OPEN ACCESS

EPIDEMIOLOGICAL SCIENCE

A meta-analysis and a functional study support the influence of mtDNA variant m.16519C on the risk of rapid progression of knee osteoarthritis

Alejandro Durán-Sotuela,[1] Mercedes Fernandez-Moreno,[1] Victoria Suárez-Ulloa,[2] Jorge Vázquez-García,[1] Sara Relaño,[1] Tamara Hermida-Gómez,[1,3] Vanesa Balboa-Barreiro,[4] Lucia Lourido-Salas,[1] Valentina Calamia,[1] Patricia Fernandez-Puente,[1] Cristina Ruiz-Romero ,[1,3] Juan Fernández-Tajes,[1] Carlos Vaamonde-García,[1] María C de Andrés,[1] Natividad Oreiro,[1,3] Francisco J Blanco ,[1,5] Ignacio Rego-Perez [1]

Handling editor Josef S Smolen

► Additional supplemental material is published online only. To view, please visit the journal online (http://dx.doi.org/10.1136/ard-2022-223570).

For numbered affiliations see end of article.

Correspondence to
Dr Ignacio Rego-Perez, Grupo de Investigación en Reumatología (GIR), Complexo Hospitalario Universitario de A Coruña (CHUAC), Sergas, Universidade da Coruña (UDC), Instituto de Investigación Biomédica de A Coruña, A Coruna, Galicia, Spain; Ignacio.Rego.Perez@sergas.es

FJB and IR-P contributed equally.

Received 2 November 2022
Accepted 17 March 2023

To cite: Durán-Sotuela A, Fernandez-Moreno M, Suárez-Ulloa V, *et al*. *Ann Rheum Dis* Epub ahead of print: [*please include* Day Month Year]. doi:10.1136/ard-2022-223570

ABSTRACT

Objectives To identify mitochondrial DNA (mtDNA) genetic variants associated with the risk of rapid progression of knee osteoarthritis (OA) and to characterise their functional significance using a cellular model of transmitochondrial cybrids.

Methods Three prospective cohorts contributed participants. The osteoarthritis initiative (OAI) included 1095 subjects, the Cohort Hip and Cohort Knee included 373 and 326 came from the PROspective Cohort of Osteoarthritis from A Coruña. mtDNA variants were screened in an initial subset of 450 subjects from the OAI by in-depth sequencing of mtDNA. A meta-analysis of the three cohorts was performed. A model of cybrids was constructed to study the functional consequences of harbouring the risk mtDNA variant by assessing: mtDNA copy number, mitochondrial biosynthesis, mitochondrial fission and fusion, mitochondrial reactive oxygen species (ROS), oxidative stress, autophagy and a whole transcriptome analysis by RNA-sequencing.

Results mtDNA variant m.16519C is over-represented in rapid progressors (combined OR 1.546; 95% CI 1.163 to 2.054; p=0.0027). Cybrids with this variant show increased mtDNA copy number and decreased mitochondrial biosynthesis; they produce higher amounts of mitochondrial ROS, are less resistant to oxidative stress, show a lower expression of the mitochondrial fission-related gene fission mitochondrial 1 and an impairment of autophagic flux. In addition, its presence modulates the transcriptome of cybrids, especially in terms of inflammation, where interleukin 6 emerges as one of the most differentially expressed genes.

Conclusions The presence of the mtDNA variant m.16519C increases the risk of rapid progression of knee OA. Among the most modulated biological processes associated with this variant, inflammation and negative regulation of cellular process stand out. The design of therapies based on the maintenance of mitochondrial function is recommended.

WHAT IS ALREADY KNOWN ON THIS TOPIC

⇒ There is a percentage of patients with osteoarthritis (OA) that suffer a rapid progression of the disease.
⇒ During the last years, the study of this 'new old' phenotype attracted the attention of scientific community to try to find out the causes of its development and to identify appropriate biomarkers to select these patients to include them in clinical trials.

WHAT THIS STUDY ADDS

⇒ This study demonstrates that mitochondrial DNA (mtDNA) variation is a part of the puzzle that predispose patients with OA to suffer this phenotype.
⇒ Specifically, the presence of the mtDNA D-loop variant m.16519C increases the risk of suffering this phenotype by acting on different aspects of cell behaviour that are related to OA, pointing to an altered inflammatory state that could predispose patients to suffer this phenotype.

HOW THIS STUDY MIGHT AFFECT RESEARCH, PRACTICE OR POLICY

⇒ The study provides a new mitochondrial genetic biomarker able to classify patients prone to suffer this phenotype, opens the door to the design of mitochondrial therapies to treat the disease and highlights the importance to treat inflammation to avoid the development of this phenotype.

INTRODUCTION

Osteoarthritis (OA) is a chronic musculoskeletal disease with a polygenic and heterogeneous nature that involves movable joints. The set of features that take place during the development of the disease lead to consider OA as a severe disease of the whole joint as an organ.[1–3]

The aetiology of knee OA (KOA) is still relatively unknown, but factors such as obesity, age, female gender, knee malalignment, previous knee injury and genetic predisposition are considered risk factors.[4 5] Heterogeneity is one of the characteristics that defines KOA in terms of aetiology and in terms of progression. Normally, KOA can

be considered a slowly progressive disorder, by which most patients gradually worsen over years or decades[6]; however, a small percentage of patients with OA (~14%) show a rapid evolution or structural progression of the disease.[7 8] During the last decades, small series, case reports and different definitions and acceptations have been given to describe this particular phenotype, although the exact aetiopathogenesis of this KOA subtype is still largely unknown.

Over the last few years, an attempt has been made to classify the different terms used to define this dramatic change in joint health: accelerated OA (AOA), rapid progression, rapidly progressive OA, rapidly destructive OA or rapid radiographic change.[8] Although no full consensus has been achieved,[9] different research groups provided definitions of AOA. AOA consists in the rapid onset and progression from preradiographic KOA (KL grade 0 or 1) to advanced end-stage radiographic KOA (KL grade 3 or 4) in 4 years or less.[10]

The underlying causes for the rapid progression phenotype are largely unknown, and the search for differential or specific risk factors seems to be challenging since most of those proposed to date use to overlap with the risk factors for disease incidence, as with those risk factors for normal progression.[11] In this sense, knee malalignment, meniscal damage, female gender, advanced age, previous injury or defects in subchondral bone such as osteoporosis or osteonecrosis, are considered potential risk factors for the rapid progression phenotype.[8 10 12–17]

During the last years, it has been consistently demonstrated that both mitochondrial dysfunction and certain mitochondrial genetic polymorhpisms (ie, mtDNA haplogroups) play an important role on different KOA-related features, including prevalence, structural damage, incidence or progression of the disease.[18–22] Moreover, a series of functional consequences of the mitochondrial genome on the development of the disease has been demonstrated using both animal models[23 24] and cellular models of transmitochondrial cybrids.[25] In addition, mtDNA haplogroups show a direct impact on the nuclear methylation pattern of human articular chondrocytes,[26] on telomere length[27] and on serum levels of specific OA-related biochemical biomarkers.[28] Thus, all these features led to consider them as complementary genetic biomarkers for the disease.[29] However, the influence of mtDNA variation on the development of the rapid progression phenotype has not been evaluated so far.

Taking into account this background, the present study aims to assess the influence of mtDNA variation on the rapidly progressive KOA (RPKOA) phenotype in order to identify potential biomarkers and therapeutic targets in the mtDNA molecule. To carry out this approach, we used longitudinal data from three prospective cohorts of patients (osteoarthritis initiative (OAI), PROspective Cohort of Osteoarthritis from A Coruña (PROCOAC) and Cohort Hip and Cohort Knee (CHECK)), applied high-throughput next-generation sequencing techniques and designed different functional studies using a cellular model based on transmitochondrial cybrids.

