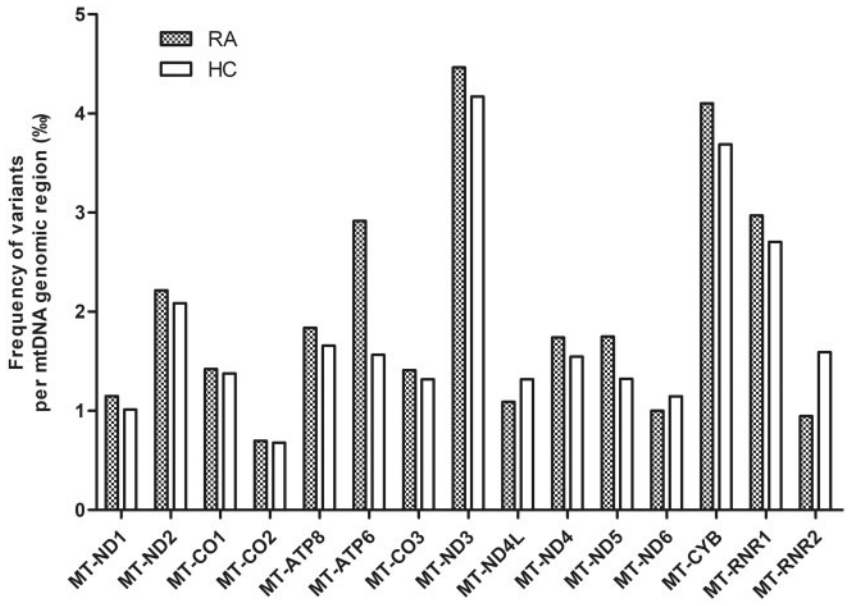
We then divided all mtDNA variants into three groups: common variants shared by RA and HC, RA-only variants and HC-only variants. Among the common variants group, a comparable number of patients carried variants in each gene region (Table 1). Most mtDNA genes displayed significant accumulation of RA-only variants or HC-only variants. In the RA-only variant group, *MT-ND5* was significantly enriched in RA patients (104/124, 83.9%, $P = 1.32 \times 10^{-49}$), followed by *MT-ATP6*, *MT-ND1* and *MT-CO1*. An abundance of HC-only variants accumulated in the *MT-CYB* gene (32/123, 26.0%, $P = 1.99 \times 10^{-11}$), followed by *MT-ND5*, *MT-ND4* and *MT-CO1*.

Potential pathogenic mtDNA variants associated with rheumatoid arthritis

We conducted gene-based analysis to clarify the relationship between mtDNA variants and RA. Both Fisher's exact and SKAT/SKAT-O tests revealed that more enrichment variants in *MT-ATP6* were present in RA patients than in the control group even after Bonferroni adjustment. The significance was not maintained for *MT-ND5* after multiple comparison correction (Table 2).

We next focused on unique variants in coding regions carried by RA patients and filtered out haplogroup or subhaplogroup defining variants based on the mtDNA mutation database, MitoMap (Fig. 1). Among the 103 remaining protein-coding genes, we selected 54 nonsynonymous variants for further analysis, because synonymous variants are likely to be functionally neutral. A scoring system combining with HmtVar, SIFT and PolyPhen was applied to evaluate pathogenicity. Twenty-nine variants showed moderate or strong evidence of disease causing and all those were rare/low frequency variants (Table 3), which was <1‰ in 48 882 human mitochondrial sequences from MITOMASTER. Notably, we included the variant m.8830C>A as RA-associated variant due to high frequency observation in

Fig. 2 Frequency distribution of average mtDNA variants per mitochondrial genome in RA patients ($n = 124$) and healthy controls (HC) ($n = 123$)



Average number of mtDNA variants per mtDNA genome was calculated by dividing the variant numbers by both region length and sample size in protein-coding gene.

