

Mitochondrial DNA haplogroup M7 confers a reduced risk of colorectal cancer in a Han population from northern China

Qing Yuan¹ | Liping Su² | Tian Wang³ | Yang Liu² | Zhenxing Lu¹ | Kaixiang Zhou² | Shanshan Guo² | Xiwen Gu⁴ | Jinliang Xing² | Xu Guo² 

¹Institute of Medical Research, Northwestern Polytechnical University, Xi'an, China

²State Key Laboratory of Cancer Biology and Department of Physiology and Pathophysiology, Fourth Military Medical University, Xi'an, China

³Laboratory Department, Shaanxi Provincial People's Hospital, Xi'an, China

⁴Key Laboratory of Shaanxi Province for Craniofacial Precision Medicine Research, Clinical Research Center of Shaanxi Province for Dental and Maxillofacial Diseases, College of Stomatology, Xi'an Jiaotong University, Xi'an, China

Correspondence

Xu Guo, Jinliang Xing, State Key Laboratory of Cancer Biology and Department of Physiology and Pathophysiology, Fourth Military Medical University, 169 Changle West Road, Xi'an 710032, China.
Emails: fmmuxguo@163.com; xingjl@fmmu.edu.cn

Funding information

This work was supported by the National Natural Science Foundation of China [grant 81602463 and 81830070].

Abstract

Mitochondria are central eukaryotic organelles in cellular metabolism and ATP production. Mitochondrial DNA (mtDNA) alterations have been implicated in the development of colorectal cancer (CRC). However, there are few reports on the association between mtDNA haplogroups or single nucleotide polymorphisms (SNPs) and the risk of CRC. The mtDNA of 286 Northern Han Chinese CRC patients were sequenced by next-generation sequencing technology. MtDNA data from 811 Han Chinese population controls were collected from two public data sets. Then, logistic regression analysis was used to determine the effect of mtDNA haplogroup or SNP on the risk of CRC. We found that patients with haplogroup M7 exhibited a reduced risk of CRC when compared to patients with other haplogroups (odds ratio [OR] = 0.532, 95% confidence interval [CI] = 0.285–0.937, $p = 0.036$) or haplogroup B (OR = 0.477, 95% CI = 0.238–0.916, $p = 0.030$). Furthermore, haplogroup M7 was still associated with the risk of CRC when the validation and combined control cohort were used. In addition, several haplogroup M7 specific SNPs, including 199T>C, 4071C>T and 6455C>T, were significantly associated with the risk of CRC. Our results indicate the risk potential of mtDNA haplogroup M7 and SNPs in CRC in Northern China.

KEYWORDS

colorectal cancer risk, haplogroup, mitochondria, single nucleotide polymorphisms

1 | INTRODUCTION

Colorectal cancer (CRC) is one of the most common cancers in both men and women, accounting for nearly 10% of the global cancer incidence.¹ Approximately 274,800 new CRC cases and 132,100 CRC-related deaths have been estimated to occur in China each year, accounting for nearly one-tenth of the global CRC burden.² Assessing the risk of CRC and then screening is the most powerful public health tool to reduce mortality or incidence.^{3,4} The major risk

factors that may influence the development of the CRC include age, male gender, obesity, diet high in fat as well as the medical history of inflammatory bowel disease (IBD), diabetes mellitus and so on.⁵ Several prediction models have been developed to quantify CRC risk based on clinical or laboratory data. However, all these models still have obvious limitations, such as the restricted age range of subjects, selection bias of subjects and incomplete assessment of risk factors in these studies.^{5–7} Meanwhile, the diagnosis of CRC mainly depends on imaging techniques, which are less efficient for the early

Qing Yuan, Liping Su, and Tian Wang contributed equally to this work as the first authors.

This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

© 2021 The Authors. *Journal of Cellular and Molecular Medicine* published by Foundation for Cellular and Molecular Medicine and John Wiley & Sons Ltd.

detection of CRC. Accurate early detection or risk prediction contributes to improved CRC survival. Unfortunately, to date, there are no favourable biomarkers for the risk prediction of CRC.

Intensive efforts have been made to understand the genetic risk factors of CRC. To date, over 40 nuclear genome variants associated with CRC risk have been identified, including SNP rs10911251, rs1321311, rs1035209 and so on.^{8–10} However, these susceptibility loci account for only about 8%–16% of CRC cases, suggesting that additional genetic risk factors of CRC may remain to be explored.⁸

Mitochondria are central eukaryotic organelles in cellular metabolism and ATP production. Notably, multiple metabolic deregulations have been linked to the pathogenesis of CRC, including amino acid metabolism, glucose metabolism and lipid metabolism.¹¹ Somatic mtDNA mutations and copy number alterations have also been frequently observed in CRC samples.¹² However, the potential involvement of germline mtDNA variations in CRC development is less known.