METHODS

Cohort's description

For this study, we used data from three independent prospective cohorts of patients: the American OAI cohort as a discovery cohort, and two European cohorts, CHECK and PROCOAC, as replication cohorts. A complete description of OAI, CHECK and PROCOAC cohorts is available elsewhere[30 31] and in the online supplemental methods section.

Rapid progression criteria

We used longitudinal data in terms of Kellgren and Lawrence (KL) grade to define, for the first time, the 'GIR-blanco criteria' to classify patients with RPKOA. This definition considers as rapid progressors those patients with a baseline KL grade of 0–1 or 2 that evolve to KL ≥3 or 4, respectively over a 48-month period (60 months for CHECK participants) (online supplemental figure 1). Radiographs in CHECK were obtained at 24-month, 60-month and 96-month follow-up visits, lacking 48-month visit. Therefore, we only included CHECK patients with a minimum follow-up period of 60 months. Patients that did not meet these criteria, namely, subjects with the same radiographic characteristics at baseline as rapid progressors but with a slower or no structural evolution over time, were classified as non-rapid progressors. We excluded from the study: (i) patients diagnosed with inflammatory arthritis, (ii) non-rapid progressors patients with baseline KL ≥3 on the contralateral knee and/or (iii) patients that did not reach a minimum follow-up period of 48 months in OAI and PROCOAC cohorts (60 months in CHECK).

Taking into account these criteria, we reached 268 rapid progressors from the discovery cohort of the OAI, and 66 and 41 from CHECK and PROCOAC, respectively. Because of their similar frequency distribution in terms of gender and rapid progressors, we performed the study using, as replication cohort, the pooling of CHECK and PROCOAC cohorts (European cohort), reaching a total of 107 rapid progressors. Thus, we got a larger sample size as well as an increased statistical power to perform the appropriate analyses. In addition, we analysed 827 non-rapid progressors from the American cohort of the OAI and 592 non-rapid progressors from the European cohort (307 from CHECK and 285 from PROCOAC).

In summary, we analysed 375 rapid progressors (268 from the discovery cohort and 107 from the replication cohort) and 1419 non-rapid progressors (827 from the discovery cohort and 592 from the replication cohort), reaching a final sample size of 1794 subjects.

In-depth sequencing of mtDNA

The in-depth sequencing of mtDNA was performed in an initial subset of 450 genomic DNA samples from the OAI (184 rapid progressors and 266 non-rapid progressors) using an Ion S5XL sequencer and an Ion Chef (ThermoFisher Scientific) for automatic library preparation and templating. A more detailed explanation of this methodology is available in the online supplemental methods section.

Data analysis

Ion S5XL reads were processed with Ion Torrent Suite software V.5.12 (ThermoFisher Scientific) for variant calling using an adapted pipeline for Ampliseq libraries using a low stringency mode in Torrent Variant Caller plug-in. In order to confirm that results are not contingent to the presence of Nuclear Mitochondrial DNA Segments (NUMTs), a parallel pipeline was carried out consisting in the cleaning-up of NUMT-derived reads.[32 33]

The resultant mitochondrial variants were initially filtered by their population frequency. Thereby, those mitochondrial genetic variants with a minor allele frequency (MAF) <0.05 were discarded. In addition, a conservative heteroplasmy detection threshold was set to 20%. Even so, after automatic variant calling, manual verification was performed using Integrative Genomics Viewer software V.2.8.0.

Statistical analyses of mtDNA sequencing data

All the statistical analyses were performed using IBM-SPSS software, release 24 (IBM, Armonk, New York, USA) and R software V.3.6.3 (The R Foundation for Statistical Computing).

The first step consisted in the selection of those mtDNA genetic variants with a MAF ≥0.05 that showed a differential frequency distribution between rapid progressors and non-rapid progressors. To proceed with this, a χ^2 test was applied to obtain the corresponding OR and 95% CI to select those differentially represented variants with a p value ≤0.15. From these variants, Benjamin-Hochberg (B-H) correction was applied to control the false rate discovery (FDR). Finally, we selected those ones with a corrected p value <0.05.

As a second step, a binary logistic regression model was constructed using, as independent variable, the selected single-nucleotide polymorphism (SNP) and, as confounding factors, a series of clinical variables classically associated with OA, including gender, age, body mass index (BMI), contralateral KOA, Western Ontario and McMaster Universities Osteoarthritis Index (WOMAC) pain score and previous injury in target knee (only available in the discovery cohort). The progression status (rapid progressors/non-rapid progressors) was used as dependent variable.

Meta-analysis

In this work, we performed a subsequent meta-analysis of rapid progression of KOA including data from the American cohort (discovery cohort) and the European cohort (replication cohort). A detailed description is included in the online supplemental methods section.

Estimation of adjusted ORs and their 95% CIs was used as the effect size measurements for the association between m.16519C and the rapid progression of KOA. Meta-analysis was planned to be performed on k=2 studies, with an estimated statistical power of 90.3% to detect as statistically significant an OR ≥1.5 associated with the presence of m.16519C, with a p value <0.05 two-tailed significance level.

Single base extension assay

The assignment of the mtDNA variant with a significantly different frequency between rapid and non-rapid progressors was carried out in the rest of the samples using the single base extension technique on a SeqStudio genetic analyzer system (ThermoFisher Scientific). A detailed description of this methodology is available in the online supplemental methods.

Functional studies using transmitochondrial cybrids

Some parts of these methods are available in the extended online supplemental methods section.

Cybrid culture and culture conditions

The human osteosarcoma cell line 143B.TK$^-$ Rho-0 was used to generate the transmitochondrial cybrids by fusing them with platelets from healthy donors carrying the haplogroup H and the mtDNA variant m.16519C or the haplogroup H and the mtDNA reference m.16519T, following the protocol described elsewhere.[25] We finally obtained two H cybrids: cybrids harbouring the mtDNA variant m.16519C and cybrids not harbouring the mtDNA variant m.16519C.

mtDNA copy number and complete mtDNA sequencing

Total DNA from cultured cells was isolated with the QIamp DNA Mini kit (QIagen) following manufacturer recommendations. mtDNA copy number was performed using SYBR Green in a LightCycler 480 II system (Roche). This measure was determined from the *Crossing Point* values for both mitochondrial 12S ribosomal gene and nuclear RNaseP gene (12S/RNaseP ratio) using standard curves.

Finally, in order to verify that the generated cybrids carry the desired mtDNA sequence in terms of haplogroup H and mtDNA variant m.16519C, two clones per cybrid were in-depth sequenced following the above described method.

RNA isolation and gene expression assays

Quantitative real-time PCR experiments were performed on a LightCycler 480 II system (Roche) using Universal Probe Library probes to assess the gene expression of genes related to: (i) mitochondrial biogenesis (peroxisome proliferator-activated receptor gamma, coactivator 1 alpha (*PPARGC1A*) and transcription factor A mitochondrial (*TFAM*)) and (ii) mitochondrial fusion (mitofusin 2 (*MFN2*)) and fission (fission mitochondrial 1 (*FIS1*)) processes.

RNA-sequencing

A whole transcriptome assay by RNA-sequencing (RNA-seq) was performed in cybrids with and without the mtDNA single nucleotide variant (mtSNV) m.16519C; 10 ng of complementary DNA were used to automatically generate libraries using the Ion AmpliSeq Transcriptome Human Gene Expression Panel, Chef-Ready Kit (ThermoFisher) and subsequently sequenced on an Ion S5XL.

