

Mitogenome germline mutations and colorectal cancer risk in Polish population

Katarzyna Skonieczna¹, Arkadiusz Jawień², Andrzej Marszałek^{3,4}, Tomasz Grzybowski¹

¹Division of Molecular and Forensic Genetics, Department of Forensic Medicine, Faculty of Medicine, Ludwik Rydygier *Collegium Medicum*, Nicolaus Copernicus University, Bydgoszcz, Poland

²Department of Vascular Surgery and Angiology, Faculty of Medicine, Ludwik Rydygier *Collegium Medicum*, Nicolaus Copernicus University, Bydgoszcz, Poland

³Department of Pathology, Faculty of Medicine, Ludwik Rydygier *Collegium Medicum*, Nicolaus Copernicus University, Bydgoszcz, Poland

⁴Department of Tumor Pathology and Prophylaxis, Poznan University of Medical Sciences and Greater Poland Cancer Center, Poznan, Poland

Submitted: 3 August 2018

Accepted: 9 December 2018

Arch Med Sci 2020; 16 (2): 366–373

DOI: <https://doi.org/10.5114/aoms.2018.80893>

Copyright © 2019 Termedia & Banach

Corresponding author:

Katarzyna Skonieczna PhD
Division of Molecular
and Forensic Genetics
Department of
Forensic Medicine
Faculty of Medicine
Ludwik Rydygier
Collegium Medicum
Nicolaus Copernicus
University
9 Skłodowskiej-Curie St
85-094 Bydgoszcz, Poland
Phone: +48 52 585 35 52
E-mail: k.skonieczna@gmail.com

Abstract

Introduction: To date, several nuclear DNA variants have been shown to be associated with increased risk of developing colorectal cancer. Despite the fact that mitochondria play an important role in carcinogenesis, little is known about inherited mitochondrial DNA mutations that could be involved in this disease. Thus, potential associations between inherited mutations in the entire mitochondrial genomes and colorectal cancer were analysed in this study.

Material and methods: Two hundred mitogenome sequences determined for colorectal cancer patients and healthy individuals from Poland were used to investigate the association between mtDNA alleles or haplogroups and colorectal cancer. Additional mtDNA control region haplotypes determined for 1353 individuals from the general Polish population were used for comparison of haplogroup and certain allele frequencies between case and control groups.

Results: The non-R clades together with their diagnostic T alleles at positions 12705 and 16223 were observed with higher frequencies in healthy individuals than in colorectal cancer patients. Nevertheless, the differences of the R macrohaplogroup (as well as 12705 or 16223 alleles) frequencies between cases and controls were statistically insignificant after Bonferroni correction. Most of the non-R clades were of Asian and African origin, but none of them were prevalent in the control group. Moreover, neither mtDNA alleles nor haplogroups were associated with clinicopathological parameters of colorectal cancer patients.

Conclusions: Contrary to some previous reports, the findings of this study do not support the hypothesis that mitochondrial DNA variants contribute to inherited predisposition to colorectal cancer.

Key words: mtDNA, mitochondrial genome, polymorphisms, hereditary mutations, colorectal cancer.

Introduction

Colorectal cancer (CC) is the third in males and second in females most commonly diagnosed type of cancer. The incidence of this disease

varies depending on geographical location. The highest morbidity of this disease was reported in Europe, Northern America and Oceania, whereas the lowest was reported in Africa [1]. Annually about 18000 individuals are diagnosed and about 12000 persons die because of colorectal cancer in Poland [2]. Genetic factors could help to identify individuals with increased risk of developing colorectal cancer. Persons at increased risk of this disease could be further subjected to appropriate screening programmes, leading to earlier cancer diagnosis and subsequent improved survival [3, 4]. To date, about 40 common alleles of nuclear DNA, which are risk factors for colorectal cancer, have been identified [5–8]. These susceptibility loci account for only about 8% to 16% [5] of CC cases. Thus, additional genetic factors that contribute to this disease still need to be discovered.

Mitochondrial DNA (mtDNA) polymorphisms were previously claimed to be associated with different cancer types, e.g. breast cancer or prostate cancer [9]. The human mitochondrial genome encodes 13 proteins which are part of the oxidative phosphorylation complexes and thus are essential for proper functioning of mitochondria [10]. The alterations in mtDNA sequence could subsequently change the mitochondrial gene expression or protein sequence, which could impair mitochondrial functions. Importantly, dysregulated expression of mitochondrial proteins encoded by nuclear genes were reported to be potentially relevant in colorectal carcinogenesis [11]. Notably, mitochondria play an essential role in the production of reactive oxygen species (ROS) and the regulation of apoptosis – processes relevant to cancer development [9]. Thus, it is necessary to verify whether mtDNA contributes to an individual's risk of cancer development.