Germline mtDNA variations are often characterized by assigning haplogroups, which can be defined by a certain set of mtDNA variants and reflect specific ancestral populations and geographic origins. At present, it has been proved that mtDNA haplogroups are associated with the risk of various cancers.^{13–16} Grzybowski et al. found that haplogroup R and its diagnostic mutations at positions 12705 and 16223 are associated with higher frequencies in Polish CRC patients when compared to healthy individuals.¹⁷ Certain mtDNA SNPs have also been associated with an increased risk of CRC in Iranian or European Americans.^{18–22} However, mtDNA haplogroups are highly variable among ethnic populations originating from different geographical locations. It is still unclear whether mtDNA haplogroups or SNPs are an important risk factor for colorectal carcinogenesis in the Han Chinese population.

To assess the contribution of mtDNA haplogroups and SNPs to the prevalence of CRC, we conducted a case-control study in a Han Chinese population.

2 | MATERIALS AND METHODS

2.1 | Patient specimens and DNA extraction

A total of 286 CRC patients were enrolled from the Tangdu Hospital affiliated with the Fourth Military Medical University (FMMU) (China) from January 2008 to December 2014. The diagnosis of CRC was confirmed by histological pathology. Patients' demographic and clinical data, including age, gender, tumour position, degree of differentiation, tumour node metastasis (TNM) stage and carcinoembryonic antigen (CEA), were collected and analysed. Tumour and paired non-tumour tissue specimens were collected from each patient during the surgical operation. Haematoxylin-eosin-stained slides of CRC tumour tissues were carefully reviewed to confirm that the cancer cell content was more than 90% in tumour tissues. Then, Genomic DNA was extracted from tissues of CRC patients who were enrolled in our previous study,²³ according to the E.N.Z.A. Tissue

DNA Kit (Omega Bio-Tek, Inc., Doraville, GA) under the manufacturer's instructions. DNA quantity and quality were analysed by Qubit (Invitrogen) and the 2100 Bioanalyzer system (Agilent Technologies), respectively. The study was approved by the Ethics Committees of the FMMU and written informed consent was obtained from each patient.

2.2 | Capture-based mtDNA next-generation sequencing

To obtain the full spectrum of germline mtDNA variations in CRC patients, we performed capture-based mtDNA next-generation sequencing as previously described.²⁴ Briefly, genomic DNA (1 µg for each sample) was randomly sonicated by Sonicator (Scientz98) to obtain fragments mainly distributed at 300–500 bp. The sonicated DNA fragments were end-repaired, ligated with sequencing adapters and amplified to generate the whole genome sequencing (WGS) library. Then, the WGS libraries of 20 samples were mixed with homemade biotinylated mtDNA capture probes for hybridization. After PCR amplification and purification, the mtDNA capture quality was determined using agarose gel electrophoresis and real-time fluorescence quantification. Finally, the captured mtDNA libraries were sequenced on an Illumina X Ten platform using paired-end runs with 2 × 150 cycles (PE 150).

2.3 | mtDNA haplogroup and mtSNPs

The FASTQ preprocessor fastp software (version 0.20.0)²⁵ was used to perform quality control and adapter trimming on the raw sequencing data. BWA software (version 0.7.17-r1188) was used to map the trimmed reads. To minimize the contamination of nuclear mtDNA segments, trimmed reads were mapped to the Revised Cambridge Reference Sequence (rCRS) of mtDNA and the reference genome (hg19). Next, Picard tools (version 1.81) were used to mark and remove duplicate reads. To reduce the false-positive rate of nearby indel positions, local realignment was performed using IndelRealigner in GATK software (version 3.2-2). Then, mtDNA sequences were extracted into FASTA format with the Perl script written in our laboratory. The obtained FASTA sequences were analysed using MitoTool (www.mitotool.org)²⁶ to determine mtDNA haplogroups and SNPs. Macro-haplogroups and micro-haplogroups were annotated based on the PhyloTree (www.phyloree.org).²⁷ The variation which was observed in both tumour and paired non-tumour tissues was defined as mtDNA SNP. SNPs with minor allele frequency (MAF) <5% in cases and controls were excluded in further analysis.

2.4 | Control cohorts

Two independent Han Chinese cohorts were used to evaluate the risk of CRC. First, we collected information on mtDNA haplogroups and

SNPs from a published data set, which included 562 normal individuals in Shaanxi Province of Northern China.²⁸ To further validate the results, we also collected another control cohort from the 1000 Genome Project,²⁹ including 249 Chinese (Southern Han Chinese, Beijing Han Chinese from Northern China and Denver Han Chinese from Colorado).

2.5 | Statistical analysis

In a case-control study, haplogroups with frequency >1% in both the controls and CRC patients were analysed to evaluate the effect of common mtDNA haplogroup on CRC. To estimate the relative risk, odds ratios (ORs) and 95% confidence intervals (CIs) were calculated for each haplogroup and SNP. Logistic regression analysis was used to assess the associations between haplogroups or SNPs and the risk of CRC. *p*-values of less than 0.05 were considered statistically significant. When other haplogroups were used as the reference haplogroup, it refers to the comparison of a specific haplogroup with all other haplogroups. Statistical analyses were performed using R software (version 4.0.2).

3 | RESULTS

3.1 | Patients characteristics

The clinical characteristics of the 286 CRC patients are summarized in Table 1. The median age of patients was 61 years old (ranging from 28–87 years) and 149 (52.10%) patients were male. Most CRC patients (93.71%) had CRC differentiation grade I–II, 164 (57.34%) patients were diagnosed with TNM stage I–II and 172 (60.14%) patients had serum CEA <5 ng/ml. The prevalence of major haplogroups is also listed in these different clinicopathological categories (Table S1).

3.2 | Association between mtDNA haplogroups and colorectal cancer risk in northern Han population

To investigate the association between genetic mtDNA variation and CRC risk, mtDNA haplogroups were annotated in 286 CRC cases and 562 Northern Han Chinese healthy controls (cohort 1). As shown in Table 2, patients were categorized into ten major haplogroups. Among them, haplogroup D was the most prevalent clade in both CRC patients and healthy controls (51 cases [17.83%] and 126 controls [22.42%]), followed by haplogroup B (51 cases [17.83%] and 86 controls [15.30%]).

When compared with other haplogroups, haplogroup M7 had a much lower percentage in CRC patients (5.24%, *n* = 15) than in control cohort 1 (9.43%, *n* = 53), which corresponded to a significantly reduced risk of CRC (OR 0.532, [95% CI 0.285–0.937], *p* = 0.036) (Table 2). The reduced CRC risk in haplogroup M7 was also evident when haplogroup B was selected as the reference haplogroup (OR 0.477, [95% CI 0.238–0.916], *p* = 0.030) (Table 3). Although

TABLE 1 Clinical-pathological characteristics of 286 Han Chinese CRC patients

Characteristics	CRC, <i>n</i> (%) <i>n</i> = 286
Age (years)	
≤61	141 (49.30%)
>61	145 (50.70%)
Gender	
Male	149 (52.10%)
Female	137 (47.90%)
Position	
Colon	194 (67.83%)
Rectum	92 (32.17%)
Differentiation	
I–II	268 (93.71%)
III	18 (6.29%)
TNM stage	
I–II	164 (57.34%)
III–IV	122 (42.66%)
Serum CEA (ng/ml)	
<5	172 (60.14%)
≥5	109 (38.11%)
Unknown	5 (1.75%)

Abbreviations: CEA, Carcinoembryonic antigen; CRC, Colorectal cancer; TNM, Tumour node metastasis.

no significant association was found when haplogroup D was used as the reference haplogroup, the selection of haplogroup D as a reference may be inappropriate due to the apparent imbalance of haplogroup D between the CRC and control cohort (Table 2).

3.3 | Association between haplogroup M7 and CRC risk with validation cohort

To further validate the association of haplogroup M7 with the risk of CRC, we used another Han Chinese data set from the 1000 genome project, which consisted of 249 individuals (Southern Han, 55; Northern Han, 121; and Denver Han, 73). As shown in Table 4, patients with haplogroup M7 were still associated with a significantly reduced risk of CRC when other haplogroups (OR 0.429, [95% CI 0.217–0.808], *p* = 0.011) and haplogroup B (OR 0.431, [95% CI 0.200–0.901], *p* = 0.028) were selected as the reference haplogroup. The association between haplogroup M7 and reduced CRC risk was also supported by a marginal significance level when haplogroup D was used as reference (OR 0.504, [95% CI 0.236–1.046], *p* = 0.070).

Based on the combined control cohort (cohort 1&2, *n* = 811), we also found that haplogroup M7 was a protective factor for the onset of CRC when referring to other haplogroups (OR 0.499, [95% CI 0.272–0.855], *p* = 0.017) and haplogroup B (OR 0.461, [95% CI 0.237–0.857], *p* = 0.018).

TABLE 2 Association between mtDNA haplogroups and CRC risk with other haplogroups as reference group

Haplogroup	Case (n = 286)	Control (n = 562)	OR (95%CI)	p-value
A	27 (9.44%)	56 (9.96%)	0.942 (0.574–1.513)	0.808
B	51 (17.83%)	86 (15.30%)	1.201 (0.818–1.751)	0.344
D	51 (17.83%)	126 (22.42%)	0.751 (0.520–1.073)	0.121
G	13 (4.55%)	30 (5.34%)	0.844 (0.420–1.611)	0.619
M7	15 (5.24%)	53 (9.43%)	0.532 (0.285–0.937)	0.036
M8	34 (11.89%)	59 (10.50%)	1.150 (0.729–1.791)	0.541
M9	7 (2.45%)	8 (1.42%)	1.737 (0.604–4.888)	0.291
M10	10 (3.50%)	10 (1.78%)	2.000 (0.812–4.929)	0.126
N9	22 (7.69%)	33 (5.87%)	1.336 (0.754–2.323)	0.310
R9	42 (14.69%)	81 (14.41%)	1.022 (0.678–1.522)	0.915
Others	14 (4.90%)	20 (3.56%)	1.395 (0.680–2.785)	0.350

Bold entries indicate statistical significance.

TABLE 3 Association between mtDNA haplogroups and CRC risk with haplogroup B and D as reference group

Haplogroup	Case (n = 286)	Control (n = 562)	OR(95%CI)	p-value	OR(95%CI)	p-value
B	51 (17.83%)	86 (15.30%)	Ref ^a		1.465 (0.911–2.360)	0.115
A	27 (9.44%)	56 (9.96%)	0.813 (0.454–1.439)	0.481	1.191 (0.674–2.082)	0.542
D	51 (17.83%)	126 (22.42%)	0.683 (0.424–1.098)	0.115	Ref ^b	
G	13 (4.55%)	30 (5.34%)	0.731 (0.341–1.503)	0.404	1.071 (0.504–2.180)	0.854
M7	15 (5.24%)	53 (9.43%)	0.477 (0.238–0.916)	0.030	0.699 (0.353–1.328)	0.287
M8	34 (11.89%)	59 (10.50%)	0.972 (0.561–1.675)	0.918	1.424 (0.833–2.423)	0.194
M9	7 (2.45%)	8 (1.42%)	1.475 (0.491–4.348)	0.477	2.162 (0.724–6.330)	0.156
M10	10 (3.50%)	10 (1.78%)	1.686 (0.650–4.381)	0.277	2.471 (0.959–6.373)	0.058
N9	22 (7.69%)	33 (5.87%)	1.124 (0.588–2.128)	0.720	1.647 (0.871–3.084)	0.121
R9	42 (14.69%)	81 (14.41%)	0.874 (0.524–1.453)	0.605	1.281 (0.780–2.101)	0.326
Others	14 (4.90%)	20 (3.56%)	1.180 (0.541–2.526)	0.671	1.729 (0.799–3.667)	0.156

^aHaplogroup B was used as reference group.

^bHaplogroup D was used as reference group.

Bold entries indicate statistical significance.

TABLE 4 Validation of association between mtDNA haplogroups and CRC risk

Haplogroup	Case(n = 286)	Validation control cohort (n = 249)	OR(95%CI)	p-value	Combined control cohort (n = 811)	OR(95%CI)	p-value
M7	15 (5.24%)	28 (11.24%)	0.427 (0.217–0.808)	0.011^a	81 (9.99%)	0.499 (0.272–0.855)	0.017^a
B	51 (17.83%)	41 (16.47%)	Ref	0.028^b	127 (15.66%)	Ref	0.018^b
M7	15 (5.24%)	28 (11.24%)	0.431 (0.200–0.901)		81 (9.99%)	0.461 (0.237–0.857)	
D	51 (17.83%)	48 (19.28%)	Ref	0.070 ^a	174 (21.45%)	Ref	0.155 ^a
M7	15 (5.24%)	28 (11.24%)	0.504 (0.236–1.046)		81 (9.99%)	0.632 (0.326–1.166)	

^aOther haplogroups were used as reference group.

^bHaplogroup B was used as reference group.

^cHaplogroup D was used as reference group.

Bold entries indicate statistical significance.

3.4 | The association between mtDNA SNPs and CRC risk

To clarify the association between mtDNA SNPs and CRC risk, we screened for common SNPs with allele frequencies higher than 5% in our cohorts. Using this criterion, 53 SNPs were identified in control

cohort 1 (Table S2), 72 SNPs were identified in control cohort 2 (Table S3) and 58 SNPs were identified in the combined control cohort (Table S4).

As shown in Figure 1, three haplogroup M7-specific SNPs, including 199T>C, 4071C>T, and 6455C>T, exhibited a stable association with the reduced risk of CRC in both cohorts. Although SNP

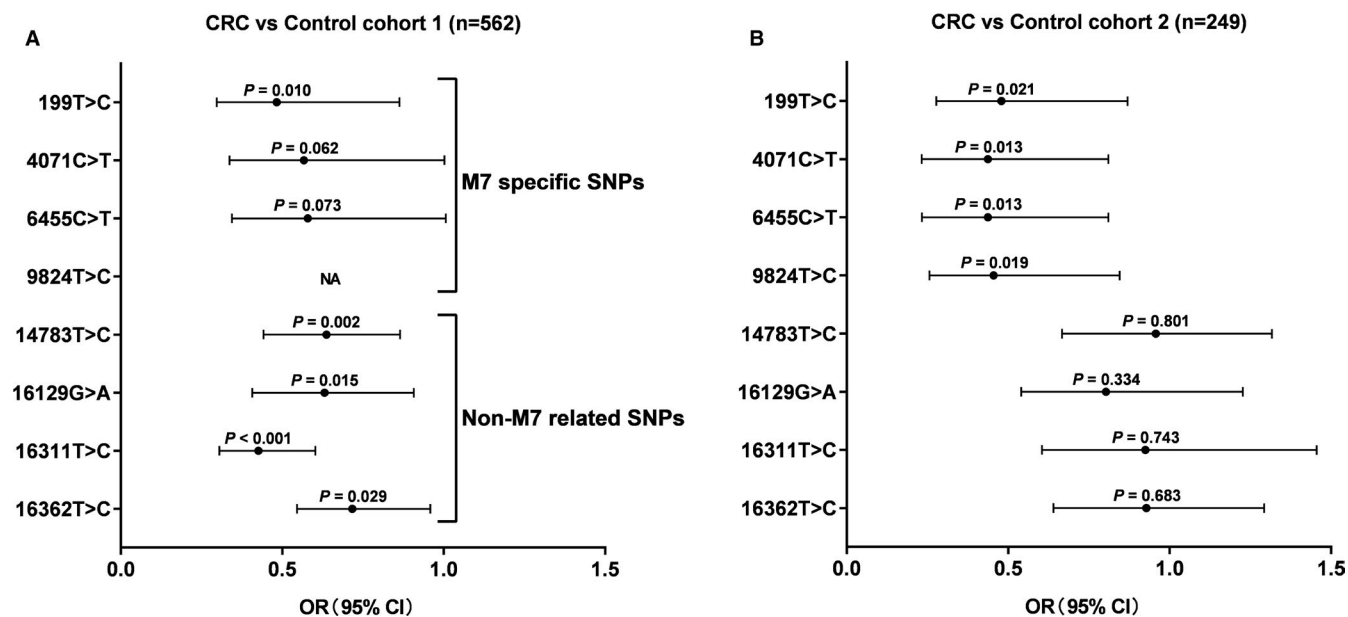


FIGURE 1 Association between mtDNA SNPs and the risk of CRC. (A) Using logistic regression to analyze the correlation between SNPs and CRC in control cohort 1 ($n = 562$). The abscissa represents the odds ratio (OR) value and 95% confidence interval (CI), and the ordinate represents different SNPs. (B) The correlation analysis between SNPs and CRC was performed in the method of logistic regression in validation control cohort ($n = 249$).

4071C>T and 6455C>T did not significantly indicate a decreased risk of CRC in control cohort 1 ($p = 0.062$ and $p = 0.073$, Figure 1A), which may partially be due to sampling bias. However, these three haplogroup M7-specific SNPs were significantly associated with the risk of CRC in control cohort 2 (Figure 1B). Another haplogroup M7 specific SNP 9824T>C also showed a significant association with CRC only in control cohort 2 ($p = 0.019$), because 9824T>C was not detected in control cohort 1.

Moreover, four SNPs catalogued to non-M7 haplogroup, including 14783T>C, 16129G>A, 16311T>C and 16362T>C, also showed significant associations with CRC risk in control cohort1 ($p = 0.002$, $p = 0.015$, $p < 0.001$, $p = 0.029$), but these associations may be artificial due to the complete lack of support in control cohort 2. Thus, the association of haplogroup M7 specific SNPs with CRC provides further evidence that haplogroup M7 is associated with reduced CRC risk in the Han Chinese population.

4 | DISCUSSION

In this study, we comprehensively analysed the association between the mtDNA haplogroups, SNPs and CRC risk in the Northern Han Chinese population. We sequenced mitochondrial genomes from 286 CRC patients and retrieved mtDNA haplogroup data from 562 and 249 Han Chinese normal cohorts from public data sets. We found that haplogroup M7 was significantly associated with a reduced risk of CRC.

Several studies have reported the association between mtDNA haplogroup M7 and disease risk in the Han Chinese population.

For instance, haplogroup M7 has been reported as a risk factor for lung cancer (OR = 2.037, 95%CI = 1.253–3.312, $p = 0.004$)³⁰ in the Southwestern Chinese population. For this study, haplogroup M7 was found to be present in cases at frequencies of 12.67% but was present in controls at frequencies of 5.95%. However, haplogroup M7 was present in controls from Northwestern Chinese at frequencies of 9.43% in our study, it suggested that the opposite associations of haplogroup M7 and diseases exist in different areas of China. We have previously found that haplogroup M7 was associated with a significant reduction in the risk of liver cancer in Northern China.³¹ More recently, Sun et al. reported that haplogroup M7 is associated with an increased risk of disability in an ageing Chinese population.³² In this study, we further found that haplogroup M7 was associated with a reduced risk of CRC. Thus, our results provide further evidence for the involvement of mtDNA haplogroup M7 in ageing-related human diseases, especially cancer.

It is currently unclear that how mtDNA haplogroup M7 may affect the risk of CRC or other diseases. As a unique set of mtDNA polymorphisms, mtDNA haplogroups represent historical mutations accumulated by discrete maternal lineages.³³ In this regard, haplogroup-associated mtDNA variations have the potential to influence mitochondrial OXPHOS function and/or the generation of reactive oxygen species,^{34,35} both of which may modify the risk of cancer. Interestingly, it has been reported that, compared to haplogroup M8, haplogroup M7 has a significantly lower respiratory chain activity and lower mitochondrial membrane potential, together with a 40% reduction in respiration-related oxygen consumption. These findings hint that mitochondrial haplogroup M7 may reduce the risk of CRC due to the presence of sub-optimal mitochondrial

functionality in this haplogroup.³² Lower mitochondrial function may promote tumorigenesis due to altered mitochondrial signals including ROS and NAD⁺/NADH, which are responsible for overactivation of mitochondrial retrograde signaling.³⁶ However, we do not know how cellular metabolism pathways are regulated by mitochondrial function in haplogroup M7 containing cells. Further studies on the causal relationship between OXPHOS function and cellular metabolism pathways are necessary to completely reveal the protective role of haplogroup M7 in CRC.

This study has some limitations. One of the limitations was the small sample size. As mitochondrial haplogroup M7 only accounted for 5.2% of our CRC patient cohort and 9.4% and 11.2% in the two control cohorts, it is necessary to validate our findings in a large, independent case-control study. The control cohort 2 was consisted with Han samples from south and north China, which will also introduce some bias to the results as the higher distribution frequency of haplogroup M7 in south China. Furthermore, the association between mtDNA haplogroup M7 and CRC risk should be further validated by functional studies, which are greatly hindered by our inability to accurately manipulate the mitochondrial genome. In addition, since the majority of mitochondrial respiratory chain proteins are encoded in the nuclear genome, the association between haplogroup M7 and reduced CRC risk may also be affected by the genetic background of the host's nuclear genome.

In conclusion, our study analysed the relationship between mitochondrial genetic variations and the risk of CRC. It revealed for the first time that mtDNA haplogroup M7 reduced the risk of CRC in the Northern Han Chinese population. Our study points to the importance of the mitochondrial genetic background in the development of colorectal cancer.

ACKNOWLEDGEMENT

We thank the patients for their participation in the study.

CONFLICT OF INTEREST

The authors declare that they have no competing interests.

AUTHOR CONTRIBUTION

Qing Yuan: Software (lead); Visualization (lead); Writing-original draft (lead). **Liping Su:** Writing-original draft (equal). **Tian Wang:** Writing-original draft (supporting). **Yang Liu:** Project administration (equal). **Zhenxing Lu:** Formal analysis (equal). **Kaixiang Zhou:** Project administration (equal). **Shanshan Guo:** Formal analysis (equal). **Xiwen Gu:** Writing-review & editing (supporting). **Jinliang Xing:** Writing-review & editing (equal). **Xu Guo:** Writing-review & editing (lead).

DATA AVAILABILITY STATEMENT

The data that supports the findings of this study are available in the supplementary material of this article.

ORCID

Xu Guo  <https://orcid.org/0000-0003-4902-8859>

REFERENCES

- McGuire Shelley. World cancer report 2014. Geneva, Switzerland: world health organization, international agency for research on cancer, WHO press, 2015. *Adv Nutr.* 2016;7(2):418-419. <https://doi.org/10.3945/an.116.012211>
- Zheng ZX, Zheng RS, Zhang SW, Chen WQ. Colorectal cancer incidence and mortality in China, 2010. *Asian Pac J Cancer Prev.* 2014;15(19):8455-8460. <https://doi.org/10.7314/APJCP.2014.15.19.8455>
- Meester RG, Doubeni CA, Zuber AG, et al. Public health impact of achieving 80% colorectal cancer screening rates in the United States by 2018. *Cancer.* 2015;121(13):2281-2285. <https://doi.org/10.1002/cncr.29336>
- Arnold M, Sierra MS, Laversanne M, Soerjomataram I, Jemal A, Bray F. Global patterns and trends in colorectal cancer incidence and mortality. *Gut.* 2017;66(4):683-691. <https://doi.org/10.1136/gutjnl-2015-310912>
- Boursi B, Mamtani R, Hwang WT, Haynes K, Yang YX. A risk prediction model for sporadic CRC based on routine lab results. *Digest Dis Sci.* 2016;61(7):2076-2086. <https://doi.org/10.1007/s10620-016-4081-x>
- Kaminski MF, Polkowski M, Kraszewska E, Rupinski M, Butruk E, Regula J. A score to estimate the likelihood of detecting advanced colorectal neoplasia at colonoscopy. *Gut.* 2014;63(7):1112-1119. <https://doi.org/10.1136/gutjnl-2013-304965>
- Freedman AN, Slattery ML, Ballard-Barbash R, et al. Colorectal cancer risk prediction tool for white men and women without known susceptibility. *J Clin Oncol.* 2009;27(5):686-693. <https://doi.org/10.1200/JCO.2008.17.4797>
- Peters U, Bien S, Zubair N. Genetic architecture of colorectal cancer. *Gut.* 2015;64(10):1623-1636. <https://doi.org/10.1136/gutjnl-2013-306705>
- Sheng WY, Yong Z, Yun Z, Hong H, Hai LL. Toll-like receptor 4 gene polymorphisms and susceptibility to colorectal cancer: a meta-analysis and review. *Arch Med Sci.* 2015;11(4):699-707. <https://doi.org/10.5114/aoms.2015.53288>
- Mik M, Dziki L, Malinowska K, Trzcinski R, Majsterek I, Dziki A. Polymorphism of MSH2 Gly322Asp and MLH1-93G>A in non-familial colon cancer—a case-controlled study. *Arch Med Sci.* 2017;13(6):1295-1302. <https://doi.org/10.5114/aoms.2017.67024>
- Martin FP, Collino S, Rezzi S, Kochhar S. Metabolomic applications to decipher gut microbial metabolic influence in health and disease. *Front Physiol.* 2012;3:113. <https://doi.org/10.3389/fphys.2012.00113>
- Lee HC, Chang CM, Chi CW. Somatic mutations of mitochondrial DNA in aging and cancer progression. *Ageing Res Rev.* 2010;9(Suppl. 1):S47-S58. <https://doi.org/10.1016/j.arr.2010.08.009>
- Blein S, Bardel C, Danjean V, et al. An original phylogenetic approach identified mitochondrial haplogroup T1a1 as inversely associated with breast cancer risk in BRCA2 mutation carriers. *Breast Cancer Res.* 2015;17:61. <https://doi.org/10.1186/s13058-015-0567-2>
- Wang C, Wang Y, Wang H, Zhang R, Guo Z. Mitochondrial DNA haplogroup N is associated good outcome of gastric cancer. *Tumour Biol.* 2014;35(12):12555-12559. <https://doi.org/10.1007/s13277-014-2575-8>
- Li XY, Guo YB, Su M, Cheng L, Lu ZH, Tian DP. Association of mitochondrial haplogroup D and risk of esophageal cancer in Taihang Mountain and Chaoshan areas in China. *Mitochondrion.* 2011;11(1):27-32. <https://doi.org/10.1016/j.mito.2010.06.005>
- Guardado-Estrada M, Medina-Martinez I, Juarez-Torres E, et al. The Amerindian mtDNA haplogroup B2 enhances the risk of HPV for cervical cancer: de-regulation of mitochondrial genes may be involved. *J Hum Genet.* 2012;57(4):269-276. <https://doi.org/10.1038/jhg.2012.17>
- Skonieczna K, Jawien A, Marszalek A, Grzybowski T. Mitogenome germline mutations and colorectal cancer risk in Polish population.

- Arch Med Sci.* 2020;16(2):366-373. <https://doi.org/10.5114/aoms.2018.80893>
18. Li Y, Beckman KB, Caberto C, et al. Association of genes, pathways, and haplogroups of the mitochondrial genome with the risk of colorectal cancer: the multiethnic cohort. *PLoS One.* 2015;10(9):e0136796. <https://doi.org/10.1371/journal.pone.0136796>
 19. Mohammed F, Rezaee Khorasany AR, Mosaieby E, Houshmand M. Mitochondrial A12308G alteration in tRNA^{Leu}(CUN) in colorectal cancer samples. *Diagn Pathol.* 2015;10(1):115. <https://doi.org/10.1186/s13000-015-0337-6>
 20. Guo Z, Zhao S, Fan H, Du Y, Zhao Y, Wang G. Identification of sequence polymorphisms in the D-Loop region of mitochondrial DNA as a risk factor for colon cancer. *Mitochondrial DNA A DNA Mapp Seq Anal.* 2016;27(6):4244-4245. <https://doi.org/10.3109/19401736.2014.1003920>
 21. Govatati S, Saradamma B, Malempati S, et al. Association of mitochondrial displacement loop polymorphisms with risk of colorectal cancer in south Indian population. *Mitochondrial DNA A DNA Mapp Seq Anal.* 2017;28(5):632-637. <https://doi.org/10.3109/24701394.2016.1160076>
 22. Kumar B, Bhat ZI, Bansal S, et al. Association of mitochondrial copy number variation and T16189C polymorphism with colorectal cancer in North Indian population. *Tumor Biology.* 2017;39(11):101042831774029. <https://doi.org/10.1177/1010428317740296>
 23. Yin C, Liu Y, Guo X, et al. An Effective strategy to eliminate inherent cross-contamination in mtDNA next-generation sequencing of multiple samples. *J Mol Diagn.* 2019;21(4):593-601. <https://doi.org/10.1016/j.jmoldx.2019.02.006>
 24. Li X, Guo X, Li DY, et al. Multi-regional sequencing reveals intratumor heterogeneity and positive selection of somatic mtDNA mutations in hepatocellular carcinoma and colorectal cancer. *Int J Cancer.* 2018;143(5):1143-1152. <https://doi.org/10.1002/ijc.31395>
 25. Chen S, Zhou Y, Chen Y, Gu J. fastp: An ultra-fast all-in-one FASTQ preprocessor. *Bioinformatics.* 2018;34(17):i884-i890. <https://doi.org/10.1093/bioinformatics/bty560>
 26. Fan L, Yao YG. MitoTool: a web server for the analysis and retrieval of human mitochondrial DNA sequence variations. *Mitochondrion.* 2011;11(2):351-356. <https://doi.org/10.1016/j.mito.2010.09.013>
 27. van Oven M, Kayser M. Updated comprehensive phylogenetic tree of global human mitochondrial DNA variation. *Hum Mutat.* 2009;30(2):E386-E394. <https://doi.org/10.1002/humu.20921>
 28. Li YC, Ye WJ, Jiang CG, et al. River valleys shaped the maternal genetic landscape of Han Chinese. *Mol Biol Evol.* 2019;36(8):1643-1652. <https://doi.org/10.1093/molbev/msz072>
 29. Zheng HX, Yan S, Qin ZD, et al. Major population expansion of East Asians began before neolithic time: evidence of mtDNA genomes. *PLoS One.* 2011;6(10):e25835. <https://doi.org/10.1371/journal.pone.0025835>
 30. Zheng S, Qian P, Li F, et al. Association of mitochondrial DNA variations with lung cancer risk in a Han Chinese population from southwestern China. *PLoS One.* 2011;6(10):e25835. <https://doi.org/10.1371/journal.pone.0025835>
 31. Chen C, Ba Y, Li D, et al. Genetic variations of mitochondrial genome modify risk and prognosis of hepatocellular carcinoma patients. *Clin Res Hepatol Gastroenterol.* 2017;41(4):378-385. <https://doi.org/10.1016/j.clinre.2016.12.002>
 32. Sun D, Yao S, Wu F, et al. Mitochondrial DNA Haplogroup M7 confers disability in a Chinese aging population. *Front Genet.* 2020;11:577795. <https://doi.org/10.3389/fgene.2020.577795>
 33. Nikitin AG, Ivanova S, Kiosak D, et al. Subdivisions of haplogroups U and C encompass mitochondrial DNA lineages of eneolithic-early bronze age kurgan populations of western North Pontic steppe. *J Hum Genet.* 2017;62(6):605-613. <https://doi.org/10.1038/jhg.2017.12>
 34. Brandon M, Baldi P, Wallace DC. Mitochondrial mutations in cancer. *Oncogene.* 2006;25(34):4647-4662. <https://doi.org/10.1038/sj.onc.1209607>
 35. Fang H, Shen L, Chen T, et al. Cancer type-specific modulation of mitochondrial haplogroups in breast, colorectal and thyroid cancer. *BMC Cancer.* 2010;10(1):10-421. <https://doi.org/10.1186/1471-2407-10-421>
 36. Ma L, Fu Q, Xu B, et al. Breast cancer-associated mitochondrial DNA haplogroup promotes neoplastic growth via ROS-mediated AKT activation. *Int J Cancer.* 2018;142(9):1786-1796. <https://doi.org/10.1002/ijc.31207>

SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section.

How to cite this article: Yuan Q, Su L, Wang T, et al.

Mitochondrial DNA haplogroup M7 confers a reduced risk of colorectal cancer in a Han population from northern China.

J Cell Mol Med. 2021;25:7538-7544. <https://doi.org/10.1111/jcmm.16789>