CHP files, normalised by reads per million, were automatically generated and imported into the Transcriptome Analysis Console (TAC) software V.4.0 (ThermoFisher) to perform the differential expression analyses between cybrids with and without the mtSNV m.16519C. Only those genes with a fold change (Fc) between groups >±2 and a corrected (FDR) p value <0.05 were considered as differentially expressed.

To explore the biological meaning of the differentially expressed genes, we loaded both the list of upregulated and downregulated genes associated with the presence of the mtSNV m.16519C into the online Enrichr web-based tool (https://maayanlab.cloud/Enrichr/).

Mitochondrial reactive oxygen species production assay

Mitochondrial superoxide anion was evaluated with MitoSox Red (ThermoFisher Scientific) at a final concentration of 5 µM in Hank's Balanced Sal Solution (Sigma) during 15 min at 37°C in darkness.

Oxidative stress response assay

To evaluate the oxidative stress response under H_2O_2 incubation, the CellTiter 96 Aqueous Assay kit (Promega) was used. A NanoQuant Infinite M200 (Tecan) microplate reader was used to measure absorbance at 490 nm.

To examine for oxidised proteins relevant to mitochondrial stress, we analysed post-translational modifications by liquid chromatography-tandem mass spectrometry in two clones from each type of cybrids.

Mitochondrial membrane potential

We analysed the mitochondrial membrane potential (ΔΨm) using the fluorescent Mito Probe JC-1 in cybrids 509 (m.16519T) and 513 (m.16519C) by flow cytometry. For this procedure, cybrids cells were seeded at 8×10^4 cells per well in 12-well plates in Dulbecco's Modified Eagle Medium (DMEM) 10% fetal bovine

serum; when the cell confluence arrived near 80%, the medium was changed to DMEM serum-free and treated with the protonophore carbonylcyanide-p-trifluoromethoxyphenylhydrazone (FCCP) at 500 µM during 3 hours as depolariser inductor.

The percentage of depolarisation in a cell population was calculated by dividing the values of red fluorescence by the values of green fluorescence. Therefore, the ratio aggregates:-monomers (red:green fluorescence) in relation to the FCCP ratio corresponding to each cybrid was subsequently used to present the data.

Autophagy determination

Autophagy was evaluated through the quantification of microtubule-associated protein 1A/1B-light chain 3 (LC3) (*I and II*), sequestosome-1 (p62/SQSTM1) and phospho-ribosomal protein S6 (p-rpS6) by western blot analysis (WB). The intensity of the bands and the WB quantification were assessed by using Amershan Imager 600 software.

RESULTS

Identification of differential mtDNA variants

The analysis of the entire mtDNA sequences obtained from the initial subset of the discovery cohort revealed 1335 mtSNVs. Then, after filtering the mtSNVs, considering only those ones with a MAF ≥0.05, the number of variants dropped to 76 (online supplemental table 1). Among these 76 mtSNVs, we considered for further analyses only those differentially distributed (nominal p≤0.15) between rapid progressors and non-rapid progressors. After these initial approaches, the final number of differential mtSNVs was 9 (online supplemental table 2): m.146T>C (nominal p=0.15), m.930G>A (nominal p=0.11), m.5147G>A (nominal p=0.044), m.10463T>C (nominal p=0.106), m.11812A>G (nominal p=0.042), m.14233A>G (nominal p=0.042), m.16294C>T (nominal p=0.07), m.16296C>T (nominal p=0.045) and m.16519T>C (nominal p=0.005). The only variant that passed the B-H correction was the mtSNV m.16519C (B-H corrected p=0.045).

Neither those rare mtSNVs with a MAF <0.05, nor multiplasmic positions at D310 poly-C tract (CCCCCTCCCCC from position 303 to 316), D16189 poly-C tract (CCCCCTCCCC from 16 184 to 16 193) and $(CA)_n$ dinucleotide repeat at position 514, were taken into account in this study, although they were identified.

mtDNA variant m.16519C and rapid progression of knee OA in subjects of the discovery cohort of the OAI

No significant differences were detected in the distribution of gender, BMI, WOMAC pain and contralateral KOA at baseline among subjects with or without the mtSNV m.16519C. However, a significantly higher average age was detected in subjects harbouring the allele C at mt.16519 (p=0.049) (table 1).

The frequency of the mtSNV m.16519C in the complete discovery cohort (268 rapid progressors and 827 non-rapid progressors) was significantly higher in patients who suffered a rapid progression of KOA, compared with those patients with a slower or no progression (OR 1.558; 95% CI 1.148 to 2.113; p=0.004) (table 2).

The association of the mtSNV m.16519C remained significant even after adjusting the model for different confounding clinical variables (OR 1.559; 95% CI 1.125 to 2.161; p=0.008). In addition, BMI (OR 1.073; 95% CI 1.037 to 1.110; p<0.001), contralateral KOA at baseline (OR 2.475; 95% CI 1.830 to 3.348; p<0.001) and WOMAC pain (OR 1.122; 95% CI 1.065 to 1.183; p<0.001) were also significant risk factors to suffer the rapid progression phenotype (table 3).

mtDNA variant m.16519C and rapid progression of knee OA in subjects of the European replication cohort

No significant differences were detected in the distribution of gender, BMI, WOMAC pain and contralateral KOA at baseline among subjects with or without the mtSNV m.16519C. However, a significantly higher average age was detected in subjects harbouring the allele C at mt.16519 (p=0.030) (table 1).

The frequency of the mtSNV m.16519C in the replication cohort (107 rapid progressors and 592 non-rapid progressors) was significantly higher in patients who suffered a rapid progression of KOA, compared with those patients with a slower or no progression (OR 1.647; 95% CI 1.013 to 2.678; p=0.043) (table 2).

The regression logistic model developed for the replication cohort showed some differences in terms of statistical significance, but with the same trend than observed in the discovery

Table 1 Demographic and clinical characteristics of the study populations grouped by mtSNV m.16519T>C

	Discovery cohort (OAI)				Replication cohort (CHECK+PROCOAC)			
Characteristic	m.16519T (n=373, 34.1%)	m.16519C (n=722, 65.9%)	P value	Total (n=1095)	m.16519T (n=215, 30.5%)	m.16519C (n=484, 68.6%)	P value	Total (n=699)
Age at baseline (years)	61.62±8.6	62.76±9.1	**0.049***	62.37±8.9	59.98±7.8	61.06±7.6	**0.030***	60.72±7.7
Gender			0.780†				0.720†	
Male	144 (38.6)	285 (39.5)		429 (39.2)	47 (21.9)	100 (20.7)		147 (21.0)
Female	229 (61.4)	437 (60.5)		666 (60.8)	168 (78.1)	384 (79.3)		552 (79.0)
BMI (kg/m^2)	28.04±4.4	28.26±4.5	0.543*	28.19±4.5	27.65±4.4	27.65±4.9	0.585*	27.65±4.8
Contralateral knee OA			0.447†				0.689†	
Yes	161 (43.3)	329 (45.7)		490 (44.9)	34 (16.3)	73 (15.1)		107 (15.5)
No	211 (56.7)	391 (54.3)		602 (55.1)	174 (83.7)	409 (84.9)		583 (84.5)
WOMAC pain	1.80±2.7	1.81±2.6	0.552*	1.80±2.6	5.23±4.2	4.91±3.9	0.484*	5.01±4.0

Values are mean±SD or number of patients with percentage in parentheses.
Significant p values are in bold.
*Mann-Whitney U test for comparison between T and C alleles at m.16519.
†χ^2 test.
BMI, body mass index; CHECK, Cohort Hip and Cohort Knee; mtSNV, mitochondrial DNA single nucleotide variant; OAI, osteoarthritis initiative; PROCOAC, Prospective Cohort of Osteoarthritis A Coruña; WOMAC, Western Ontario and McMaster Universities Osteoarthritis Index.

Table 2 Frequencies and ORs of mtSNV m.16519T>C in rapid and non-rapid progressors of the discovery cohort of the OAI (upper) and of the European replication cohort (lower)

Discovery cohort (OAI)						
m.16519	**Rapid (n=268)**	**Non-rapid (n=827)**	**Total (n=1095)**	**OR**	**95% CI**	**P value***
T	72 (26.9)	301 (36.4)	373 (34.1)			
C	196 (73.1)	526 (63.6)	722 (65.9)	1.558	1.148 to 2.113	**0.004**
Replication cohort (CHECK+PROCOAC)						
m.16519	**Rapid (n=107)**	**Non-rapid (n=592)**	**Total (n=699)**	**OR**	**95% CI**	**P value***
T	24 (22.4)	191 (32.3)	215 (30.8)			
C	83 (77.6)	401 (67.7)	484 (69.2)	1.647	1.013 to 2.678	**0.043**

Values are number of patients with percentage in parentheses.
Significant p values are in bold.
*χ^2 test.
CHECK, Cohort Hip and Cohort Knee; mtSNV, mitochondrial DNA single nucleotide variant; OAI, osteoarthritis initiative; PROCOAC, Prospective Cohort of Osteoarthritis A Coruña.

cohort. The association of the mtSNV m.16519C borderlined the statistical significance maintaining the same trend as the discovery cohort (OR 1.506; 95% CI 0.848 to 2.675; p=0.162). Among the rest of confounding factors, only the contralateral KOA remained significant as a risk factor for suffering the rapid progression phenotype (OR 2.056; 95% CI 1.005 to 4.208; p=0.048) (table 3).

Meta-analysis

The search process did not identify any article or abstract describing the association of the mtSNV m.16519C with the rapid progression phenotype, so that only the two association studies performed herein were selected for subsequent meta-analysis. Considering both analyses, a total of 1794 subjects (1095 patients from the discovery cohort of the OAI and 699 from the European replication cohorts—CHECK and PROCOAC) were included in the meta-analysis.

No between-study heterogeneity was detected for the mtSNV m.16519C (I^2=0%, p=0.9173), however, the random-effect model was evenly used. The discovery cohort of the OAI showed a higher contribution or relative weight (W) than the European replication cohort (75.5% vs 24.5%). The results showed that the mtSNV m.16519C significantly associates with the RPKOA phenotype (combined OR 1.546; 95% CI 1.163 to 2.054; p=0.0027) (figure 1).

Table 3 Multivariable logistic regression model to predict the risk of rapid knee OA progression in subjects belonging to the OAI cohort (upper) and subjects belonging to the European replication cohort (lower)

Variable	B	Adjusted OR	95% CI	P value
Discovery cohort (OAI)				
Age	0.052	1.054	1.036 to 1.072	**<0.001**
Gender (female)	0.230	1.258	0.926 to 1.710	0.142
BMI	0.070	1.073	1.037 to 1.110	**<0.001**
Contralateral knee OA	0.906	2.475	1.830 to 3.348	**<0.001**
WOMAC pain	0.115	1.122	1.065 to 1.183	**<0.001**
m.16519C	0.444	1.559	1.125 to 2.161	**0.008**
Replication cohort (CHECK+PROCOAC)				
Age	−0.028	0.972	0.937 to 1.009	0.138
Gender (female)	0.459	1.582	0.796 to 3.144	0.190
BMI	0.010	1.010	0.957 to 1.066	0.727
Contralateral knee OA	0.721	2.056	1.005 to 4.208	**0.048**
WOMAC pain	0.060	1.062	0.997 to 1.130	0.060
m.16519C	0.410	1.506	0.848 to 2.675	0.162

Statistical significance declared at p<0.05 (in bold).
B, regression coefficient; BMI, body mass index; CHECK, Cohort Hip and Cohort Knee; OAI, osteoarthritis initiative; PROCOAC, Prospective Cohort of Osteoarthritis A Coruña; WOMAC, Western Ontario and McMaster Universities Osteoarthritis Index.

Functional study using transmitochondrial cybrids

Basal conditions of transmitochondrial cybrids were established following the protocol described elsewhere.[34]

In-depth sequencing of mtDNA

The complete sequencing of mtDNA from those cybrids used in this study confirmed the haplogroup H and the presence or absence of the mtDNA variant m.16519C. Uploading of the variant call format file into the Haplogrep V.2.4.0 software (https://haplogrep.i-med.ac.at/app/index.html) revealed that cybrids carrying the mtSNV m.16519C belong to the subhaplogroup H76 and cybrids not harbouring this mtSNV belong to the subhaplogroup H6b2. In addition, a series of differential SNPs between both cybrids have also been detected. Thus, cybrids H6b2 and H76 showed 15 and 14 differential mtSNVs, respectively with respect to revised Cambridge Reference Sequence (rCRS). Six of these variants (m.263G, m.750G, m.1438G, m.4769G, m.8860G and m.15326G) were common to both cybrids; therefore, in addition to the mtSNV m.16519C, the number of differential mtDNA variants between them was 16 (online supplemental table 3). No traces of heteroplasmy were detected in the obtained sequences.

mtDNA copy number

After a set of 7 independent measures per clone, in each of the 2 clones, we obtained a total of 14 values per cybrid. The results showed that cybrids with mtSNV m.16519C have 1.68-fold more copies of mtDNA than cybrids without this variant (447.4±39.3 vs 266.3±41.14; p=0.001) (figure 2A).

Mitochondrial biogenesis and mitochondrial fusion/fission processes

We performed different gene expression experiments to analyse the effect of the mtSNV m.16519C on: (i) mitochondrial biogenesis, by assaying the expression of both *PPARGC1A* and *TFAM* and (ii) mitochondrial fusion and fission processes, by assaying the expression of *MFN2* and *FIS1*, respectively.

In terms of mitochondrial biogenesis, the results obtained showed a significantly lower expression of *PPARGC1A* in cybrids harbouring the mtSNV m.16519C than in cybrids with the wild-type m.16519T (1.02±0.23 vs 2.64±0.62; p=0.0024);

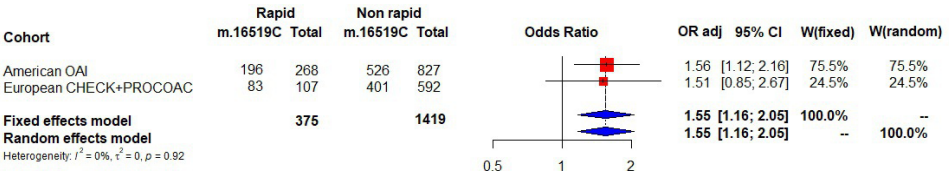


Figure 1 Forest plot of the associations analysed in this work involving the mtDNA variant m.16519C and the risk of rapid progression of knee OA. CHECK, Cohort Hip and Cohort Knee; OAI, Osteoarthritis Initiative; PROCOAC, Prospective cohort of osteoarthritis A Coruña; Rapid m.16519C, number of subjects with the mtDNA variant m.16519C that developed and did not develop a rapid knee OA progression; W, relative weight.

regarding *TFAM*, the trend was similar than for *PPARGC1A* but without reaching the statistical significance (0.93±0.15 vs 1.24±0.14; p=0.063) (figure 2B). Regarding mitochondrial fusion and fission processes, the results revealed a lower expression of the fission-related gene *FIS1*, that borderlined the statistical significance, in cybrids harbouring the mtSNV m.16519C (0.79±0.12 vs 1.56±0.35; p=0.052) (figure 2C).

Cybrids transcriptome analysis by RNA-seq

An initial unsupervised hierarchical clustering analysis based on principal component analysis showed two well-differentiated clusters representing both types of cybrids, those with the mtSNV m.16519C and those without this variant. The first component explained the 60.6% of the variance (online supplemental figure 2).

Expression analysis using TAC software showed 460 differentially expressed genes between both types of cybrids, of which 167 were upregulated in cybrids harbouring the mtSNV m.16519C, and 293 were downregulated (online supplemental table 4). Interestingly, these analyses confirmed the most robust association regarding mitochondrial biogenesis-related gene *PPARGC1A* downregulation in cybrids with the risk allele C at m.16519 (*PPARGC1A*; Fc=−3.52; nominal p=0.0009) (online supplemental table 4). On the other hand, downregulation of *FIS1* in these cybrids was also observed throughout RNA-seq, although did not reach the statistical significance (*FIS1*;

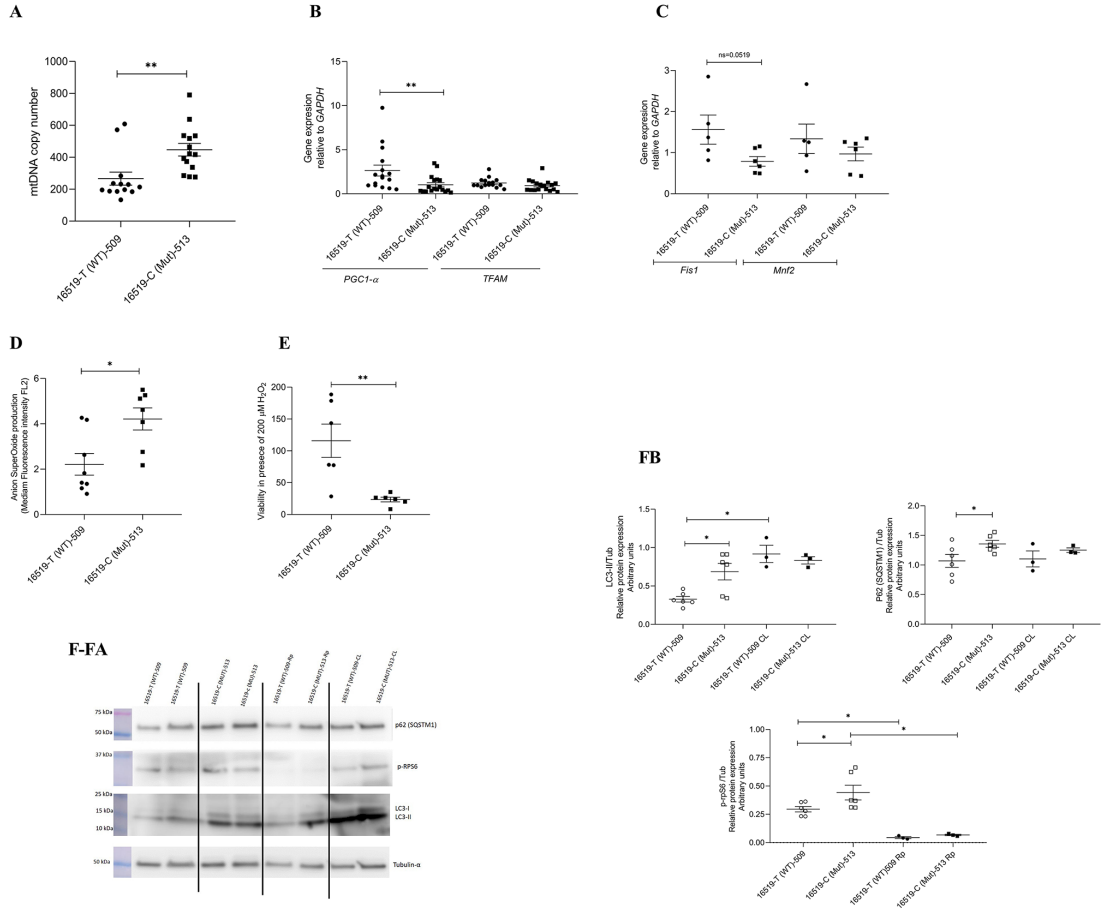


Figure 2 Functional analysis in transmitochondrial cybrids with and without the mtDNA variant m.16519C. (A) mtDNA copy number; (B) mRNA expression of genes related to mitochondrial biogenesis PGC1a (*PPARGC1A*) and *TFAM*; (C) mRNA expression of genes related to mitochondrial dynamics *FIS1* and *MFN2*; (D) superoxide anion production; (E) percentage of cells surviving oxidative stress after incubation with 300 µM H_2O_2. (F) Analysis of autophagy flux: (FA) Western blot analysis using anti-p62 (SQSTM1), LC3 (I and II), p-rpS6 and TUB; (FB) numeric data obtained by densitometry (relative protein expression). All data were obtained from three independent experiments. *p<0.05; **p<0.005. CI, cloroquine; *FIS1*, fission mitochondrial 1; LC3, light chain 3; *MFN2*, mitofusin 2; mRNA, messenger RNA; mtDNA, mitochondrial DNA; mut, mutant; *PPARGC1A*, peroxisome proliferator-activated receptor gamma, coactivator 1 alpha; p-rpS6, phospho-ribosomal protein S6; RP, rapamycin; *TFAM*, transcription factor A mitochondrial; TUB, α-tubulin; WT, wild type.

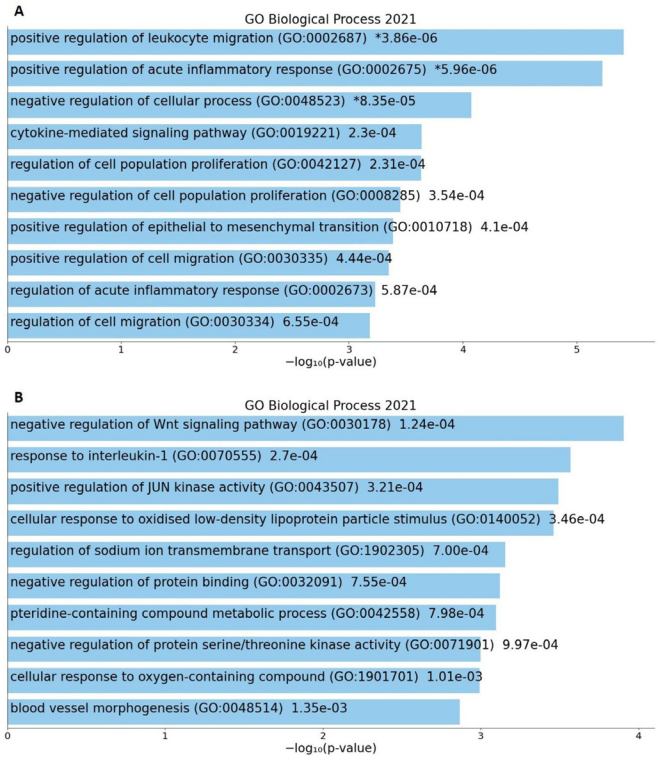


Figure 3 GO analysis using Enrichr online tool. (A) Top 10 biological processes associated with genes significantly upregulated in cybrids with m.16519C. (B) Top 10 biological processes associated with genes significantly downregulated in cybrids with m.16519C. Asterisk (*) indicates a biological process significantly altered after corrected (FDR) p values (FDR p<0.05). FDR, false rate discovery; GO, Gene Ontology.

Fc=−1.46; nominal p=0.0829) (data not shown). On the contrary, no significant differences were detected in relation to the expression of *TFAM* and *MFN2*.

Gene Ontology (GO) analysis showed three biological processes significantly modulated (FDR p<0.05), all of them among upregulated genes in cybrids harbouring the risk allele m.16519C: positive regulation of acute inflammatory response (GO:0002675), positive regulation of leucocyte migration (GO:0002687) and negative regulation of cellular process (GO:0048523) (figure 3A,B). Among the upregulated genes involved in these processes, insulin growth factor binding protein-1, interleukin 6 (*IL6*), interleukin 24 (*IL24*), prostaglandin-endoperoxide synthase 2 (*PTGS2*), colony-stimulating factor 1, vascular endothelial growth factor A, transforming growth factor beta 2 or oncostatin M receptor stand out (online supplemental table 4).

Validation of *IL6* and *PTGS2* expression by quantitative PCR (qPCR) revealed the same trend, showing a significant overexpression in cybrids harbouring the risk allele C at m.16519 (Fc=14.06 and 3.03, respectively; p=0.0286) (data not shown).

Mitochondrial ROS production and oxidative stress response

Cybrids harbouring the mtSNV m.16519C showed a significantly higher production of mitochondrial superoxide anion than cybrids without this mtDNA variant (4.21±0.49 vs 2.21±0.47; p=0.014) (figure 2D). Moreover, cybrids carrying the mtSNV m.16519C showed a lower mean percentage of surviving cells in presence of H_2O_2 than cybrids carrying the allele T at m.16519 (23.42±3.67 vs 115.7±26.17; p=0.0043) (figure 2E).

In addition, cybrids harbouring the mtSNV m.16519C showed a significantly increased amount of oxidation in four mitochondrial stress-related proteins (p<0.001), including chaperonin 10 (CH10), chaperonin 60 (CH60), glucose-regulated protein 75 (GRP75) and thioredoxin-dependent peroxide reductase (PRDX3) (online supplemental figure 3A).

From these results, it can be concluded that the presence of the allele C at m.16519 implies a greater production of the mitochondrial superoxide anion, and cells carrying this variant are more sensitive to oxidative stress.

Mitochondrial membrane potential

The results obtained showed a significantly decreased red:green ratio in cybrids with the mtDNA variant m.16519C than in cybrids harbouring the wild-type variant m.16519T (p=0.042). These data indicate a higher depolarisation in cybrids with this mtDNA variant under adverse conditions (FCCP incubation) (online supplemental figure 3B).

Autophagy

Autophagy flux analysis revealed that cybrids harbouring the mtSNV m.16519C show significantly higher protein p62 (1.36±0.06 vs 1.07±0.11; p=0.041) and LC3-II (0.69±0.11 vs 0.33±0.04; p=0.026) accumulation than cybrids without this mtDNA variant. In addition, protein levels of p-rpS6, target of mammalian target of rapamycin (mTOR) and indicator of its activity, were also significantly higher in cybrids with the mtSNV m.16519C variant (0.44±0.06 vs 0.29±0.02; p=0.026). Incubation with chloroquine to prevent autophagosome formation shows no significant differences between both cybrids on protein levels of p62 and LC3-II. However, a significantly increased expression of LC3-II only on cybrids harbouring the wild-type variant m.16519T incubated with chloroquine in relation to its basal non-incubated condition was detected (0.33±0.04 vs 0.92±0.11; p=0.0238). Finally, the incubation with rapamycin as a positive control, which inhibits mTOR, led to the almost total repression of p-rpS6 both in cybrids with the mtSNV m.16519C (0.44±0.06 vs 0.07±0.01; p=0.0238) and cybrids without this variant (0.29±0.02 vs 0.04±0.01; p=0.0238) (figure 2F). Altogether, these results point to an impairment of autophagy flux in cybrids with mtSNV m.16519C variant, which could lead to an excessive accumulation of damaged material, namely a less effective autophagy.

DISCUSSION

In this study we report, for the very first time, an association study and meta-analysis of the mtDNA variant m.16519T>C on the risk of developing a RPKOA. In addition, we provide data supporting possible explanations for this association by using a cellular model based on transmitochondrial cybrids.

We aimed to identify subjects prone to suffer a rapid evolution of the disease in terms of radiographic (KL grade) worsening of the joint, describing for the first time the 'GIR-blanco criteria' for RPKOA. Hence, we included subjects with preradiographic KOA at baseline (KL grade 0–1) as well as subjects with consolidated KOA at baseline (KL grade 2). Patients selected under these inclusion criteria showed a significant correlation with the rapidly progressive osteoarthritis type 1 (RPOA 1) criteria (loss in joint space width ≥2 mm within 1 year, without gross structural failure) in subjects from the discovery cohort of the OAI. Thus, up to 64.7% of (n=68) genotyped patients classified as rapid progressors using RPOA 1 criteria were also classified as rapid progressors using the 'GIR-blanco criteria', and up to

75.4% of (n=529) genotyped patients classified as non-rapid progressors using RPOA 1 criteria corresponded with non-rapid progressors according to 'GIR-blanco criteria' (p<0.0001) (online supplemental table 5). Besides, the frequency of the mtSNV m.16519C was also significantly higher in patients classified as rapid progressors following RPOA 1 criteria (OR 1.853; 95% CI 1.016 to 3.379; p=0.042) (online supplemental table 5).

The robustness of the association described herein manifests in different aspects. On the one hand, we identified a significant association in subjects of the discovery cohort of the OAI that was replicated in a different cohort, composed of subjects from PROCOAC and CHECK, and subsequently meta-analysed. On the other hand, even analysing these two cohorts separately, the frequency of the mtSNV m.16519C followed the same trend with an OR >1.4 (online supplemental table 6). Although it was not significant due to the small sample size, the subsequent meta-analysis of the ORs from the univariate analysis of the three cohorts did reach the statistical significance (combined OR 1.583; 95% CI 1.222 to 2.049; p=0.0005) (online supplemental figure 4). Finally, by removing from the study those patients who remained radiologically stable in terms of KL grade during the follow-up, the association remained significant (p=0.016 for the discovery cohort and p=0.006 for the replication cohort) (online supplemental table 7). Altogether, these data indicate that mtSNV m.16519C is able to classify, specifically, patients at increased risk of suffering a RPKOA progression in terms of their KL grade evolution.

From the results obtained in this study and others,[7 10] it can also be concluded that both rapid and non-rapid progressors share a set of clinical conditions that predispose them to progression. Thus, in our regression model, common OA-related clinical variables such as age, BMI, contralateral KOA and WOMAC pain at baseline associate, in an independent manner, with the risk of rapid progression. In addition, previous injury in target knee (data not available in the European cohorts) associates with the risk of rapid progression too (online supplemental table 8). Therefore, given the high frequency of the C allele at m.16519 (above 60%), as well as its robust statistical association as a risk factor, the interaction of this mtSNV with another unknown (clinical and/or genetic) risk factor/s could increase the predisposition of patients with KOA to suffer a more severe progression of the disease, as has been described in other disorders such as breast cancer.[35] The combination of this mitochondrial variant with some of the multiple independent loci associated with OA in the largest and most powerful genome-wide association studies to date,[36] is an example of genetic interactions that could be potentially explored. Interestingly, among these 100 independently reported loci associated with OA risk, only the signal rs753350451 belongs to a nuclear region that codes for a mitochondrial protein, the mitochondrial translation release factor in rescue.[36] This protein plays a main role in the regulation of the synthesis of the 13 essential subunits of the oxidative phosphorylation complexes by preventing aberrant translation. Therefore, due to the role of the mtSNV 16519T>C in mtDNA replication and transcription,[37 38] of special interest could be to explore the impact of the interaction between these two SNPs.

Different associations involving the mtSNV m.16519T>C have been described. The T allele has been associated with different disorders, including pancreatic cancer-associated diabetes mellitus and shorter life expectancy in these patients,[37] irritable bowel syndrome,[39] high altitude pulmonary oedema,[40] migraine headache and cycling vomiting syndrome[38] or multiple functional syndromes.[41] On the other hand, the C allele at mtSNV m.16519 has been identified as a risk factor for type 2 diabetes mellitus,[42] colorectal cancer,[43] breast cancer,[44] sudden death infant syndrome in females[45] and decreased change rate of VO_{2max} and citrate synthase activity after endurance training.[46] To our knowledge, the association reported in the present study between this mtSNV and the RPKOA phenotype is the first association that describes the involvement of this mitochondrial SNP in OA pathology, and the first genetic polymorphism associated with this phenotype. In addition, despite patients with this variant are significantly older in both cohorts, we consider that this has no clinical relevance since this difference is barely 1 year and the study design allows a minimum mean difference of 0.3 between C and T alleles to be detected as significant.

This variant is present in all major human races and has arisen multiple times in human evolution; in fact, it is widely distributed among the different mtDNA haplogroups, but with an increased frequency of the C allele among H subjects and a lower frequency among J subjects.[47 48] The latter would be in agreement with our previous results, in which we proposed both haplogroups H and J as risk and protective OA-related biomarkers, respectively.[19] The SNP is located in the 1 kb non-coding mtDNA control region, not far from the origin of heavy-strand replication and putative mitochondrial internal membrane-attachment site.[39] Thus, as stated before, variations at this position could act by altering the transcription levels of mitochondrial proteins as well as mtDNA replication.[37 38] Moreover, since the T>C change at m.16519 originates a potential CpG site in the D-loop region, a possible effect of this variant on mtDNA methylation patterns should not be discarded; altered methylation levels in the D-loop region have been associated with different human disorders including cancer, diabetes, neurodegenerative diseases, ageing and senescence.[49] In addition, different inherited mtDNA variants affecting the non-coding D-loop region have also been associated with N-formylmethionine levels, an initiation amino acid for intramitochondrial translation[50] that has huge effects on proteostasis and degradation processes, thereby modifying the risk of late-onset diseases.[51]

Nevertheless, this study is the first to demonstrate a series of functional consequences of this variant by using transmitochondrial cybrids with uniform nuclear background. Thus, cells harbouring the mtSNV m.16519C show increased mtDNA copy number, decreased expression of the mitochondrial biogenesis-related gene *PPARGC1A (PGC1α)* and decreased expression of *FIS1* gene. In addition, they produce higher amounts of mitochondrial reactive oxygen species (ROS), are less prone to cope with oxidative stress and manifest a different behaviour of the autophagy flux than cells harbouring the T allele at m.16519. Besides, the presence of the risk variant m.16519C induces significant differences in the transcriptome of cybrids, specifically in terms of inflammation. Although most of these features may be partially regulated by the nuclear background through an anterograde regulation of mitochondrial function,[26 52] the simultaneous occurrence of all of them, together with other unknown risk factors, could dramatically accelerate the progression of the disease.

mtDNA copy number is regulated to ensure that mitochondria are able to generate appropriate amount of energy and intracellular signals to maintain cellular functions.[53] However, it has also been demonstrated that this parameter is increased in a variety of diseases,[54–56] including OA.[18] As for other human disorders, the increased levels of mtDNA could represent a mechanism to compensate for the lower activity of electronic transport chain complexes.[56] In agreement with previous studies carried out in human OA chondrocytes,[57] mitochondrial biogenesis seems to

be diminished in cybrids with m.16519C. In this sense, activation of mitochondrial biogenesis has been proposed as a potential therapeutic strategy to limit OA progression.[57]

An interesting finding of this study is the relationship between the decreased expression levels of *FIS1* and the results of the autophagy assay. It has been demonstrated that mitochondrial dynamic proteins, including *FIS1* and *MFN2*, are associated with autophagy.[58 59] In vitro studies revealed that *FIS1* is significantly decreased in OA chondrocytes, and that *FIS1* depletion leads to an accumulation of *LC3* aggregates, inhibiting autophagy and, thereby, enhancing chondrocyte apoptosis.[60] Moreover, in mice *FIS1* small interfering RNA with destabilisation of the medial meniscus surgery, these authors also detected an increased cartilage destruction.[60] The results obtained in the present study are in agreement with the above findings, since they show significantly increased levels of *LC3-II* and decreased expression of *FIS1* in cybrids harbouring m.16519C, therefore pointing to an accumulation of damaged material that is not properly removed due to an impairment of the autophagy flux.

Because of mitochondrial dysfunction, the balance between ROS production and antioxidant capacity in chondrocytes is broken. This imbalance leads to oxidative stress-mediated cell injury, upregulation of metalloproteinases, overproduction of pro-inflammatory cytoquines, disrupting of cell signalling and cartilage degradation, therefore contributing to the progression of OA.[61–63] In agreement with this, cybrids harbouring the risk variant m.16519C produce higher amounts of mitochondrial ROS and are less prone to cell survival after exposure to hydrogen peroxide, and they show an increased oxidation degree in mitochondrial stress-related proteins such as CH10, CH60, GRP75 and PRDX3, as well as a higher depolarisation under stress conditions. The involvement of certain chaperones in OA is not new[64 65]; hence, their potential role in the development of this phenotype should be explored. Altogether, these data provide evidence of increased oxidation and mitochondrial stress in cybrids carrying the m.16519C variant. In this sense, some studies point to the development or use of targeted antioxidant-based therapies to treat the progression of OA.[63 66 67]

Transcriptome analysis of cybrids with and without the mtSNV m.16519C revealed an interesting role of inflammation as one of the most significant biological process related to the overexpression of specific genes in cybrids harbouring the risk allele C. The role of inflammation in the development and progression of OA is well known.[68–70] Therefore, the presence of this mtSNV could create a pro-inflammatory genetic background that increases the predisposition of specific patients with OA to suffer a rapid evolution of the disease in terms of structural progression. Within this inflammatory environment, in which *PTGS2* is also significantly increased, the potential role of IL6 is of special interest. *IL6* mRNA expression is significantly increased in cybrids with the risk variant m.16519C. Interestingly, preliminary data obtained by our group, in a small subset of (n=26) rapid and (n=200) non-rapid progressors of the OAI, show that baseline serum levels of this cytokine are non-significantly elevated in patients with RPKOA after adjusting for the confounder factors of age, gender, BMI, contralateral KOA and WOMAC (OR 1.934; 95% CI 0.761 to 4.912; p=0.165) (data not shown). In line with these evidences, some studies highlight the importance of this cytokine in both OA incidence and severity.[71] The implication of this cytokine in the development of this phenotype deserves further investigation.

The influence of mitochondrial dysfunction and mtDNA variation in OA is not new.[19] Recent examples of this association include the influence of specific mtDNA haplogroups on both structural damage and the rate of incidence and progression of KOA.[20 22 34 72] Moreover, this influence has also been demonstrated using animal models, specifically conplastic mice and prematurely ageing mutator mice Polg^{D275A}.

Conplastic mice BL/6^{C57} show, after surgically inducing OA, a decreased autophagy, a higher apoptosis as well as an increased Osteoarthritis Research Society International (OARSI) histopathology score, subchondral bone, menisci score and synovitis compared with the BL/6^{NZB} strain.[24] Interestingly, 91 was the number of differential mtDNA mutations between both strains, including 14 missense mutations, 4 transfer RNA mutations or 8 ribosomal RNA mutations.[73] Prematurely ageing mtDNA mutator mice Polg^{D275A} are more prone to suffer elevated subchondral bone turnover and hypertrophy in calcified cartilage.[74] In addition, both mitochondrial dysfunction and the accumulation of somatic mtDNA mutations have been related to impaired osteogenesis and accelerated bone loss in mtDNA mutator mice Polg^{D275A}.[23]

The study presented herein has some limitations that must be drawn. On the one hand, in addition to m.16519T>C, there are a few differential mitochondrial genetic polymorphisms between H6b2 and H76 cybrids and, unfortunately, the application of mtDNA editing techniques to manage homoplasmic mtDNA polymorphisms are far from effective actually. However, mtDNA haplogroup H has been established as a good candidate to perform genetic association studies due to its relatively low intragroup sequence variability.[38] Moreover, m.16519T>C was the only mtSNV that showed a significantly robust differential distribution between rapid and non-rapid progressors. On the other hand, we pooled subjects from CHECK and PROCOAC in order to achieve an improved statistical power as well as an increased sample size of rapid progressors in the (European) replication cohort; however, even performing the meta-analysis using data from the three cohorts separately, the results of the meta-analysis remained significant. In relation to this, taking into account the wide distribution of this variant among the different mitochondrial lineages, the former does not exclude the fact that these findings should be replicated in another independent (non-Caucasian) cohort of patients with KOA. Finally, this study analyses the effect of the mtSNV m.16519C on the uniform nuclear background of the osteosarcoma cell line 143B.TK^{-} Rho-0, but not its impact on a different nuclear background. Because of the well-documented mitochondria-nucleus interactions,[35 52 75] functional studies using chondrocytes-derived cybrids would be desirable.

The discovery of biomarkers that allow the identification of patients prone to suffer a rapid progression of KOA to include them in clinical trials, has been one of the research priorities of the EULAR[76] and the Applied Public-Private Research enabling OsteoArthritis Clinical Headway.[77] After analysing 1794 subjects, the results of this work show, for the first time, that the presence of the mtSNV m.16519C increases the risk of suffering a rapid progression of KOA in terms of KL grade. From this study, it can be concluded that: (i) the design of therapies based on the maintenance of mitochondrial function or mtDNA editing should become a reality as soon as possible; (ii) the genetic predisposition to an increased inflammatory state emerges as a potential risk factor for the development of this phenotype and (iii) the identification of other unknown (clinical and/or genetic) risk factors, whose interaction with this mitochondrial variant increases the predisposition to a rapid progression of the disease, should be mandatory.

Author affiliations
[1]Grupo de Investigación en Reumatología (GIR), Complexo Hospitalario Universitario de A Coruña (CHUAC), Sergas, Universidade da Coruña (UDC), Instituto de Investigación Biomédica de A Coruña, A Coruna, Galicia, Spain
[2]Grupo de Avances en Telemedicina e Informática Sanitaria (ATIS), Plataforma de Bioinformática, Complexo Hospitalario Universitario de A Coruña (CHUAC), Sergas, Universidade da Coruña (UDC), Instituto de Investigación Biomédica de A Coruña, A Coruna, Galicia, Spain
[3]Grupo GBTTC-CHUAC, Centro de Investigación Biomédica en Red Bioingeniería Biomateriales y Nanomedicina, Madrid, Spain
[4]Unidad de apoyo a la investigación, Grupo de Investigación en Enfermería y Cuidados en Salud, Complexo Hospitalario Universitario de A Coruña (CHUAC), Sergas, Universidade da Coruña (UDC), Instituto de Investigación Biomédica de A Coruña, A Coruna, Galicia, Spain
[5]Grupo de Investigación en Reumatología y Salud (GIR-S), Departamento de Fisioterapia, Medicina y Ciencias Biomédicas, Facultad de Fisioterapia, Campus de Oza, Universidade da Coruña, A Coruna, Galicia, Spain

Twitter Ignacio Rego-Perez @nacho_rego

Acknowledgements We would like to thank the participants, principal investigators, co-investigators and staff of CHECK. Part of the protein analysis has been performed by the ICTS "NANBIOSIS", specifically by the Proteomics Unit of the CIBER-BBN at SERGAS (Spain).

Contributors FJB and IR-P contributed equally in the design and coordination of the study; both conceived the study, participated in its design and helped to draft the final version of the manuscript; AD-S carried out the sequencing techniques and helped to draft the manuscript and data interpretation; JV-G helped to perform the sequencing techniques; MF-M carried out the cybrid experiments; VS-U and JF-T performed the bioinformatic analyses; SR and TH-G helped to carry out the cybrids cell culture; VB-B supervised the statistical procedures; LL-S and MCdA carried out the serum determinations of IL6; VC and PF-P carried out the proteomic analyses; CR-R supervised both serum and proteomic determinations; CV-G performed the autophagy experiments; NO helped the understanding of the clinical and radiological variables included in the different OAI, CHECK and PROCOAC datasets. All the authors approved the final version of the manuscript. IR-P and JFB are guarantors.

Funding The Osteoarthritis Initiative (OAI) is a public-private partnership comprising five contracts (N01-AR-2-2258; N01-AR-2-2259; N01-AR-2-2260; N01-AR-2-2261; N01-AR-2-2262) funded by the National Institutes of Health, a branch of the Department of Health and Human Services, and conducted by the OAI Study Investigators. Private funding partners include Pfizer; Novartis Pharmaceuticals; Merck Research Laboratories and GlaxoSmithKline. Private sector funding for the OAI is managed by the Foundation for the National Institutes of Health. CHECK-cohort study is funded by the Dutch Arthritis Association. It is led by a steering committee comprising 16 members with expertise in different fields of OA. It is chaired by Professor J W J Bijlsma and coordinated by J Wesseling, MSc. The following institutions are involved: Erasmus Medical Center Rotterdam; Kennemer Gasthuis Haarlem; Leiden University Medical Center; Maastrich University Medical Center; Martini Hospital Groningen/Allied Health Care Center for Rheum and Rehabilitation Groningen; Medical Spectrum Twente Enschede/Ziekenhuisgroep Twente Almelo; Reade, formerly Jan van Breemen Institute/VU Medical Center Amsterdam; St Maartens-kliniek Nijmegen; University Medical Center Utrecht and Wilhelmina Hospital Assen. ISCIII ('PI17/00210', 'PI19/01206', 'PI19/01213', 'PI20/00614', 'RICORS RD21/0002/0009'), co-funded by ERDF/ESF, 'A way to make Europe'/'Investing in your future'). Xunta de Galicia ('IN607A2021/07'). IR-P is supported by Contrato Miguel Servet-II Fondo de Investigación Sanitaria (CPII17/00026) estabilizado SERGAS. JV-G is supported by grant IN606A-2022/048 from Xunta de Galicia, Spain. LL-S is supported by contrato Sara Borrell Fondo de Investigación Sanitaria (CD19/00229). The Biomedical Research Networking Center (CIBER) is an initiative from ISCIII.

Disclaimer This manuscript was prepared using an OAI public use dataset and does not necessarily reflect the opinions or views of the OAI investigators, the NIH or the private funding partners.

Competing interests None declared.

Patient and public involvement Patients and/or the public were not involved in the design, or conduct, or reporting, or dissemination plans of this research.

Patient consent for publication Not applicable.

Ethics approval This study has received the favourable opinion of the ethics committee of XUNTA de Galicia, registration number 2017/451. All the clinical centres involved in the recruitment of the patients have made provisions to ensure the safety, confidentiality and ethical treatment of study participants according to the Declaration of Helsinki. In this sense, all the participants signed an informed consent. In the case of platelet donors, informed consent and the agreement of the ethical committee from Galician Health Administration were obtained.

Provenance and peer review Not commissioned; externally peer reviewed.

Data availability statement Data are available on reasonable request.

ORCID iDs
Cristina Ruiz-Romero http://orcid.org/0000-0001-7649-9803
Francisco J Blanco http://orcid.org/0000-0001-9821-7635
Ignacio Rego-Perez http://orcid.org/0000-0003-1754-1164

REFERENCES

1 Kraus VB, