

Ulcerative Colitis Is Under Dual (Mitochondrial and Nuclear) Genetic Control

Alexandra Rosa, PhD,^{*,†} Patrícia Abrantes, PhD,^{†,‡} Inês Sousa, PhD,^{†,‡} Vânia Francisco, MSc,^{†,‡} Patrícia Santos, MSc,^{†,‡} David Francisco, MSc,^{†,‡} Joana M. Xavier, PhD,[§] and Sofia A. Oliveira, MD, PhD^{†,‡}

Background: Cellular oxidative stress and genetic susceptibility have been implicated in the multifactorial etiology of ulcerative colitis (UC). The nuclear genome association with UC has been intensely investigated, but the role of the mitochondrial DNA (mtDNA) has received far less attention and may account for part of the missing heritability. This study is a comprehensive analysis of the mtDNA contribution to UC susceptibility.

Methods: The association of mitochondrial single-nucleotide polymorphisms (mtSNPs) and haplogroups with UC was tested in 488 cases and 833 controls of European ancestry from the NIDDK IBD Genetics Consortium Ulcerative Colitis Genome-Wide Association Study available through dbGaP and from the Illumina Genotype Control Database (studies 64 and 65).

Results: No evidence of population stratification could be detected using 218 ancestry informative markers for European Americans. Seven of the 58 tested mtSNPs were nominally associated with UC, and A10550G in *MT-ND4L* withstands the Bonferroni correction ($P = 1.29E-06$, odds ratio [OR_G] [95% confidence interval (CI)] = 4.80 [2.54–9.05], 10550G allele: 8.1% of patients and 1.9% of controls). A10550G remains equally associated after conditional analyses on the 11 UC genome-wide association studies (GWAS) top SNPs ($6.35E-07 < P_{\text{cond}} < 4.58E-06$), which suggests that it constitutes an independent risk factor from nuclear-encoded susceptibility *loci*. We detected additive (but not multiplicative) epistatic interactions between A10550G and all 11 top GWAS hits. Subhaplogroup K1 ($P = 0.021$, OR [95% CI] = 1.71 [1.08–2.69]) increased the risk for UC, whereas the U5b lineage conferred protection ($P = 0.016$, OR [95% CI] = 0.34 [0.14–0.82]).

Conclusions: These results suggest that UC has a dual mitochondrial and nuclear genetic control that warrants further replication in independent data sets and reinforces its etiopathogenic complexity.

(*Inflamm Bowel Dis* 2016;22:774–781)

Key Words: ulcerative colitis, mitochondrial DNA, single-nucleotide polymorphism, genetic association studies, genetic predisposition to disease

Crohn's disease and ulcerative colitis (UC) are the most prevalent forms of inflammatory bowel disease (IBD), affecting over 2.5 million people of European ancestry.¹ UC is a chronic relapsing inflammatory condition characterized by mucosal ulcers in the rectum and colon. Several key players have been implicated

in its multifactorial etiology, including environmental factors (e.g., smoking, diet, drugs, geographical and social status, stress), commensal microbiota, epithelial barrier dysfunction, innate and adaptive immune responses, cellular oxidative stress, and genetic susceptibility.^{2–5}

Genome-wide association studies (GWAS) and meta-analyses have brought the most rapid novel insights into UC pathogenesis. Among the 163 IBD-associated *loci*, 110 are shared between UC and Crohn's disease, and the 23 UC-specific *loci* are primarily immune response mediators (*IL10*, *IL23R*, *IL26*, and *MHC* genes in chromosome 6p21) and genes involved in the epithelial barrier function (*ECM1*, *CDH1*, *HNF4α*, and *LAMB1*).⁵ However, these genetic risk factors collectively explain only approximately 20% to 25% of the heritability⁶ and less than 10% of the total disease variance,⁵ which suggests that additional genetic players (e.g., rare variants, extranuclear genome) must be identified.

Hypothesis-free GWAS have pinpointed common genetic risk factors in the nuclear-encoded genome, but the association of the second genome, the mitochondrial DNA (mtDNA), and of their interaction, has received far less attention. mtDNA variation has been associated with non-Mendelian, nonmaternally inherited, complex autoimmune,⁷ and inflammatory⁸ disorders. Although mitochondrial

Supplemental digital content is available for this article. Direct URL citations appear in the printed text and are provided in the HTML and PDF versions of this article on the journal's Web site (www.ibdjournal.org).

Received for publication September 25, 2015; Accepted November 19, 2015.

From the *Unidade de Ciências Médicas, Centro de Competências das Ciências da Vida, Universidade da Madeira, Funchal, Portugal; †Instituto de Medicina Molecular, Faculdade de Medicina, Universidade de Lisboa, Lisboa, Portugal; ‡Instituto Gulbenkian de Ciência, Oeiras, Portugal; and §Centre for Biomedical Research (CBMR), Universidade do Algarve, Faro, Portugal.

Supported by the Fundação para a Ciência e a Tecnologia (FCT-Portugal) through several grants (PTDC/IIM-GES/5015/2012 and CMUP-ERI/TPE/0028/2013) fellowships and research contracts (SFRH/BPD/35737/2007 to P. Abrantes, SFRH/BPD/70008/2010 to I. Sousa, and Ciência and Investigator-FCT contracts to S. A. Oliveira).

The authors have no conflicts of interest to disclose.

Reprints: Sofia A. Oliveira, MD, PhD, Instituto de Medicina Molecular, Av. Prof. Egas Moniz, Edifício Egas Moniz, 1649-028 Lisboa, Portugal (e-mail: aaoliveira@medicina.ulisboa.pt).

Copyright © 2016 Crohn's & Colitis Foundation of America, Inc.