Table 1 Number of subjects carried mtDNA variants in each gene region

	RA-only variants		HC-only variants		Common variants		
	n	*P*-value	*n*	*P*-value	*n*	*n*	*P*-value
MT-RNR1	11	7.27×10^{-4}	7	0.01	124	123	1.00
MT-RNR2	11	7.27×10^{-4}	21	1.72×10^{-7}	124	123	1.00
MT-ND1	36	1.01×10^{-10}	19	8.30×10^{-7}	68	65	0.80
MT-ND2	27	3.04×10^{-9}	18	1.81×10^{-6}	123	121	0.62
MT-CO1	35	3.54×10^{-12}	26	3.04×10^{-9}	124	123	1.00
MT-CO2	19	1.81×10^{-6}	15	1.82×10^{-5}	28	27	1.00
MT-ATP8	12	3.7×10^{-4}	3	0.12	22	33	0.09
MT-ATP6	56	8.30×10^{-21}	13	8.28×10^{-5}	83	82	1.00
MT-CO3	19	0.06	20	3.78×10^{-7}	77	71	0.52
MT-ND3	7	0.01	11	3.69×10^{-4}	100	95	0.54
MT-ND4L	4	0.12	9	0.00	26	29	0.65
MT-ND4	34	8.42×10^{-12}	26	3.04×10^{-9}	124	123	1.00
MT-ND5	104	1.32×10^{-49}	28	5.83×10^{-10}	113	111	0.83
MT-ND6	13	1.75×10^{-4}	20	3.78×10^{-7}	46	49	0.70
MT-CYB	32	4.67×10^{-11}	32	1.99×10^{-11}	124	123	1.00

RA-only variants, mtDNA variants only carried by RA patients. HC-only variants, mtDNA variants only carried by HC subjects. Common variants, mtDNA variants identified both in RA patients and HC group.

Table 2 Potential associated mitochondrial gene of RA observed by gene-based analyses

	RA	HC	Fisher's exact test *P*-value	SKAT *P*-value	SKAT-O *P*-value
MT-ATP6[a]	229	127	$\mathbf{1.18 \times 10^{-5}}$	$\mathbf{2.90 \times 10^{-5}}$	$\mathbf{2.45 \times 10^{-7}}$
MT-ATP8	47	42	1.00	0.61	0.69
MT-CO1	273	261	0.34	0.50	1.00
MT-CO2	59	57	0.64	0.19	0.81
MT-CO3	137	127	0.71	0.11	0.57
MT-CYB	580	517	0.87	0.39	1.00
MT-ND1	136	119	0.95	0.41	0.36
MT-ND2	286	267	0.50	0.06	1.00
MT-ND3	191	177	0.67	0.08	0.39
MT-ND4	297	262	1.00	0.16	1.00
MT-ND4L	40	41	0.50	0.04	0.86
MT-ND5	393	295	**0.03**	0.00	1.00
MT-ND6	65	74	0.14	0.34	0.39
MT-RNR1	356	317	0.94	0.37	1.00
MT-RNR2	297	305	0.05	0.56	1.00
MT-tRNA	69	62	0.93	0.54	0.75

Statistical significance set at a Bonferroni adjusted $P < 2.94 \times 10^{-3}$ (0.05/17 genes). [a]Only *MT-ATP6* remained significance after Bonferroni adjustment. Fisher's exact test was used for gene-burden test.

RA, although SIFT classified it as a benign variant. Twenty-five out of these variants (83.3%) were in a heteroplasmic state, with levels ranging from 2.0% to nearly 100%. But it should be noted that extremely low levels (<5%) were observed in over half of the heteroplasmic variants (16/25, 72.0%).

We then grouped these thirty variants into two groups according to their frequency in RA patients. Only four variants (m.8830C>A, m.8833G>C and m.8843T>A in *MT-ATP6* and m.14135T>G in *MT-ND5*) were frequently detected in RA patients (>5%), with frequencies of 29.03% (36/124), 24.19% (30/124), 8.06% (10/124) and 72.58% (90/124), respectively. The three variants, except m.8843T>A, were significantly associated with RA using SKAT-O test at a Bonferroni adjusted threshold ($P < 4.00 \times 10^{-5}$). Of these four variants, only m.8830C>A and m.8843T>A showed evolutionary conservation (80.00% and 86.67%); the conservation index of m.8833G>C and m.14135T>G was as low as 26.67% and 42.22%, respectively, among 45 vertebrates. Quite a few variants that failed to satisfy our filtering criteria were significantly associated with RA

Table 3 Potential associated mtDNA variants of rheumatoid arthritis obtained from next-generation sequencing followed by filtering[a]

mtDNA variants	Heteroplasmic rate	Case	Gene	Mutation type	Amino acid change	Frequency	Conservation Index (%)	HmtVar Prediction Score	SIFT[b]	PolyPhen[c]
m.5268A>G	84.96%	1	MT-ND2	Missense	I267V	0.02%	88.89	0.49	4	2
m.6355A>T	3.13%	1	MT-CO1	Missense	H151L	0.00%	97.78	0.76	3	3
m.6376T>C	3.13%	1	MT-CO1	Missense	I158T	0.00%	100	0.8	3	3
m.6396A>G	4.69%	1	MT-CO1	Missense	I165V	0.00%	100	0.66	3	2
m.6534A>G	2.17%	1	MT-CO1	Missense	T211A	0.00%	100	0.69	3	3
m.6654T>C	2.17%	1	MT-CO1	Missense	F251L	0.00%	100	0.8	3	3
m.6744G>T	2.08%	1	MT-CO1	Missense	G281C	0.00%	100	0.85	3	3
m.7216G>T	2.89%	1	MT-CO1	Missense	R438L	0.00%	100	0.87	3	3
m.7652T>C	99.87%	1	MT-CO2	Missense	F23L	0.00%	100	0.82	3	3
m.8531A>C	2.74%	2	MT-ATP8/MT-ATP6	Missense	ATPase8: T56PA/ATPase6: N2T	0.00%	93.33	0.82/0.63	4	3
m.8542T>C	96.56%	1	MT-ATP8/MT-ATP6	Missense	ATPase8: C59C/ATPase6: F6L	0.01%	6.67	–/0.73	4	3
m.8830C>A[d]	26.81%	36	MT-ATP6	Missense	L102M	0.00%	80	0.68	1	3
m.8833G>C	28.03%	30	MT-ATP6	Missense	A103P	0.00%	26.67	0.88	4	3
m.8843T>A	32.88%	10	MT-ATP6	Missense	I106N	0.00%	86.67	0.73	4	2
m.8935C>T	100.00%	1	MT-ATP6	Missense	L137F	0.01%	100	0.81	4	3
m.8978T>C	99.79%	1	MT-ATP6	Missense	I151T	0.03%	100	0.77	4	3
m.9874T>C	2.00%	1	MT-CO3	Missense	L223P	0.00%	77.78	0.81	3	2
m.9909T>C	99.59%	1	MT-CO3	Missense	F235L	0.02%	100	0.72	3	3
m.11018G>T	2.40%	1	MT-ND4	Nonsense	E87TERM	0.00%	88.89	–	–	–
m.11038A>C	3.13%	1	MT-ND4	Missense	K93N	0.00%	86.67	0.78	3	3
m.11101A>T	99.90%	1	MT-ND4	Missense	E114D	0.00%	93.33	0.81	3	3
m.11397A>G	2.08%	1	MT-ND4	Missense	H213R	0.00%	93.33	0.87	3	3
m.12760A>G	95.63%	1	MT-ND5	Missense	I142V	0.00%	93.33	0.63	3	3
m.13906A>C	2.29%	1	MT-ND5	Missense	N524H	0.00%	60	0.76	3	2
m.13910T>G	2.33%	1	MT-ND5	Nonsense	M525TERM	0.00%	35.56	–	–	–
m.13911A>T	2.14%	1	MT-ND5	Missense	M525I	0.00%	35.56	0.77	3	3
m.13937A>G	2.14%	1	MT-ND5	Missense	H534R	0.00%	95.56	0.81	3	3
m.14135T>G	24.32%	90	MT-ND5	Missense	L600R	0.00%	42.22	0.54	3	2
m.15209T>C	99.53%	1	MT-CYB	Missense	Y155H	0.01%	100	0.85	3	3
m.15767C>G	95.38%	1	MT-CYB	Missense	Q341E	0.00%	93.33	0.79	3	3

[a]The filtering strategy refers to Fig. 1 for more information. [b]PolyPhen prediction: 1 benign, 2 possibly damage, 3 probably damage. [c]SIFT prediction: 1 tolerated, 2 tolerated with low confidence, 3 deleterious with low confidence, 4 deleteriouse. [d]m.8830C>A was recorded in the table due to high frequency observed in RA patients, although low evidence predicted by SIFT.

according to SKAT-O before adjustment (Supplementary Table S6, available at *Rheumatology* online). The significance level only held up to m.14766C>T due to the difference in the heteroplasmic fraction between the groups. We also found that several variants in the control region were significant associated with RA, but no significance maintained after Bonferroni correction (Supplementary Table S6, available at *Rheumatology* online).

With regard to tRNA, we utilized HmtVar, Mamit-tRNA, MitoTIP and PON-mt-tRNA to assess the predicted pathogenicity of each variant. We identified a reported variant, m.15940delT, located in a 5T stretch and part of the TΨC loop of $tRNA^{Thr}$ (Supplementary Table S7, available at *Rheumatology* online); this variant has been associated with various phenotypes, such as hyperammonaemia, hypertrophic cardiomyopathy and dysuria [22]. HmtVar indicated m.7472A>C to be a disease-causing variant with mild evolutionarily conservation (62.22%) in the variable loop region; however, in a previous study, it was speculated that this variant has an inhibitory effect on MERRF-related pathogenic 7472insC [23]. Thus, it was unlikely to be a pathogenic mutation.

Association between potential pathogenic mtDNA variants and phenotypes in RA

As bilirubin has been supported as an effective antioxidant to scavenge ROS, we investigated the association between serum bilirubin levels and the four frequently identified potential pathogenic mtDNA variants. We observed high total/indirect bilirubin in patients carrying the m.8830C>A variant ($P = 0.006$ and $P = 0.002$), whereas no significance was detected for other variants (Supplementary Table S8, available at *Rheumatology* online).

Association analysis of mtDNA haplogroups with RA

Given the observation that the mtDNA haplogroup background itself confers protective or harmful effects on various diseases, we determined the mtDNA haplogroup for RA patients and healthy individuals based on the mtDNA tree in the PhyloTree and MitoTool. All subjects were categorized into the known East Asian haplogroups except one JT haplogroup found in HC individuals. The distribution of the common southern (e.g. B, F, M7, R9a'b and N9a) and northern (e.g. A, G, D, CZ, M8a and Y) haplogroups showed no significant difference between RA patients and HC individuals (data not shown).

As presented in Table 4, we found a lower frequency of subhaplogroup F1 ($P = 0.01$, OR = 0.31, 95% CI: 0.13, 0.76), particularly F1a'c and F1a ($P = 0.04$, OR = 0.37, 95% CI: 0.14, 0.98; $P = 0.01$, OR = 0.26, 95% CI: 0.08, 0.81, respectively), in the RA group than the HC group, indicating that these haplogroups decrease the risk of RA. No associations were detected for the other haplogroups, even for macro-haplogroups R9 and F, which included subhaplogroups F1, F1a'c and F1a. No haplogroups were significantly associated with RA after multiple comparison correction ($P > 1.47 \times 10^{-3}$). To assess the potential role of F1, F1a'c and F1a-defined mutations, we conducted χ^2 analysis and observed a significantly lower frequency of all mutations except m.G13759A in RA patients compared with the control group before Bonferroni correction (Supplementary Table S9, available at *Rheumatology* online). We further analysed the degree of evolutionary conservation for F1, F1a'c and F1a-defined mutations. Strikingly, only two synonymous mutations (m.G6962A and m.C12882T) were found to be highly conserved.

Discussion

In the present study, we described the landscape of germline and somatic mtDNA variants in RA patients and identified enrichment variants in *MT-ATP6* for RA patients using next-generation sequencing. In particular, three potential pathogenic variants (m.8833G>C, m.8843T>A and m.14135T>G) and the m.8830C>A variant were frequently observed in the RA group at low levels of heteroplasmy (ranging from 15.09% to 44.44%). Haplogroup F1 and its subhaplogroup F1a might be considered protective factors for RA in the Taizhou population, but the significance was not maintained after multiple comparisons.

Oxidative stress is believed to be involved in the pathogenesis of RA [24, 25]. Previous reports have demonstrated a strong association of enhanced ROS and reduced free radical scavenging activity in peripheral blood and synovial tissue with RA [26, 27]. Dysfunctional mitochondrial respiratory complexes are the major source of ROS generation leading to mtDNA mutations. Furthermore, the damaged mtDNA in RA might be recognized as nonself, stimulating the immune system by inducing the expression of major histocompatibility complex I and II (MHC-I and MHC-II) [16, 28]. Recently, Tseng CC *et al.* identified four potential pathogenic mtDNA genes (*MT-CO3*, *MT-TA*, *MT-TC* and *MT-TT*) in gout and provided a potential association between *MT-CO3* and levels of high density lipoprotein [29]. In our study, a higher incidence of mtDNA variants was found in peripheral blood samples from RA compared with samples from the control group, especially for *MT-ATP6*. More interestingly, a set of variants with similar locations in *MT-ATP6* were frequently detected in RA patients. These observations further confirmed the relationship between mtDNA alteration and RA development and raised the possibility that certain *MT-ATP6* mutations might confer RA susceptibility. *MT-ATP6* encodes a subunit of the membrane proton channel, a key component of F_1F_0 ATP synthase or complex V. Mitochondrial complex V has been considered as a key mediator in ATP synthesis, ROS production and apoptosis pathways, which were associated with RA progression documented in previous studies [30]. A series of clinical manifestations, such as neurogenic muscle

TABLE 4 Frequencies of mtDNA haplogroups in RA patients and healthy controls

Haplogroup	RA (124) Case (%)	HC (123) Case (%)	*P*-value	OR	95% CI
M	64 (51.61)	53 (43.09)	0.20	1.39	0.84, 2.29
D	24 (19.35)	25 (20.33)	0.87	0.95	0.51, 1.78
D4	14 (11.29)	21 (17.07)	0.20	0.62	0.30, 1.29
D4a	5 (4.03)	6 (4.88)	0.76	0.83	0.25, 2.78
D4b	4 (3.22)	7 (5.69)	0.36	0.56	0.16, 1.96
D5	10 (8.06)	3 (2.44)	0.05	3.50	0.95, 13.19
M*	40 (32.26)	28 (22.76)	0.13	1.56	0.88, 2.74
M7	11 (8.87)	8 (6.50)	0.49	1.40	0.54, 3.61
M7b	3 (2.42)	6 (4.88)	0.33	0.48	0.12, 1.98
M7c	8 (6.45)	2 (1.63)	0.10	4.17	0.87, 20.06
M8	14 (11.29)	10 (8.13)	0.40	1.44	0.61, 3.38
M8a	6 (4.84)	5 (4.07)	0.77	1.20	0.36, 4.04
CZ	8 (6.45)	5 (4.07)	0.40	1.63	0.52, 5.12
M9	0 (0.00)	1 (0.81)	1.000	0	—
M10	5 (4.03)	2 (1.63)	0.45	2.54	0.48, 13.36
M12'G	6 (4.88)	3 (2.44)	0.50	2.03	0.50, 8.32
G	6 (4.84)	3 (2.44)	0.50	2.03	0.50, 8.32
M#	9 (7.26)	6 (4.88)	0.43	1.53	0.53, 4.43
N	60 (48.39)	70 (56.91)	0.18	0.71	0.43, 1.17
R	44 (35.48)	53 (43.09)	0.22	0.73	0.44, 1.21
R9	20 (16.13)	26 (21.14)	0.31	0.72	0.38, 1.37
R9b	1 (0.81)	1 (0.81)	1.00	0.99	0.06, 16.04
F	19 (15.32)	25 (20.33)	0.30	0.71	0.37, 1.37
F1	7 (5.65)	20 (16.26)	0.01	0.31	0.13, 0.76
F1a'c	6 (4.84)	15 (12.20)	0.04	0.37	0.14, 0.98
F1a	4 (3.23)	14 (11.38)	0.01	0.26	0.08, 0.81
F2	8 (6.45)	2 (1.63)	0.10	4.17	0.87, 20.06
F#	4 (3.22)	3 (2.44)	1.00	1.33	0.29, 6.09
B	22 (17.74)	26 (21.14)	0.50	0.81	0.43, 1.51
B4	14 (11.29)	19 (15.45)	0.34	0.70	0.33, 1.46
B5	8 (6.45)	7 (5.70)	0.80	1.14	0.40, 3.26
A	10 (8.06)	9 (7.32)	0.83	1.11	0.44, 2.84
N9	5 (4.03)	8 (6.50)	0.38	0.60	0.19, 1.90
N#	1 (0.81)	1 (0.81)	1.00	0.99	0.06, 16.04

Statistical significance set at a Bonferroni adjusted $P < 1.47 \times 10^{-3}$ (0.05/34 haplogroups). No significance was found after multiple comparison correction.

weakness, peripheral neuropathy, optic atrophy and metabolic syndrome, have been attributed to pathogenic mutations in *MT-ATP6* differing by heteroplasmic load [31–33].

Curiously, up to 90 blood samples from RA patients harboured low levels of heteroplasmic m.14135T>G mutation in *MT-ND5*, a subunit of mitochondrial complex I. Consistent with our findings, Mitsunaga and colleagues suggested that aggregation of the rare/low-frequency variants in *NDUFA7*, another subunit of complex I, was associated with severe erosive arthritis and played an important role in the pathogenesis of RA [4]. According to Da Sylva [16], synovial cells adopt mtDNA mutations as a means of survival, which results in hyperplasia due to high levels of ROS. We speculate that this mutation might be accumulated in synovial tissue, causing a low mutation burden in blood. Unfortunately, we could not perform analysis on synovial tissues due to difficulties in obtaining samples. Moreover, as moderate evidence of pathogenicity and poor conservation index were predicted, the effect of this variant remains to be further fully elucidated.

Previous studies have reported that haplogroup B/B4 and J, both belonging to macrohaplogroup R, were associated with low susceptibility to OA in southern China and Spain, respectively. Inconsistent with our results, subhaplogroup B/B4 showed a lower frequency in RA patients, but no significant difference in the southern Chinese population. Instead, haplogroup F1/F1a'c, another major subhaplogroup of macro haplogroup R, was considered a protective factor for RA. No such protective effect was observed in Chinese OA patients after adjustment for age and gender, though a decreasing tendency was noted. This distribution difference might be partially due to the small sample size or different character of RA and OA. Although the exact biological function of haplogroup F remained unclear, recent studies revealed that haplogroup F (especially for

subhaplogroup F1) exhibited lower ROS levels and mitochondrial complex I activities than other haplogroups [34–36]. Fang *et al.* clarified that the underlying mechanism by which haplogroup G conferred a risk of OA was by shifting the metabolic pattern from glycolysis to OXPHOS but that the protective haplogroup B4 preferred ATP generation from glycolysis [37]. Moreover, gene expression profiles based on trans-mitochondrial technology have identified haplogroup G-enriched OA-related pathways, such as the rheumatoid arthritis, glycolysis/gluconeogenesis and NF-κB signalling pathways, compared with haplogroup B4 using. Evidence has shown that exogenous administration of an intermediate of glycolysis, fructose 1,6-biphosphate, could effectively reduce the disease activity in a zymosan-induced arthritis model by promoting anti-inflammatory pathways [38]. These observations support our finding that haplogroup F1 and its subhaplogroup F1a were protective factors for RA, as it displayed similar features as haplogroup B4 with high glycolysis activity. However, the exact role of haplogroup F1 in RA requires further validation and functional investigation.

There are some limitations to our study. First, we did not analyse the mtDNA mutation spectrum of synovial tissues in RA due to difficulties in sample collection. In addition, we did not validate our results in large samples and frequent heteroplasmic variants with other methods, such as qPCR or pyrosequencing assays. Finally, the sample size was relatively small, particularly for the analysis of the mtDNA haplogroup background. We are currently collecting more samples to confirm our findings.

In conclusion, we conducted high-throughput sequencing to investigate the comprehensive landscape of mtDNA variants and the influence of the mtDNA background on RA in the Taizhou area, China. Four RA-specific somatic variants that drove the enrichment variants in *MT-ATP6* and *MT-ND5* were frequently observed. In addition, the protective effect of haplogroup F1/F1a could not hold up in multiple comparison correction in our present study. The association should be evaluated in a large sample size, and targeted functional studies of *MT-ATP6/MT-ND5* variants and haplogroup F1/F1a will provide new insights into the role of mtDNA mutations in RA.

Acknowledgements

We thank all the colleagues at the department of Clinical Laborary in Taizhou Hospital of Zhejiang Province for assistance in sample and clinical data collection. We thank Haixia He for her help in organizing the data. We also thank Shuanglong Chen for his assistant in the analyses of SKAT and SKAT-O.