The mitogenome is inherited only from the mother and its mutation rate is about 10 times higher than nuclear DNA [12]. In consequence, every few generations, a random mutation is introduced into the mitochondrial DNA sequence and modifies its mutational signature. All currently known mitogenome lineages evolved from a common ancestor (called “Mitochondrial Eve”) through the last 200 000 years [13]. Since then many new mtDNA lineages have arisen in human populations. As different mutations were independently introduced into various mtDNA lineages, each of them harbours a specific pattern of mutations. Thus, on the basis of the presence of certain mutational motifs, mitochondrial DNA sequences (haplotypes) are classified into haplogroups (groups of sequences that share a set of mutations inherited from a common ancestor).

The mitochondrial DNA sequence is highly variable among populations originating from different

geographical locations. For instance, H, J, K, T, V, X and U haplogroups are common in Europeans, whereas A, B, C and D clades occur frequently in Asians [14]. The recent landscape of mtDNA sequence variability results from complex, ancient migration events. As mtDNA allele and haplogroup frequencies highly vary between different populations, it is important to evaluate the genetic risk of cancer in a specific ethnic context. To date, several studies have been carried out to investigate the associations of single mitochondrial DNA polymorphic sites (SNPs) with colorectal cancer in different ethnic groups [15–20]. The analysed SNPs were hotspots or positions diagnostic for major haplogroups. Nevertheless, huge variability within the major haplogroups in human mtDNA phylogenetic tree could be observed. For instance, within the H haplogroup a total of 108 subhaplogroups were identified, and they are defined by a variety of mutations [21]. Some of these subhaplogroups were found to be characteristic for specific ethnic groups [22]. Thus, it is relevant to analyse the variability of the entire mitochondrial genome sequence within a specific ethnic group to investigate the mtDNA associations with colorectal cancer.

Previous studies suggested that certain polymorphic mtDNA positions may be associated with increased risk of colorectal cancer in Iranian, Indian or European Americans [16–20]. In contrast, the studies performed on other ethnic groups (e.g. British, African Americans, Asian Americans, Latinos or Native Hawaiians) excluded the potential impact of mtDNA variability in developing CC [15, 16]. Nevertheless, it is still unclear whether the mtDNA variability might be a risk factor enabling predicting the colorectal cancer development in the Polish population. It should also be noted that the association studies performed so far were limited to certain mtDNA positions that encompass less than 10% of all positions in the mtDNA molecule [15–20]. Thus, it could not be excluded that mtDNA variability outside those positions might be associated with the occurrence of colorectal cancer. This study considers the complete mitochondrial genome nucleotide variation with respect to colorectal cancer development.

To comprehensively investigate whether mitochondrial DNA is a risk factor for colorectal cancer the entire mitochondrial genome sequences of 100 colorectal cancer patients and 100 healthy individuals of Polish origin from the Pomerania-Kujawy region (Poland) were analysed in this study. Additionally, for variants located within hypervariable regions I and II of the mtDNA control region, the statistical calculations were performed using the control group increased to 1353 individuals based on reported data from the general

Polish population [22–25]. The analysis has shown that the R macrohaplogroup and its diagnostic mutations at positions 12705 and 16223 were observed with higher frequencies in colorectal cancer patients compared to healthy individuals. Nevertheless, the underrepresentation of non-R clades in Polish colorectal cancer patients may not have any relevance to colorectal cancer.

Material and methods

Ethics approval and consent to participate

The experiments were approved by the local Bioethics Committee of the Ludwik Rydygier Collegium Medicum, Nicolaus Copernicus University Bydgoszcz, Poland (statements no. KB/32/2002, KB/299/2003, KB/414/2008, KB 432/2008, KB/466/2010, and KB 43/2016). Written consent from all the participants was obtained before the sample collection and subsequent analysis.

mtDNA sequences used in this study

All cases and controls were of Polish origin. Haplotypes of all mitochondrial DNA sequences were determined based on the comparison with the revised Cambridge Reference Sequence, rCRS [5] and published previously by Mielnik-Sikorska *et al.* [22], Skonieczna *et al.* [26], Malyarchuk *et al.* [23], Malyarchuk *et al.* [24] and Grzybowski *et al.* [25].

Entire mtDNA haplotypes were deposited in GenBank under the accession numbers JX128058, JX128059, KM047188–KM047235, MF177134–MF177183 and KY782150–KY782249. Haplogroup assignment was performed according to the updated human mtDNA phylogenetic tree [21] (phylo-tree build 17). Insertions and deletions within the mononucleotide repeat region between positions 303–315 and 16184–16193 were disregarded from the analysis.

Case group – mitogenomes

The patients were recruited to the study at the Department of Vascular Surgery and Angiology NCU CM in Bydgoszcz. The recruitment procedure, clinicopathological features of the patients and somatic mtDNA mutations were described by Skonieczna *et al.* [26]. The case group consisted of 100 unrelated colorectal cancer patients from the Pomerania-Kujawy region of Poland (41 females and 59 males). The mean age of colorectal cancer patients recruited to the study was 65.3 ± 10.7 years. Forty-four patients were diagnosed with right-sided CC, whereas 56 persons had left-sided CC. Lymph node metastases were identified in 38 colorectal cancer patients. Stage I (T1–T2, N0, M0) was diagnosed in 29%, stage II (T3–T4, N0, M0) in 33%, stage III (any T, N1–2, M0) in 34%, and

stage IV (any T, any N, M1) in 4% of CC patients. Mitochondrial genome haplotypes determined for normal cells of those patients were described by Skonieczna *et al.* [26]. Concomitant presence of somatic mtDNA substitutions and indels was observed in 35 colorectal cancer patients. Exclusively somatic substitutions in mitogenomes were identified in 41 CC patients. Solely somatic mtDNA indels were observed in 10 colorectal cancer patients.

Control group I – mitogenomes

The control group consisted of 100 unrelated, healthy individuals of matched ethnicity, from the Pomerania-Kujawy region (54 females and 46 males), for which mitogenome sequences were described by Malyarchuk *et al.* [23]. The entire mitochondrial genome sequences (nucleotide positions 1–16569 bp) were used to analyse the associations between allele or haplogroup status and colorectal cancer risk.

Control group II–HVI and HVII regions

For comparison purposes, phylogenetically checked hypervariable region I (HVI) and II (HVII) haplotypes determined for 1353 individuals from the general Polish population and described by Mielnik-Sikorska *et al.* [22], Malyarchuk *et al.* [23], Malyarchuk *et al.* [24] and Grzybowski *et al.* [25] were also used. The above-mentioned HVI and HVII haplotypes were determined for individuals from Pomerania-Kujawy ($n = 536$), Pomerania ($n = 166$), Kaszuby ($n = 290$), Suwalszczyzna ($n = 73$), Podhale ($n = 201$) and Upper Silesia ($n = 87$) regions. HVI and HVII haplotypes were used to analyse colorectal cancer risk with respect to haplogroup status or mtDNA allele observed within nucleotide position ranges: 16024–16400 (encompassing the HVI region) or 30–407 (containing the HVII region).

Statistical analysis

To analyse the genetic relationships of the Polish subpopulations (individuals from Kashubia; Podhale; Pomerania; Pomerania-Kujawy; Suwałki; Upper Silesia regions as well as colorectal cancer patients from Pomerania-Kujawy region) involved in this study, the principal component analysis (PCA) was performed on mtDNA haplogroup frequencies using *MVSP*, v. 3.22 (Kovach Computing Services, Anglesey, UK).

Statistical significance of the differences between the two groups (cases and controls) and the mtDNA allele or haplogroup status was calculated using the χ^2 test with Yates' correction. Statistical significance of the differences between clinicopathological features and the mtDNA allele or haplogroup status was calculated using the χ^2

test with Yates' correction. Mitochondrial DNA nucleotide variants with frequencies less than 4% or above 96% were excluded from the statistical analysis between colorectal cancer patients and healthy individuals. Statistical calculations were performed using Statistica, v. 13.1 software (StatSoft Inc, Round Rock, TX, USA). The odds ratio (OR) and 95% confidence intervals (95% CI) were also calculated. The p -values of less than 0.05 were considered as statistically significant. Correction for multiple testing was performed using the Bonferroni adjustment.

Statistical power and target number of individuals in cases and controls were calculated with Mitopower software [27].

Results

The reliability of the mtDNA haplotypes used in this study was confirmed by phylogenetic methods in previous reports [22–26].

Haplogroup assignment and colorectal cancer risk

Principal component analysis (PC) has shown that the Polish population is quite homogeneous in terms of mitochondrial haplogroup frequencies. In particular, four Polish subpopulations from Pomerania-Kujawy, Kashubia, Pomerania and Podhale regions group together (Supplementary Figure S1). The Suwałki region appears slightly distant based on PCA1 variance, whereas colorectal cancer patients from the Pomerania-Kujawy region are slightly distant based on PCA2 variance (Supplementary Figure S1). Taken together, the results of PCA provide no proof for hidden stratification within the Polish population (Supplementary Figure S1).

The most prevalent clade in cases and controls was the R macrohaplogroup. Within this clade, the most commonly observed was the H haplogroup in all studied groups. Also the J, T, U and W clades were frequently observed in cases and controls. The frequencies of other haplogroups were below 4%. The frequency of the R macrohaplogroup was significantly higher in colorectal cancer patients than in the general Polish population ($p = 0.0387$; OR = 3.4955 with 95% CI: 1.0926–11.183) or exclusively in healthy controls from the Pomerania-Kujawy region ($p = 0.0066$; OR = 5.7059 with 95% CI: 1.597–20.3861; Table I). Nevertheless, these differences were statistically insignificant after Bonferroni correction ($p^B = 0.4644$ and $p^B = 0.0792$, respectively). The frequencies of the other mitochondrial haplogroups identified in the Polish population did not differ significantly between the colorectal cancer patients and the persons from the general Polish population or healthy individuals from the Pomerania-Kujawy region solely (Table I).

MtDNA allele frequency and colorectal cancer risk

The comparison of one hundred mitogenomes from healthy individuals of Polish origin to the rCRS sequence revealed 2527 polymorphisms in 605 different positions. Twenty-seven (about 1%) polymorphisms in 12 positions were insertions, whereas 12 polymorphisms (about 0.5%) in three positions were deletions. The majority (about 98.5%) of hereditary mtDNA mutations were substitutions located in 596 positions (Supplementary Table S1). Hereditary substitutions were predominantly homoplasmic, and only seven were in a heteroplasmic state (Supplementary Table S1). In 110 out of 611 mtDNA polymorphic positions, the minor allele frequency exceeded the 4% level in healthy persons from the Pomerania-Kujawy region. In mitogenomes of Polish colorectal cancer patients 2325 polymorphisms at 536 different positions were found. Approximately 1.1% of all polymorphisms in six positions were insertions, whereas about 0.5% of all hereditary mutations were deletions located in one position. Hereditary mtDNA mutations were predominantly substitutions (about 98.4%) located in 557 positions (Supplementary Table S1). The majority of those substitutions were homoplasmic (about 98.1%), and only about 1.9% were in a heteroplasmic state (Supplementary Table S1). In 98 polymorphic positions of colorectal cancer mitogenomes, the minor allele exceeded the level of 4%.

Comparison of allele frequencies between the studied and control groups have shown that the frequencies of the T allele at positions 12705 and 16223 (C alleles in both positions constitute a complete set of mutations diagnostic for the R macrohaplogroup) were more than four times higher in healthy individuals than in colorectal cancer patients ($p = 0.0052$; OR = 5.7059 with 95% CI: 1.597–20.3861 and $p = 0.0046$; OR = 4.9157 with 95% CI: 1.5909–15.1888, respectively; Supplementary Table SII). Nevertheless, these differences were not statistically significant after Bonferroni adjustment ($p^B = 1.000$ for each allele). The frequencies of other mitochondrial alleles did not differ significantly between the cases and controls (Supplementary Table SII).

Mitogenome variability and clinicopathological features of colorectal cancer patients

The molecular data were also associated with the clinicopathological features. Associations were performed for haplogroups or mtDNA alleles that were observed in Polish CC patients with frequencies equal to or above 4%. The analysis of the haplogroup distribution between the pa-

tients with different clinicopathological parameters showed no association between haplogroup status and gender, age, tumour site (left or right), lymph node metastasis or somatic mitogenome substitutions and/or indels in colorectal cancer cells (Supplementary Table SIII). Also, the mtDNA alleles at 99 different polymorphic positions were not associated with clinicopathological features

Table I. Associations between mtDNA haplogroups and colorectal cancer in a Polish population

| Haplogroups | Case group (N = 100) n (%) | Control group I (N = 100) | P-value | P ^B -value | Control group II (N = 1353) | P-value | P ^B -value |
|-------------|----------------------------------|---------------------------------|---------|-----------------------|-----------------------------------|---------|-----------------------|
| R* | 97.0 (97) | 85.0 (85) | 0.0066 | 0.0792 | 90.2 (1221) | 0.0387 | 0.4644 |
| HV | 48.0 (48) | 35.0 (35) | 0.0851 | 1.0000 | 41.2 (632) | 0.8844 | 1.0000 |
| H | 44.0 (44) | 35.0 (35) | 0.2472 | 1.0000 | 45.3 (568) | 0.7721 | 1.0000 |
| HV0 | 4.0 (4) | 0.0 (0) | 0.1297 | 1.0000 | 4.1 (55) | 0.8175 | 1.0000 |
| V | 3.0 (3) | 0.0 (0) | 0.2447 | 1.0000 | 4.0 (54) | 0.8214 | 1.0000 |
| J | 7.0 (7) | 8.0 (8) | 1.0000 | 1.0000 | 9.4 (118) | 0.6838 | 1.0000 |
| T: | 13.0 (13) | 13.0 (13) | 0.8335 | 1.0000 | 11.1 (139) | 0.4900 | 1.0000 |
| T1 | 2.0 (2) | 4.0 (4) | 0.6785 | 1.0000 | 2.5 (31) | 0.8735 | 1.0000 |
| T2 | 11.0 (11) | 9.0 (9) | 0.8137 | 1.0000 | 7.9 (107) | 0.3668 | 1.0000 |
| U: | 28.0 (28) | 26.0 (26) | 0.8735 | 1.0000 | 25.5 (319) | 0.3792 | 1.0000 |
| U1 | 1.0 (1) | 0.0 (0) | 1.0000 | 1.0000 | 0.5 (6) | 0.9782 | 1.0000 |
| U2 | 2.0 (2) | 1.0 (1) | 1.0000 | 1.0000 | 1.2 (15) | 0.7505 | 1.0000 |
| U3 | 1.0 (1) | 0.0 (0) | 1.0000 | 1.0000 | 1 (12) | 0.6640 | 1.0000 |
| U4 | 5.0 (5) | 10.0 (10) | 0.2829 | 1.0000 | 5.4 (68) | 0.8214 | 1.0000 |
| U5*: | 13.0 (13) | 9.0 (9) | 0.4978 | 1.0000 | 12.1 (151) | 0.6912 | 1.0000 |
| U5a | 8.0 (8) | 6.0 (6) | 0.7817 | 1.0000 | 7.3 (92) | 0.8004 | 1.0000 |
| U5b | 5.0 (5) | 3.0 (3) | 0.7182 | 1.0000 | 4.7 (59) | 0.9616 | 1.0000 |
| U7 | 0.0 (0) | 0.0 (0) | 1.0000 | 1.0000 | 0.2 (2) | 0.3112 | 1.0000 |
| U8 | 7.0 (7) | 6.0 (6) | 1.0000 | 1.0000 | 4.7 (64) | 0.4380 | 1.0000 |
| K | 6.0 (6) | 5.0 (5) | 1.0000 | 1.0000 | 4.2 (53) | 0.4498 | 1.0000 |
| RO | 0.0 (0) | 1.0 (1) | 1.0000 | 1.0000 | 0.5 (6) | 0.8881 | 1.0000 |
| I | 1.0 (1) | 1.0 (1) | 1.0000 | 1.0000 | 1.9 (26) | 0.7834 | 1.0000 |
| W | 2.0 (2) | 6.0 (6) | 0.2790 | 1.0000 | 3.5 (47) | 0.6165 | 1.0000 |
| N | 0.0 (0) | 2.0 (2) | 0.4773 | 1.0000 | 1.0 (14) | 0.6229 | 1.0000 |
| A | 0.0 (0) | 1.0 (1) | 1.0000 | 1.0000 | 0.1 (2) | 0.3112 | 1.0000 |
| C | 0.0 (0) | 1.0 (1) | 1.0000 | 1.0000 | 0.5 (7) | 0.9782 | 1.0000 |
| D | 0.0 (0) | 1.0 (1) | 1.0000 | 1.0000 | 0.2 (3) | 0.5028 | 1.0000 |
| E | 0.0 (0) | 0.0 (0) | 1.0000 | 1.0000 | 0.1 (1) | 0.0884 | 1.0000 |
| G | 0.0 (0) | 1.0 (1) | 1.0000 | 1.0000 | 0.3 (4) | 0.6567 | 1.0000 |
| M | 0.0 (0) | 0.0 (0) | 1.0000 | 1.0000 | 0.1 (1) | 0.0884 | 1.0000 |
| L | 0.0 (0) | 0.0 (0) | 1.0000 | 1.0000 | 0.3 (4) | 0.6567 | 1.0000 |
| X | 0.0 (0) | 2.0 (2) | 0.4773 | 1.0000 | 1.7 (23) | 0.3686 | 1.0000 |

Case group – 100 colorectal cancer patients for which mitogenome sequences were described by Skonieczna et al. [26]; control group I – 100 healthy persons from Polish population for which mitogenome sequences were described by Malyarchuk et al. [23]; control group II – general Polish population of 1353 individuals for which HVI and HVI regions were described by Malyarchuk et al. [23]; Malyarchuk et al. [24]; Grzybowski et al. [25] and Mielnik-Sikorska et al. [22]. Numbers of haplotypes belonging to specific haplogroups are in parentheses. Major haplogroups are shown in bold. Frequencies of each subhaplogroup within a particular major clade are added to the overall frequency of this major haplogroup. Statistical significance was calculated with the chi-square test with Yates' correction. *Macrohaplogroup; p^B – Bonferroni corrected p-values.

such as age, sex, tumour location (left or right site), lymph node metastasis or somatic mtDNA mutations (substitutions and/or indels) in CC cells (Supplementary Table SIV).

Discussion

A limited number of mtDNA polymorphisms analysed in previous studies focused on genetic risk factors for colorectal cancer [15–20]. In some of them, only a single SNP was analysed [17, 20]. To provide a more comprehensive insight into the relationship between mtDNA variation and colorectal cancer occurrence, the entire mitochondrial genome sequences were analysed in this study. According to the guidelines recommended by internationally recognized experts, the studies of the mtDNA associations with cancer should fulfil several requirements [28–34]. First and foremost, to avoid the problem of population stratification, there is a need for recruiting individuals of the same ethnic/biogeographical origin to both case and control groups. In line with this recommendation, the strength of the research described here lies in proper matching of the control group to the group of colorectal cancer patients. Indeed, persons from both groups were of Polish origin and were recruited from the same geographical part of Poland, reduced to the small Pomerania-Kujawy region. Moreover, the PC analysis showed no hidden stratification within the Polish subpopulation. Additionally, the case and control groups used for mitogenome comparisons were equal in the number of investigated individuals.

In the studies conducted so far no associations between mtDNA variation and colorectal cancer risk have been reported in large British [15], African American, Asian American, Latino and Native Hawaiian [16] cohorts. Other studies performed on much smaller groups of patients suggested an association between single mtDNA polymorphic positions (e.g. m.12308A>G or m.16189T>C) and colon or rectal cancer in Iranian [17] and Indian [19, 20] populations. Also, the T haplogroup in European Americans [16] was suggested to be associated with colorectal cancer.

Nevertheless, caution should be taken regarding the above - mentioned mtDNA risk factors. In particular, *a posteriori* phylogenetic analysis of the putative associations of mtDNA with breast cancer showed a variety of shortcomings [35]. These errors were related to genotyping artefacts, inadequate control group structure or a small number of patients analysed [35]. Similar discrepancies could also be observed in some of the colorectal cancer association studies performed previously. For instance, the association of the m.12308A>G allele with colorectal cancer was evaluated based on the analysis of only 30 patients and 100 con-

trols [17]. Moreover, Mohammed *et al.* [17] used tumour samples instead of normal tissues for genotyping mtDNA at position 12308 in colorectal cancer patients. This approach could be misleading as the tumour sample could have a somatic mtDNA mutation at this position. In this respect it is worth noting that the mtDNA haplotypes from this study were determined for normal colon cells for patients. Moreover, in line with the international recommendations [28–34], the accuracy of all mitogenome sequences used here was confirmed with phylogenetic methods in previous reports [22–26]. The above - mentioned quality measures guarantee the authenticity of the genotyping and further association results presented in this study.

Furthermore, the strength of this research lies in the phylogenetic approach used for grouping of mtDNA sequences for the purpose of comparisons between cases and controls, which is highly recommended by internationally recognized experts [33]. Indeed, classification of each mtDNA clade presented in this study (Table I, Supplementary Table III) was based on a specific set of mtDNA mutations fixed during evolution, which are delineated in a reconstructed worldwide human phylogenetic tree [21, build 17, updated 18 February 2016].

According to the international recommendations, to avoid the inflation of type 1 error, the statistical evaluations should be corrected for multiple comparisons using the Bonferroni method (determined by a number of mtSNPs or independent haplogroups) or a permutation approach. In this respect, it should be noted that while the R macrohaplogroup and its diagnostic C alleles at positions 12705 and 16223 [21] were more frequent in cases than in controls in this study, the Bonferroni corrected *p*-values were not significant. Considering this, the results obtained in this study should be taken with caution and further validation studies using independent cohorts are needed to confirm this apparent difference in frequencies. The 12705 position, located in the *MT-ND5* gene, is highly evolutionarily stable in human mtDNA [10]. In fact, transition at this position is specific for two lineages only, African L0g and the R macrohaplogroup, widely distributed across the world [21]. Importantly, all European lineages are descendants of the R clade [14]. The C to T transition at 12705 is a synonymous mutation, for which the conservation index is about 31%, which suggests little if any effect on cell phenotype. Also, the transition at position 16223, located at hypervariable region I [10], does not seem to play any role in modifying cell phenotype. Indeed, 16223 is a frequently mutated site in the human mitogenome and transition at this position is observed in different mutational patterns of 33 distant

haplogroups [21]. Altogether, wide distribution of both 12705 and 16223 transitions across different human populations and the potential lack of any effect on cell phenotype suggest that they might not be associated with colorectal cancer development. Thus, colorectal cancer must be determined by other molecular factors.

Importantly, transitions at both 12705 and 16223 positions are the only variants that are diagnostic for the R macrohaplogroup [21]. It is worth noting that the non-R clades identified in the general Polish population have non-European origin and are components widely found in Africa (e.g. L clade) and Asia (e.g. A, C, D and G haplogroups) [14]. Those African and Asian lineages might have been introduced into the Polish population as a result of medieval migrations of Asian individuals from Siberia and Western Asia as well as during recent migrations of Ashkenazi Jews from Germany to Poland [22]. Interestingly, the incidence of colorectal cancer in Asian and African populations is much lower than in European ethnic groups [1]. Thus, underrepresentation of the non-R haplogroups in the group of Polish colorectal cancer patients may reflect the lower risk for this disease in Asians and Africans. Importantly, none of the Asian and African lineages in Polish controls was significantly more prevalent in healthy Polish individuals. Altogether, the obtained results may suggest that some other factors, e.g. nuclear DNA mutations, in those non-R Polish individuals are protective against colorectal cancer.

On the other hand, frequencies of non-R haplogroups were low and the frequency difference of the R macrohaplogroup between cases and controls was not large enough to consider a strong association according to the international recommendations [28–34]. Furthermore, this potential association was lost after Bonferroni correction. The next limitation of this study is the small sample size of the analysed case and control groups, which in consequence contributed to the low statistical power of the obtained results. Indeed, assuming two categories (R and non-R) and the χ^2 test with 5000 permutations, our study had only 65.8% (assuming 1353 controls) or 40.2% power (assuming 100 controls) [27]. Assuming twelve categories (twelve major haplogroups found in the Polish population) our study had even lower power of 3.4% (assuming 1353 controls). The underpowered results suggest that the association of the R haplogroup (and its diagnostic m.12705C>T and m.16223C>T transitions) with colorectal cancer risk in the Polish population can also be explained by chance. Thus, following international recommendations [28–34], further validation studies replicated in larger cohorts are necessary to finally resolve the issue of the obtained associations. In

fact, subsequent calculations have shown that to achieve 80% power for the two categories (R and non-R clades) the case and control groups should consist of at least 689 individuals each, whereas assuming twelve categories (haplogroups), the sample size should consist of at least 1336 individuals [27].

In conclusion, although macrohaplogroup R and its diagnostic mutations at positions 12705 and 16223 were observed with higher frequency in colorectal cancer patients than in healthy individuals, the findings of this study do not support the hypothesis that mitochondrial DNA variants play a significant role in inherited predisposition to colorectal cancer. Also, further analysis has not shown any mtDNA factors that could be associated with clinicopathological outcomes of colorectal cancer patients. Nevertheless, further studies performed on larger cohorts are necessary to finally verify the underrepresentation of non-R haplogroups in Polish colorectal cancer patients.

Acknowledgments

The work was supported partially by a Polish Ministry of Science and Higher Education grant (no. N N301 075839) and a grant from the Faculty of Medicine, Collegium Medicum, Nicolaus Copernicus University in Poland (no. MN-4/WL/2016).

Conflict of interest

The authors declare no conflict of interest.

References

1. 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: 683-91.
2. Didkowska J, Wojciechowska U, Olasek P. Cancer in Poland in 2015. Polish National Cancer Registry, Department of Epidemiology and Cancer Prevention, Poland. 2015; 1-102.
3. Ahmadinejad M, Pouryaghobi SM, Bayat F, et al. Surgical outcome and clinicopathological characteristics of emergency presentation elective cases of colorectal cancer. *Arch Med Sci* 2018; 14: 826-9.
4. Matyja M, Pasternak A, Szura M, Wysocki M, Pędziwiatr M, Rembiasz K. How to improve the adenoma detection rate in colorectal cancer screening? Clinical factors and technological advancements. *Arch Med Sci* In press; <https://doi.org/10.5114/aoms.2018.74863>.
5. Peters U, Bien S, Zubair N. Genetic architecture of colorectal cancer. *Gut* 2015; 64: 1623-36.
6. 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: 699-707.
7. 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: 1295-302.

8. Lamami Y, Mesediyeva R, Arikan S, et al. Preliminary report: one of the PD-1 gene variants may be a valuable marker for colorectal cancer. *Arch Med Sci Civil Dis* 2018; 3: e34-40.
9. Brandon M, Baldi P, Wallace DC. Mitochondrial mutations in cancer. *Oncogene* 2006; 25: 4647-62.
10. Andrews RM, Kubacka I, Chinnery PF, Lightowlers RN, Turnbull DM, Howell N. Reanalysis and revision of the Cambridge reference sequence for human mitochondrial DNA. *Nat Genet* 1999; 23: 147.
11. Zhang K, Chen Y, Huang X, et al. Expression and clinical significance of cytochrome c oxidase subunit IV in colorectal cancer patients. *Arch Med Sci* 2016; 12: 68-77.
12. Brown WM, Prager EM, Wang A, Wilson AC. Mitochondrial DNA sequences of primates: tempo and mode of evolution. *J Mol Evol* 1982; 18: 225-39.
13. Soares P, Ermini L, Thomson N, et al. Correcting for purifying selection: an improved human mitochondrial molecular clock. *Am J Hum Genet* 2009; 84: 740-59.
14. Soares P, Achilli A, Semino O, et al. The archaeogenetics of Europe. *Curr Biol* 2010; 20: R174-83.
15. Webb E, Broderick P, Chandler I, et al. Comprehensive analysis of common mitochondrial DNA variants and colorectal cancer risk. *Br J Cancer* 2008; 99: 2088-93.
16. 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: e0136796.
17. Mohammed F, Rezaee Khorasany AR, Mosaieby E, Houshmand M. Mitochondrial A12308G alteration in tRNA(Leu(CUN)) in colorectal cancer samples. *Diagn Pathol* 2015; 10: 115.
18. 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: 4244-5.
19. 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: 632-7.
20. Kumar B, Bhat ZI, Bansal S, et al. Association of mitochondrial copy number variation and T16189C polymorphism with colorectal cancer in North Indian population. *Tumour Biol* 2017; 39: 1010428317740296.
21. van Oven M, Kayser M. Updated comprehensive phylogenetic tree of global human mitochondrial DNA variation. *Hum Mutat* 2009; 30: E386-94.
22. Mielnik-Sikorska M, Daca P, Malyarchuk B, et al. The history of Slavs inferred from complete mitochondrial genome sequences. *PLoS One* 2013; 8: e54360.
23. Malyarchuk B, Litvinov A, Derenko M, et al. Mitogenomic diversity in Russians and Poles. *Forensic Sci Int Genet* 2017; 30: 51-6.
24. Malyarchuk BA, Grzybowski T, Derenko MV, Czarny J, Woźniak M, Miścicka-Sliwka D. Mitochondrial DNA variability in Poles and Russians. *Ann Hum Genet* 2002; 66: 261-83.
25. Grzybowski T, Malyarchuk BA, Derenko MV, Perkova MA, Bednarek J, Woźniak M. Complex interactions of the Eastern and Western Slavic populations with other European groups as revealed by mitochondrial DNA analysis. *Forensic Sci Int Genet* 2007; 1: 141-7.
26. Skonieczna K, Malyarchuk B, Jawień A, et al. Mitogenomic differences between the normal and tumor cells of colorectal cancer patients. *Hum Mutat* 2018; 39: 691-701.
27. Pardo-Seco J, Amigo J, González-Manteiga W, Salas A. A generalized model to estimate the statistical power in mitochondrial disease studies involving 2×k tables. *PLoS One* 2013; 8: e73567.
28. Zhang AM, Jia X, Bi R, et al. Mitochondrial DNA haplogroup background affects LHON, but not suspected LHON, in Chinese patients. *PLoS One* 2011; 6: e27750.
29. Wang D, Su LY, Zhang AM, et al. Mitochondrial DNA copy number, but not haplogroup, confers a genetic susceptibility to leprosy in Han Chinese from Southwest China. *PLoS One* 2012; 7: e38848.
30. Zhang W, Tang J, Zhang AM, et al. A matrilineal genetic legacy from the last glacial maximum confers susceptibility to schizophrenia in Han Chinese. *J Genet Genomics* 2014; 41: 397-407.
31. Bi R, Zhang W, Yu D, et al. Mitochondrial DNA haplogroup B5 confers genetic susceptibility to Alzheimer's disease in Han Chinese. *Neurobiol Aging* 2015; 36: 1604.e7-16.
32. Fachal L, Mosquera-Miguel A, Pastor P, et al. No evidence of association between common European mitochondrial DNA variants in Alzheimer, Parkinson, and migraine in the Spanish population. *Am J Med Genet B Neuropsychiatr Genet* 2015; 168B: 54-65.
33. Salas A, Elson JL. Mitochondrial DNA as a risk factor for false positives in case-control association studies. *J Genet Genomics* 2015; 42: 169-72.
34. Zhang AM, Hu QX, Liu FL, et al. Mitochondrial DNA haplogroup a decreases the risk of drug addiction but conversely increases the risk of HIV-1 infection in Chinese addicts. *Mol Neurobiol* 2016; 53: 3873-81.
35. Salas A, García-Magariños M, Logan I, Bandelt HJ. The saga of the many studies wrongly associating mitochondrial DNA with breast cancer. *BMC Cancer* 2014; 14: 659.