DOI 10.1097/MIB.0000000000000694

Published online 1 March 2016.

single-nucleotide polymorphisms (mtSNPs) are represented in most genotyping platforms used in GWAS, their analysis has been neglected or underreported.⁵

Mitochondria are the main intracellular source of reactive oxygen species (ROS) because their primary function is production of ATP through oxidative phosphorylation in the electron transport chain. High levels of ROS and oxidative stress have been implicated in the pathophysiology of IBD,^{3,4} while the Krebs cycle and oxidative phosphorylation in the mitochondria are among the few downregulated pathways in active UC biopsies.⁹ These organelles also play a central role in intracellular signaling, apoptosis, and metabolic pathways such as tricarboxylic acid cycle and the metabolism of amino acids, lipids, cholesterol, and steroids.

The human mtDNA is a haploid, nonrecombining circular genome of approximately 16,600 nucleotides encoding for 13 electron transport chain polypeptides, 22 transfer RNAs, and 2 ribosomal RNAs. Particular combinations of mtSNPs define haplogroups and their subclades that tend to be associated with broad geographic areas and/or populations.¹⁰ Although presumed to be selectively neutral, mtSNPs of phylogenetic relevance and especially those in the mtDNA-coding region are likely to assume a functional role on the expression of some complex traits. mtDNA polymorphisms and haplogroups may act independently or synergistically with nuclear genetic factors and mtDNA-encoded gene/nuclear-encoded gene interactions may explain part of the UC genetic vacuum.

To determine whether the mitochondrial genome may account for part of the missing heritability in UC, in this study we tested the association of mtSNPs and haplogroups in 1321 UC cases and healthy controls of European ancestry from the NIDDK IBD Genetics Consortium Ulcerative Colitis (NIDDK IBD UC) GWAS¹¹ available through the NCBI Database of Genotypes and Phenotypes (dbGaP),¹² and from Illumina Genotype Control Database (iControlDB) studies, controlling for population stratification with the analysis of ancestry informative markers (AIMs).

MATERIALS AND METHODS

Study Subjects and Genotyping

The criteria and procedures for the recruitment of patients with UC are described in detail elsewhere.¹¹ Briefly, cases of self-reported white, European non-Hispanic ancestry, were ascertained in North American institutions and were selected to have either left-sided or extensive UC (i.e., patients only with proctitis were excluded). Genotyping was performed at the Feinstein Institute for Medical Research with the Illumina HumanHap550v3 Genotyping BeadChips (detailed methodology described in Silverberg et al¹¹). The phenotype and genotype data investigated in this study is deposited in the Database of Genotypes and Phenotypes (dbGaP)¹² under the dbGaP study accession number phs000345.v1.p1 (<http://www.ncbi.nlm.nih.gov/sites/entrez?db=gap>). Control data were

obtained from studies 64 and 65 ($n = 833$) deposited in the Illumina iControlDB at <http://genomeinformaticsalliance.org/science/icontrolldb.ilmn>. A χ^2 test was used to compare the sex distribution between UC cases and controls.

Population Stratification Assessment

Among the 300 AIMs described to be most informative for discriminating northwest and southeast European ancestries among European Americans and which may be used to correct for population substructure (<http://genepath.med.harvard.edu/~reich/EUROSNP.htm>),¹³ 156 were directly genotyped in the UC data set as they were included in the Illumina HumanHap550v3 microarrays. Sixty-two ungenotyped AIMs were replaced by a proxy SNP available in the array and identified using the SNP Annotation and Proxy Search (SNAP) tool (<http://www.broadinstitute.org/mpg/snap/>)¹⁴ ($r^2 = 1$ in the CEU population panel and the 1000 Genomes Pilot 1 SNP data set, maximum distance 500 kb), and the remaining 82 AIMs could not be confidently substituted. Of the original 300 AIMs, a panel of 218 AIMs (see Table, Supplemental Digital Content 1, <http://links.lww.com/IBD/B189>) was therefore used for population stratification of this data set.

Principal component analysis (PCA), F_{ST} statistics, and the inflation factor λ were computed using EIGENSOFT 3.0, with default parameters. PCAs were computed for 488 patients with UC and 833 controls with or without HapMap reference samples from 4 populations (174 CEU—US Utah residents with Northern and Western European ancestry; 176 YRI—Yoruba people of Ibadan, Nigeria; 86 CHB—individuals from Beijing, China; 89 JPT—individuals from Tokyo, Japan). The genotypes for these 525 HapMap individuals at the 218 AIMs were downloaded from the International HapMap Project Web site (<http://hapmap.ncbi.nlm.nih.gov/biomart/martview>, Release 27).¹⁵ F_{ST} values indicate which proportion of the total genetic variance is attributable to the differences between populations and to the differences between individuals of the same population. This measure of genetic structure varies between 0 (no genetic differentiation of population units) and 1 (complete differentiation of population units). The λ inflation factor indicates population stratification when $\lambda > 1.000$ and is used to correct the chi-square association tests when stratification exists.¹⁶

mtDNA Haplogroup Classification

Given the nonrecombining nature of mtDNA, haplogroups are defined by the haplotypic combination of multiple mtSNPs throughout the molecule, rather than the status at any single point mutation. Among the genotyped mtSNPs, and considering the updated human mtDNA phylogeny in PhyloTree (build 16; <http://phylotree.org/>),¹⁷ we selected the SNPs with phylogenetic relevance to classify our samples into the most prevalent haplogroups in North American population of European ancestry: West Eurasian haplogroups R0/HV, H, V, J, T, U, K, N1, I, and W; Eastern Eurasian and Native American haplogroups A, B2, C, and D4;

Near Eastern and Central Asian haplogroups M and Z; and African haplogroup L. The nomenclature of clades follows that proposed by Torroni et al,¹⁸ Macaulay et al,¹⁹ and Richards et al²⁰. Sixteen samples were excluded from the mtDNA haplogroup analysis because it was not possible to classify them within the above-mentioned haplogroups.

Association Analyses

For the mtSNPs analysis, a quality filtering was applied to consider only those with an SNP call rate $\geq 95\%$ and minor allele frequency (MAF) $\geq 1\%$. Heterozygote genotypes were considered as missing. For the haplogroup classification, all mtSNP genotypes available were used because rare variants may be of phylogenetic relevance. For instance, the A5656G mtSNP did not pass the quality filtering, but its single occurrence in haplogroup U phylogeny in PhyloTree proved useful for classifying U5b1 samples (further analyzed in the U5b cluster).

For haplogroups' analyses, we compared each haplogroup with all others pooled together. Rare haplogroups (<2%) were included in the "others" category (haplogroups A, B2, C, D4, N1, L, M, and Z). Because no phylogeographic reasoning exists for these clades, the "others" category was not considered for the association analysis. To adjust the association analyses for relevant confounding factors, sex was included as covariate in a binomial logistic regression (log-additive model) computed with the R freeware version 3.1.2.²¹ Pairwise conditional analyses, additive, and epistatic interactions between mtSNPs and top GWAS nuclear SNPs were performed sex-adjusted by logistic regression (log-additive model) in R.

Results were considered nominally significant below the conventional level of 0.05. Because some mtSNPs are in linkage disequilibrium and the haplogroup comparisons are not independent, no corrections for multiple testing were performed in both SNP and haplogroup association analyses, and uncorrected *P* values are reported.

ETHICAL CONSIDERATIONS

This study was approved by the ethics committee of the Centro Hospitalar Lisboa Norte, E.P.E./Faculdade de Medicina de Lisboa, Lisboa, Portugal.

RESULTS

Characterization of the Dataset

The NIDDK IBD UC GWAS¹¹ data available through dbGaP (study accession number phs000345.v1.p1) consists of phenotypes and genotypes of 488 patients obtained on the Illumina HumanHap550v3 Genotyping BeadChips (which include 156 mtSNPs) and are the focus of the work presented here. The remaining data for 540 UC cases genotyped on the Illumina HumanHap300v2 BeadChips could not be further investigated here because mtSNPs were not represented in these arrays. Control data (*N* = 833) was obtained from the Illumina iControlDB

database. No information was available regarding the age at examination (AAE) of the cases, but the controls' average AAE \pm SD was 48.2 ± 11.3 years (*n* = 798). The 2 groups differ significantly (*P* = 4.09E-33) in the sex distribution (49.2% and 18.0% of males in the cases and controls, respectively), and because UC seems to occur more frequently in women than in men,¹ the association analyses reported in this study are adjusted for sex (very similar results were obtained in unadjusted association tests, data not shown).

Population Stratification Assessment

The effect of population stratification is particularly problematic in mtDNA studies owing to its smaller effective population size when compared with autosomal markers.²² Because we investigated the mtDNA association on part of the data set used in the NIDDK IBD UC GWAS, we assessed whether a hidden ancestry and/or geographic substructure was present in this subset of samples using genotypes at 218 biallelic AIMs (see Table, Supplemental Digital Content 1, <http://links.lww.com/IBD/B189>).

To validate the intercontinental resolution of these 218 AIMs, we performed a PCA with patients with UC and controls plus HapMap CEU, CHB, JPT, and YRI individuals (Fig. 1A). HapMap populations cluster into distinct groups (except CHB and JPT samples which overlap; $F_{ST}[\text{CHB, JPT}] = 0.006$), with significant differences between populations (*P* < 2.00E-09, pairwise $F_{ST} = 0.186\text{--}0.243$). UC samples are in closer proximity to the CEU cluster ($F_{ST}[\text{patients with UC, CEU}] = 0.003$, $F_{ST}[\text{UC controls, CEU}] = 0.004$) than to any other HapMap group (pairwise $F_{ST} = 0.167\text{--}0.218$), and again the distribution of UC cases and controls overlaps and no genetic differentiation is detected among them ($F_{ST}[\text{patients with UC, UC controls}] = 0.000$).

Figure 1B depicts the top 2 principal components of the PCA analysis performed with the patients with UC and controls only, which shows in greater detail that these 2 groups seem to overlap, with no major outliers. Analysis of variance for population differences along the eigenvectors demonstrated that UC cases and controls did not differ with regard to the first 2 principal components, nor did they differ in the overall PCA test findings (*P* = 0.810). The inflation factor ($\lambda = 1.000$) and F_{ST} index ($F_{ST}[\text{patients with UC, UC controls}] = 0.000$) support the inexistence of population stratification and genetic differentiation, respectively, between these subsets, and therefore adjustment for ancestry in association tests is not required.

SNP Association Analyses

We first tested the association of top NIDDK IBD UC GWAS findings (rs3806308, rs10753575, rs6426833, rs1004819, rs11465804, rs11209026, rs10889677, rs2395185, rs7134599, rs1558744, rs2870946) in this subgroup of samples. These autosomal SNPs were associated with UC with $1.36\text{E-}05 \leq P \leq 0.019$ (see Table, Supplemental Digital Content 2, <http://links.lww.com/IBD/B190>), therefore validating this subset. As expected, the associations were not as significant as originally

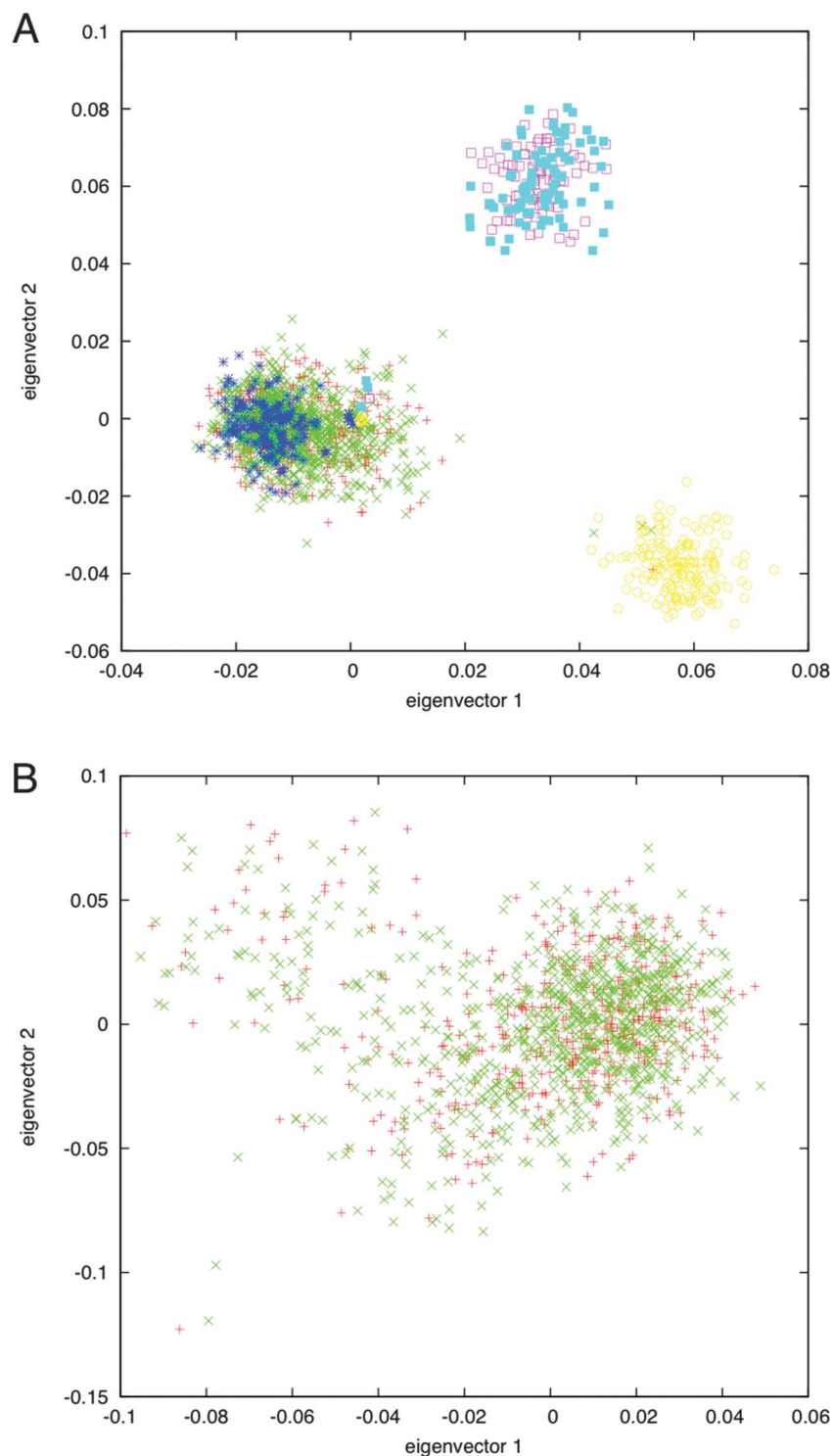


FIGURE 1. Population stratification analysis. The top 2 axes of genetic variation, determined as eigenvectors in principal components analysis (PCA) of genotypes at 218 ancestry-informative markers in 488 patients with UC (red symbols) and 833 controls (green symbols) with (A) or without (B) HapMap reference samples (174 CEU—US Utah residents with Northern and Western Europe ancestry; 176 YRI—Yoruba people of Ibadan, Nigeria; 86 CHB—individuals from Beijing, China; 89 JPT—individuals from Tokyo, Japan). CEU, CHB, JPT, and YRI HapMap samples are shown in dark blue, pink, light blue, and yellow, respectively.

Downloaded from <https://academic.oup.com/ibdjournal/article/22/4/774/4561709> by B-On Consortium Portugal user on 03 March 2022

reported ($5.5E-10 \leq P \leq 3.9E-04$),¹¹ most likely due to the loss of power.

Of the 156 mtSNPs genotyped in our data set, 58 passed quality control and were further analyzed. The mean genotype call rate was 98.1% and 97.6% in patients and controls, respectively. Table 1 shows the association results of mtSNPs nominally associated with UC risk that were analyzed using a sex-adjusted logistic regression (log-additive model), and Supplemental Digital Content 3 (see Table, Supplemental Digital Content 3, <http://links.lww.com/IBD/B191>) presents the association results at all 58 mtSNPs tested. Seven mtSNPs were significantly associated with UC, and A10550G survives the conservative Bonferroni correction for multiple testing ($P \leq 8.62E-04$). The 10550G allele was observed in 8.1% of patients with UC and 1.9% of controls ($P = 1.29E-06$, odds ratio [OR] [95% confidence interval (CI)] = 4.80 [2.54–9.05]). This mtSNP is equally associated ($6.35E-07 \leq P_{\text{cond}} \leq 4.58E-06$) after pairwise conditional analyses on the 11 top NIDDK IBD UC GWAS SNPs when adjusted for sex (see Table, Supplemental Digital Content 4, <http://links.lww.com/IBD/B192>). Association of top UC GWAS SNPs was also not affected ($1.09E-05 \leq P_{\text{cond}} \leq 0.033$) by conditional analysis on A10550G (see Table, Supplemental Digital Content 4, <http://links.lww.com/IBD/B192>), which suggests that the mtSNP A10550G constitutes an independent risk factor from nuclear-encoded susceptibility loci.

Because the interplay between multiple independent genetic variants may be important in conferring disease susceptibility to UC, we tested whether the A10550G mtSNP genetically interacts with each of the top 11 nuclear-encoded GWAS SNPs. Epistasis was explored using an additive interaction model that suggests the existence of multiple genes acting in parallel, and a multiplicative interaction model that addresses synergistic effects among genes. Additive effects were specifically tested using sex-adjusted logistic regression analyses for a score that ranges from 0 to 3 and is the sum of the number of risk alleles each individual has at A10550G (number of G alleles that is 0 or 1) and at the nuclear SNP being tested (ranges from 0 to 2). We detected additive interactions between A10550G and the top 11 nuclear SNPs because the ORs were greater and the P values were

lower in the additive model (see Table, Supplemental Digital Content 5, <http://links.lww.com/IBD/B193>) when compared with their individual association (see Table, Supplemental Digital Content 2, <http://links.lww.com/IBD/B190>). The most significant additive effect identified was between the risk (G) alleles at A10550G and rs2870946 in *IL26* ($P_{\text{add}} = 1.99E-09$, OR [95% CI] = 2.49 [1.85–3.35]) (see Table, Supplemental Digital Content 5, <http://links.lww.com/IBD/B193>). For this particular SNP, the number (and percentage) of controls and cases with 0, 1, 2, and 3 risk alleles were 695 (88.3%), 88 (11.2%), 4 (0.5%), and 0 (0%); and 349 (74.1%), 114 (24.2%), 8 (1.7%), and 0 (0%), respectively. As tested using a linear logistic regression model ($\log\text{Odds} = \beta_0 + A10550G + \text{nuclear SNP} + A10550G * \text{nuclear SNP} + \text{sex}$), no multiplicative effects ($P_{\text{mult}} \geq 0.395$) were detected between the A10550G mtSNP and any of the top 11 GWAS SNPs (see Table, Supplemental Digital Content 5, <http://links.lww.com/IBD/B193>).

Haplogroup Association Analyses

Figure 2 depicts the simplified phylogenetic tree used for haplogroup classification of individuals in our data set. The most common European mtDNA haplogroups were observed in our data set of UC cases and controls, and their frequencies in the control data set are in agreement with those previously reported for the European populations.¹⁸ Table 2 shows the sex-adjusted haplogroup association results with UC. Unlike other subclades, U5a, U5b, and K1 were tested for association because missing calls did not prevent U and K individuals, respectively, to be further subclassified. The K1 subcluster was significantly ($P = 0.021$, OR [95% CI] = 1.71 [1.08–2.69]) more frequent in patients with UC than in controls (8.8% versus 5.6%, respectively). Additionally, the U5b lineage seems to confer protection against UC (1.4% in patients with UC versus 3.6% in controls; $P = 0.016$, OR [95% CI] = 0.34 [0.14–0.82]), in accordance with its defining A7768G mtSNP ($P = 1.03E-02$, OR [95% CI] = 0.30 [0.12–0.75]). The K haplogroup was marginally associated ($P = 0.051$) despite the nominal ($P = 0.030$) and strong ($P = 1.29E-06$) association of it defining A3480G and A10550G mtSNPs, respectively. Because haplogroups are defined by a particular combination of

TABLE 1. Mitochondrial Single Nucleotide Polymorphisms (mtSNPs) Associated with Ulcerative Colitis Risk

mtSNP	dbSNP ID	Gene	Variant Type	Allele	Patients (n/N, %)	Controls (n/N, %)	P	OR (95% CI)
T250C	rs41400048	—	—	C	22/486 (4.5)	18/826 (2.2)	0.036	2.05 (1.05–4.01)
A3480G	rs28358584	<i>ND1</i>	Synonymous	G	49/479 (10.2)	53/819 (6.5)	0.030	1.61 (1.05–2.48)
A7768G	rs41534044	<i>CO2</i>	Synonymous	G	6/484 (1.2)	29/832 (3.5)	0.010	0.30 (0.12–0.75)
T10034C	rs41347846	<i>TG</i>	—	C	22/484 (4.5)	18/827 (2.2)	0.035	2.06 (1.05–4.03)
G10398A	rs2853826	<i>ND3</i>	Thr114Ala	A	358/484 (74.0)	661/831 (79.5)	0.008	0.68 (0.52–0.90)
A10550G	rs28358280	<i>ND4L</i>	Synonymous	G	38/471 (8.1)	15/787 (1.9)	1.29E-06	4.80 (2.54–9.05)
C10873T	rs2857284	<i>ND4</i>	Synonymous	T	454/468 (97.0)	815/827 (98.5)	0.027	0.40 (0.18–0.90)

Sex-adjusted P values and their respective OR and 95% CI are shown only for nominally significant associations. $P < 0.05$.

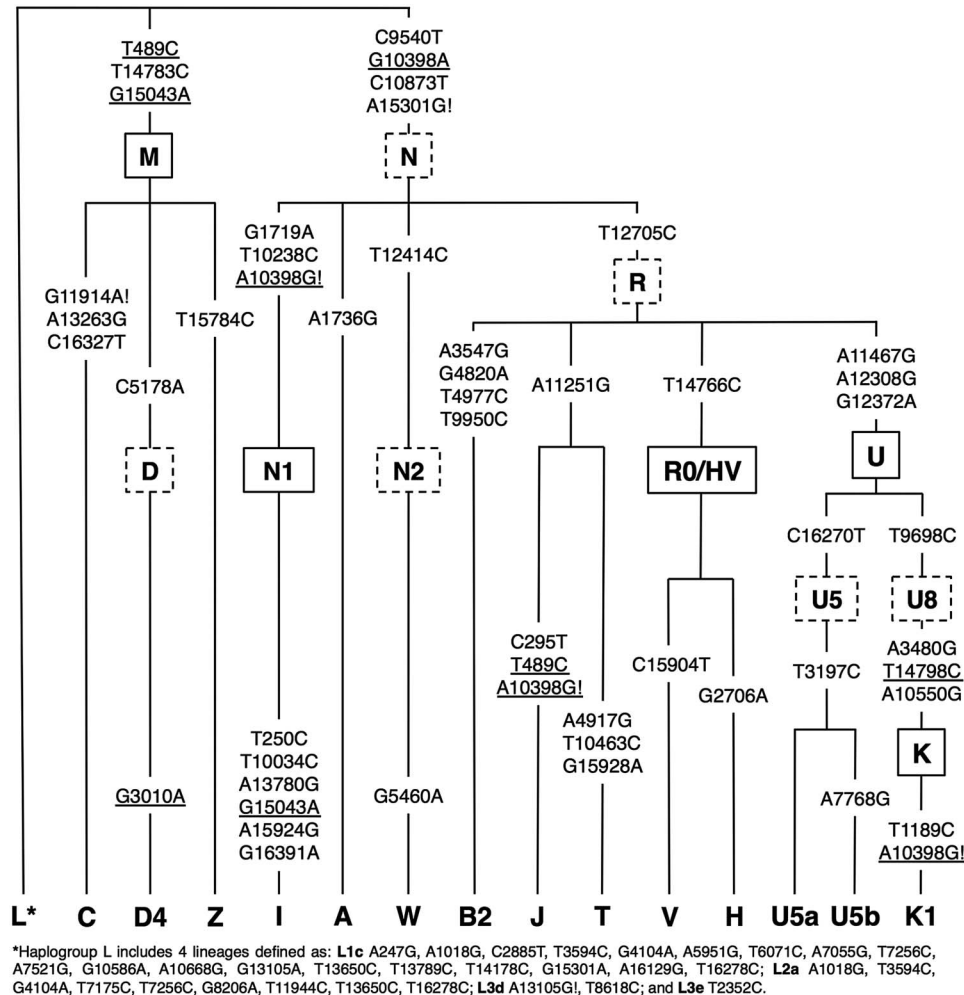


FIGURE 2. Simplified phylogenetic tree depicting the mtSNPs used for haplogroup classification. mtSNPs are named after their base pair position in the mitochondrial reference sequence (NC_001807) and alleles are given in forward evolutionary time direction ([ancestral base] [position number] [derived base], according to www.phyloree.org). Mutations that are reversions to an ancestral state (back mutations) are indicated with an exclamation mark. mtSNPs found to be homoplasic in our data set are underlined in this tree. Full boxes indicate phylogenetic clades found in our data set, whereas dashed boxes represent intermediate nodes.

alleles at several defining SNPs, this apparent inconsistency may result from the synergistic effect of other haplogroup-defining SNPs over A3480G and A10550G, and/or from the protective 10398A allele in haplogroup K and its state reversion to 10398G in K1 (Fig. 2). Furthermore, it could be partially explained by a higher number of control individuals having the K haplogroup (58/833) than the 3480G (53/819) or the 10550G (15/787) alleles, because of missing calls (Tables 1 and 2). Some of the 58 control individuals carrying the K haplogroup had missing calls in A3480G or A10550G, but all of them were classified in the K haplogroup as their genotypes at other SNPs (e.g., A11467G, A12308G, G12372A, T9698C, and T14798C) supported this classification.

CONCLUSIONS

Since the colon of patients with UC is characterized by an energy deficiency state and elevated ROS,^{3,4,9} mitochondria may

contribute to the pathogenic process because of their pivotal role in cellular energy metabolism and ROS generation. Given that ATP depletion compromises the regeneration and barrier function of the colonic mucosa^{23,24} and high ROS levels help to sustain the chronic inflammation, certain mtSNPs and haplogroups may influence the susceptibility to UC. We specifically tested whether the mtDNA genome contributes to the missing heritability in UC in a subset of samples with no detectable population stratification and in which previously reported top GWAS findings were validated. The chief finding of this association study was that the A10550G variant in *MT-ND4L* increases UC risk by approximately 4.8-fold, which is higher than the increase in the risk conferred by any single nuclear genome-encoded variant associated thus far with UC. This supports the notion that UC has a dual (nuclear and mitochondrial) genetic control.

MT-ND4L encodes the core subunit of the respiratory chain NADH dehydrogenase (complex I) and is believed to belong to

TABLE 2. Results of Mitochondrial Haplogroup Association Testing with Ulcerative Colitis Risk

Haplogroup Subhaplogroup	Patients (N, %)	Controls (N, %)	<i>P</i>	OR (95% CI)
H	190 (38.9)	327 (39.3)	0.771	
R0/HV	23 (4.7)	41 (4.9)	0.770	
V	10 (2.0)	31 (3.7)	0.091	
J	46 (9.4)	88 (10.6)	0.942	
T	49 (10.0)	89 (10.7)	0.633	
U	53 (10.9)	105 (12.6)	0.376	
U5a	19 (3.9)	33 (4.0)	0.963	
U5b	7 (1.4)	30 (3.6)	0.016	0.34 (0.14–0.82)
K	50 (10.2)	58 (7.0)	0.051	
K1	43 (8.8)	47 (5.6)	0.021	1.71 (1.08–2.69)
W	8 (1.6)	21 (2.5)	0.266	
I	22 (4.5)	23 (2.8)	0.099	
Others ^a	29 (5.9)	42 (5.0)	—	
Undetermined	8 (1.6)	8 (1.0)	—	

Only clades with frequency above 2% were tested for association. Sex-adjusted *P* values are shown for all haplogroups but OR and 95% CI are highlighted in bold only for nominally significant associations.

^aIncludes minor clades A, B2, C, D4, N1, L, M, and Z. Because this class includes an assortment of lineages not phylogenetically related, association tests were not performed.

the minimal assembly for the transfer of electrons from NADH to ubiquinone. A reduced activity in mitochondrial respiratory chain complex II has been observed in the colonic mucosa of patients with UC and UC-induced mice models, whereas the activity of other complexes has not been consistently described as altered.^{3,25,26} Synonymous SNPs such as A10550G (Met27Met) are assumed not to affect protein function because no change occurs in the amino acid composition. However, naturally occurring synonymous codon substitutions in the nuclear genome have been shown to alter the *in vivo* translation kinetics and thus protein folding and function.²⁷ Nonsynonymous mutations and SNPs in *MT-ND4L* have been associated with several diseases, including colorectal cancer and body mass index.^{28,29} Given that we only assessed the association of 58 mitochondrial SNPs with UC, it is conceivable that the associated SNP is a proxy for another functional genetic variation in strong linkage disequilibrium that was not tested.

Hudson et al³⁰ probed the mtDNA association with eleven common late-onset diseases (ankylosing spondylitis, ischemic stroke, multiple sclerosis, Parkinson's disease, primary biliary cirrhosis, psoriasis, schizophrenia, coronary artery disease, hypertension, type-2 diabetes, and UC) investigated under the Wellcome Trust Case Control Consortium (WTCCC). Among these phenotypes, UC was the disease with the largest number of nominal mtDNA associations (23 mtSNPs), 5 of which were from effectively genotyped SNPs and the remaining were imputed,³⁰

but none of these associations overlapped with those described herein. Several factors confound the comparison of these 2 studies, starting with the choice of commercial arrays (Affymetrix 6.0 and Illumina HumanHap550v3) with mostly different mitochondrial polymorphisms. Although the WTCCC data set is larger (2855 cases and 5033 controls after quality control), almost two-thirds of the SNPs tested for association with UC in the WTCCC data set were imputed³⁰ and not effectively genotyped as in the NIDDK data set. Additionally, population stratification in the WTCCC study was assessed by PCA performed with mtSNPs,³⁰ not with nuclear-encoded SNPs, and the numerous clusters observed most likely correspond to the major European haplogroups.³¹ Therefore, the reported lack of population stratification among cases and controls using this method in the WTCCC data set is consistent with the modest association of relatively rare mitochondrial subhaplogroups with UC in the NIDDK data set, but does not exclude the possibility that mtSNP associations in the WTCCC data set are confounded by the geographical origin, which is known to correlate well with autosomal principal components.³¹ Future studies are therefore warranted to further validate and fine-map these associations in other populations and to subsequently determine how this variation contributes to UC etiopathogenesis.

Genetic polymorphisms in nuclear-encoded genes involved in mitochondrial energy metabolism (e.g., *OCTN*, *UCP2*, *DLD*, and *NF-κB*) have been associated with UC.^{5,32,33} *OCTN* and *UCP2* have been implicated in intestinal barrier function *in vivo*,^{34,35} and *NF-κB* is known to regulate epithelial cell proliferation,^{24,36} 2 biological processes disrupted in patients with UC. In our study, we did not test genetic interactions between the mitochondrial genome and the above-mentioned genes because they were not the top findings in the GWAS under investigation, but we detected an epistatic effect between A10550G and the 11 top SNPs in the NIDDK IBD UC GWAS. The genetic interaction was of an additive and not of a multiplicative nature, as expected, because the functions of *ND4L* and of the nuclear genes to which these polymorphisms map (e.g., *IL23R*, *IL26*) do not seem to be correlated.

Gene expression and proteomic characterization of active UC biopsies have reported significant downexpression of citrate cycle and respiratory chain mitochondrial proteins⁹ and show high levels of a number of oxidative stress response proteins (e.g., selenium-binding protein, SOD, and thioredoxin-dependant peroxide reductase) and energy-generation proteins (e.g., isocitrate dehydrogenase, L-lactate dehydrogenase B-chain, inorganic phosphatase, and enoyl-CoA hydratase).³⁷ Because mitochondrial-related pathways were not enriched among UC GWAS hits, Cardinale et al⁹ suggested that perturbation of the mitochondrial respiration and ATP production are not etiopathogenic drivers but rather secondary to inflammation. It is however important to note that the mitochondrial genome was not investigated in this meta-analysis of UC GWAS scans,⁵ therefore underscoring the putative contribution of 13 mtDNA-encoded polypeptides related to oxidative phosphorylation. The importance of oxidative stress contribution to UC is further supported by the successful use of therapeutic drugs in patients

and animal models that, directly or indirectly, scavenge free radicals, increase antioxidative capacity of cells, and inhibit prooxidative enzymes.^{38,39} Finally, it would be of interest to develop a more comprehensive line of research regarding the role of mtDNA variation in UC risk (e.g., full mtDNA sequencing, heteroplasmy levels) and to test the interactions between the mtSNPs, polymorphisms in the nuclear genome and environmental factors previously reported for UC. Deeper understanding of the cellular and genetic contributors to oxidative damage in UC may expand the existing therapeutic targets toward cellular redox homeostasis.

ACKNOWLEDGMENTS

The NIDDK IBD Genetics Consortium Ulcerative Colitis (NIDDK IBD UC) Genome-Wide Association Study was conducted by the NIDDK IBD UC Investigators and supported by the National Institute of Diabetes and Digestive and Kidney Diseases (NIDDK). The data from the NIDDK IBD UC reported here were supplied by the NIDDK Central Repositories. This manuscript was not prepared in collaboration with Investigators of the NIDDK IBD UC study and does not necessarily reflect the opinions or views of the NIDDK IBD UC study, the NIDDK Central Repositories, or the NIDDK. The authors sincerely thank the NIDDK IBD UC Investigators for sharing their data through dbGaP.

REFERENCES

- Molodecky NA, Soon IS, Rabi DM, et al. Increasing incidence and prevalence of the inflammatory bowel diseases with time, based on systematic review. *Gastroenterology*. 2012;142:46–54.
- Danese S, Sans M, Fiocchi C. Inflammatory bowel disease: the role of environmental factors. *Autoimmun Rev*. 2004;3:394–400.
- Jena G, Trivedi PP, Sandala B. Oxidative stress in ulcerative colitis: an old concept but a new concern. *Free Radic Res*. 2012;46:1339–1345.
- Zhu H, Li YR. Oxidative stress and redox signaling mechanisms of inflammatory bowel disease: updated experimental and clinical evidence. *Exp Biol Med*. 2012;237:474–480.
- Jostins L, Ripke S, Weersma RK, et al. Host-microbe interactions have shaped the genetic architecture of inflammatory bowel disease. *Nature*. 2012;491:119–124.
- Zhang YZ, Li YY. Inflammatory bowel disease: pathogenesis. *World J Gastroenterol*. 2014;20:91–99.
- Xavier JM, Shafiee NM, Ghaderi F, et al. Association of mitochondrial polymorphism m.709G>A with Behçet's disease. *Ann Rheum Dis*. 2011;70:1514–1516.
- Canter JA, Olson LM, Spencer K, et al. Mitochondrial DNA polymorphism A4917G is independently associated with age-related macular degeneration. *PLoS One*. 2008;3:e2091.
- Cardinale CJ, Wei Z, Li J, et al. Transcriptome profiling of human ulcerative colitis mucosa reveals altered expression of pathways enriched in genetic susceptibility loci. *PLoS One*. 2014;9:e96153.
- Kivisild T, Shen P, Wall DP, et al. The role of selection in the evolution of human mitochondrial genomes. *Genetics*. 2006;172:373–387.
- Silverberg MS, Cho JH, Rioux JD, et al. Ulcerative colitis-risk loci on chromosomes 1p36 and 12q15 found by genome-wide association study. *Nat Genet*. 2009;41:216–220.
- Tryka KA, Hao L, Sturcke A, et al. NCBI's database of genotypes and phenotypes: dbGaP. *Nucleic Acids Res*. 2014;42:D975–D979.
- Price AL, Butler J, Patterson N, et al. Discerning the ancestry of European Americans in genetic association studies. *Plos Genet*. 2008;4:e236.
- Johnson AD, Handsaker RE, Pulit S, et al. SNAP: a web-based tool for identification and annotation of proxy SNPs using HapMap. *Bioinformatics*. 2008;24:2938–2939.
- The International HapMap Consortium. Integrating common and rare genetic variation in diverse human populations. *Nature*. 2010;467:52–58.
- Devlin B, Roeder K. Genomic control for association studies. *Biometrics*. 1999;55:997–1004.
- van Oven M, Kayser M. Updated comprehensive phylogenetic tree of global human mitochondrial DNA variation. *Hum Mutat*. 2009;30:E386–E394.
- Torroni A, Huoponen K, Francalacci P, et al. Classification of European mtDNAs from an analysis of three European populations. *Genetics*. 1996;144:1835–1850.
- Macaulay V, Richards M, Hickey E, et al. The emerging tree of West Eurasian mtDNAs: a synthesis of control-region sequences and RFLPs. *Am J Hum Genet*. 1999;64:232–249.
- Richards M, Macaulay V, Hickey E, et al. Tracing European founder lineages in the Near Eastern mtDNA pool. *Am J Hum Genet*. 2000;67:1251–1276.
- R Core Team. *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing; 2014. Available at: <http://www.R-project.org/>. Accessed January 1, 2014.
- Salas A, Fachal L, Marcos-Alonso S, et al. Grupo de investigación ESIGEM (Estudio Sobre la Influencia Genética en la Enfermedad Meningocócica). Investigating the role of mitochondrial haplogroups in genetic predisposition to meningococcal disease. *PLoS One*. 2009;4:e8347.
- Lewis K, Caldwell J, Phan V, et al. Decreased epithelial barrier function evoked by exposure to metabolic stress and nonpathogenic *E. coli* is enhanced by TNF-alpha. *Am J Physiol Gastrointest Liver Physiol*. 2008;294:G669–G678.
- Bär F, Bochmann W, Widok A, et al. Mitochondrial gene polymorphisms that protect mice from colitis. *Gastroenterology*. 2013;145:1055–1063.
- Sifroni KG, Damiani CR, Stoffel C, et al. Mitochondrial respiratory chain in the colonic mucosal of patients with ulcerative colitis. *Mol Cell Biochem*. 2010;342:111–115.
- Santhanam S, Rajamanickam S, Motamarry A, et al. Mitochondrial electron transport chain complex dysfunction in the colonic mucosa in ulcerative colitis. *Inflamm Bowel Dis*. 2012;18:2158–2168.
- Kimchi-Sarfaty C, Oh JM, Kim IW, et al. A “silent” polymorphism in the MDR1 gene changes substrate specificity. *Science*. 2007;315:525–528.
- Polyak K, Li Y, Zhu H, et al. Somatic mutations of the mitochondrial genome in human colorectal tumours. *Nat Genet*. 1998;20:291–293.
- Flaquer A, Baumbach C, Kriebel J, et al. Mitochondrial genetic variants identified to be associated with BMI in adults. *PLoS One*. 2014;9:e105116.
- Hudson G, Gomez-Duran A, Wilson IJ, et al. Recent mitochondrial DNA mutations increase the risk of developing common late-onset human diseases. *Plos Genet*. 2014;10:e1004369.
- Biffi A, Anderson CD, Nalls MA, et al. Principal-component analysis for assessment of population stratification in mitochondrial medical genetics. *Am J Hum Genet*. 2010;86:904–917.
- Waller S, Tremelling M, Bredin F, et al. Evidence for association of OCTN genes and IBD5 with ulcerative colitis. *Gut*. 2006;55:809–814.
- Yu X, Wieczorek S, Franke A, et al. Association of UCP2 -866 G/A polymorphism with chronic inflammatory diseases. *Genes Immun*. 2009;10:601–605.
- Shekhawat PS, Srinivas SR, Matern D, et al. Spontaneous development of intestinal and colonic atrophy and inflammation in the carnitine-deficient *jvs* (OCTN2(-/-)) mice. *Mol Genet Metab*. 2007;92:315–324.
- Zhang H, Kuai XY, Yu P, et al. Protective role of uncoupling protein-2 against dextran sodium sulfate-induced colitis. *J Gastroenterol Hepatol*. 2012;27:603–608.
- Eckmann L, Neish AS. NF-kappaB and mucosal homeostasis. *Curr Top Microbiol Immunol*. 2011;349:145–158.
- Poulsen NA, Andersen V, Moller JC, et al. Comparative analysis of inflamed and non-inflamed colon biopsies reveals strong proteomic inflammation profile in patients with ulcerative colitis. *BMC Gastroenterol*. 2012;12:76.
- Damiani CR, Benetton CA, Stoffel C, et al. Oxidative stress and metabolism in animal model of colitis induced by dextran sulfate sodium. *J Gastroenterol Hepatol*. 2007;22:1846–1851.
- Piechota-Polanczyk A, Fichna J. The role of oxidative stress in pathogenesis and treatment of inflammatory bowel diseases. *Naunyn Schmiedebergs Arch Pharmacol*. 2014;387:605–620.