Funding: This work was supported by grants from the National Natural Science Foundation of China (81672086), Science and Technology Department of Zhejiang Province (LQ19H100001), as well as the Zhejiang Medical and Health Science and Technology Plan (2017176310 and 2018KY913).

Disclosure statement: The authors have declared no conflicts of interest.

Supplementary data

Supplementary data are available at *Rheumatology* online.

References

1 Phillips DC, Dias HK, Kitas GD, Griffiths HR. Aberrant reactive oxygen and nitrogen species generation in rheumatoid arthritis (RA): causes and consequences for immune function, cell survival, and therapeutic intervention. Antioxidants Redox Signaling 2010;12:743–85.

2 Bohuslav J, Chen LF, Kwon H, Mu Y, Greene WC. p53 induces NF-kappaB activation by an IkappaB kinase-independent mechanism involving phosphorylation of p65 by ribosomal S6 kinase 1. J Biol Chem 2004;279:26115–25.

3 Chen L, Xu B, Liu L *et al*. Cadmium induction of reactive oxygen species activates the mTOR pathway, leading to neuronal cell death. Free Radic Biol Med 2011;50:624–32.

4 Mitsunaga S, Hosomichi K, Okudaira Y *et al*. Aggregation of rare/low-frequency variants of the mitochondria respiratory chain-related proteins in rheumatoid arthritis patients. J Hum Genet 2015;60:449–54.

5 Bobba A, Amadoro G, Valenti D *et al*. Mitochondrial respiratory chain Complexes I and IV are impaired by beta-amyloid via direct interaction and through Complex I-dependent ROS production, respectively. Mitochondrion 2013;13:298–311.

6 Drechsel DA, Patel M. Differential contribution of the mitochondrial respiratory chain complexes to reactive oxygen species production by redox cycling agents implicated in Parkinsonism. Toxicol Sci 2009;112:427–34.

7 Drose S, Bleier L, Brandt U. A common mechanism links differently acting complex II inhibitors to cardioprotection: modulation of mitochondrial reactive oxygen species production. Mol Pharmacol 2011;79:814–22.

8 Siebels I, Drose S. Q-site inhibitor induced ROS production of mitochondrial complex II is attenuated by TCA cycle dicarboxylates. Biochim Biophys Acta 2013;1827:1156–64.

9 Diaz-Morales N, Lopez-Domenech S, Iannantuoni F *et al*. Mitochondrial DNA haplogroup JT is related to impaired glycaemic control and renal function in type 2 diabetic patients. J Clin Med 2018;7:220.

10 Fang H, Hu N, Zhao Q *et al*. mtDNA haplogroup N9a increases the risk of type 2 diabetes by altering mitochondrial function and intracellular mitochondrial signals. Diabetes 2018;67:1441–53.

11 Feng J, Zhang J, Liu M *et al*. Association of mtDNA haplogroup F with healthy longevity in the female

Chuang population, China. Exp Gerontol 2011;46: 987–93.

12 Koo BS, Song Y, Lee S *et al*. Association of Asian mitochondrial DNA haplogroup B with new development of knee osteoarthritis in Koreans. Int J Rheum Dis 2019; 22:411–6.

13 Yu X, Wieczorek S, Franke A *et al*. Association of UCP2 -866 G/A polymorphism with chronic inflammatory diseases. Genes Immun 2009;10:601–5.

14 Duhn PH, Sode J, Hagen CM, Christiansen M, Locht H. Mitochondrial haplogroups in patients with rheumatoid arthritis: no association with disease and disease manifestations. PLoS One 2017;12:e0188492.

15 Biniecka M, Fox E, Gao W *et al*. Hypoxia induces mitochondrial mutagenesis and dysfunction in inflammatory arthritis. Arthritis Rheum 2011;63:2172–82.

16 Da Sylva TR, Connor A, Mburu Y, Keystone E, Wu GE. Somatic mutations in the mitochondria of rheumatoid arthritis synoviocytes. Arthritis Res Ther 2005;7:R844–51.

17 Hajizadeh S, DeGroot J, TeKoppele JM, Tarkowski A, Collins LV. Extracellular mitochondrial DNA and oxidatively damaged DNA in synovial fluid of patients with rheumatoid arthritis. Arthritis Res Ther 2003;5: R234–40.

18 Yu X, Wester-Rosenlof L, Gimsa U *et al*. The mtDNA nt7778 G/T polymorphism affects autoimmune diseases and reproductive performance in the mouse. Hum Mol Genet 2009;18:4689–98.

19 Lee S, Abecasis GR, Boehnke M, Lin X. Rare-variant association analysis: study designs and statistical tests. Am J Hum Genet 2014;95:5–23.

20 van Oven M, Kayser M. Updated comprehensive phylogenetic tree of global human mitochondrial DNA variation. Hum Mutat 2009;30:E386–94.

21 Fan L, Yao YG. MitoTool: a web server for the analysis and retrieval of human mitochondrial DNA sequence variations. Mitochondrion 2011;11:351–6.

22 Seneca S, Lissens W, Liebaers I *et al*. Pitfalls in the diagnosis of mtDNA mutations. J Med Genet 1998;35: 963–4.

23 Pulkes T, Liolitsa D, Eunson LH *et al*. New phenotypic diversity associated with the mitochondrial tRNA(SerUCN) gene mutation. Neuromuscul Disord 2005;15:364–71.

24 Balogh E, Veale DJ, McGarry T *et al*. Oxidative stress impairs energy metabolism in primary cells and synovial tissue of patients with rheumatoid arthritis. Arthritis Res Ther 2018;20:95.

25 Li Y, Goronzy JJ, Weyand CM. DNA damage, metabolism and aging in pro-inflammatory T cells: rheumatoid arthritis as a model system. Exp Gerontol 2018;105:118–27.

26 Veselinovic M, Barudzic N, Vuletic M *et al*. Oxidative stress in rheumatoid arthritis patients: relationship to diseases activity. Mol Cell Biochem 2014;391:225–32.

27 Pradhan A, Bagchi A, De S *et al*. Role of redox imbalance and cytokines in mediating oxidative damage and disease progression of patients with rheumatoid arthritis. Free Radic Res 2019;53:768–79.

28 Gu Y, Wang C, Roifman CM, Cohen A. Role of MHC class I in immune surveillance of mitochondrial DNA integrity. J Immunol 2003;170:3603–7.

29 Tseng CC, Chen CJ, Yen JH, Huang HY *et al*. Next-generation sequencing profiling of mitochondrial genomes in gout. Arthritis Res Ther 2018;20:137.

30 Natera-Naranjo O, Kar AN, Aschrafi A *et al*. Local translation of ATP synthase subunit 9 mRNA alters ATP levels and the production of ROS in the axon. Mol Cell Neurosci 2012;49:263–70.

31 Ye W, Chen S, Jin S, Lu J. A novel heteroplasmic mitochondrial DNA mutation, A8890G, in a patient with juvenileonset metabolic syndrome: a case report. Mol Med Rep 2013;8:1060–6.

32 Larson AA, Balasubramaniam S, Christodoulou J *et al*. Biochemical signatures mimicking multiple carboxylase deficiency in children with mutations in MT-ATP6. Mitochondrion 2019;44:58–64.

33 Felhi R, Mkaouar-Rebai E, Sfaihi-Ben Mansour L *et al*. Mutational analysis in patients with neuromuscular disorders: detection of mitochondrial deletion and double mutations in the MT-ATP6 gene. Biochem Biophys Res Commun 2016;473:61–6.

34 Zhou H, Nie K, Qiu R *et al*. Generation and bioenergetic profiles of cybrids with East Asian mtDNA haplogroups. Oxid Med Cell Long 2017;2017:1062314.

35 Chalkia D, Chang YC, Derbeneva O *et al*. Mitochondrial DNA associations with East Asian metabolic syndrome. Biochim Biophys Acta Bioenerg 2018;1859:878–92.

36 Hwang IW, Kim K, Choi EJ, Jin HJ. Association of mitochondrial haplogroup F with physical performance in Korean population. Genomics Inform 2019;17:e11.

37 Fang H, Zhang F, Li F *et al*. Mitochondrial DNA haplogroups modify the risk of osteoarthritis by altering mitochondrial function and intracellular mitochondrial signals. Biochim Biophys Acta 2016;1862:829–36.

38 Veras FP, Peres RS, Saraiva AL *et al*. Fructose 1,6-bisphosphate, a high-energy intermediate of glycolysis, attenuates experimental arthritis by activating anti-inflammatory adenosinergic pathway. Sci Rep 2015;5: 15171.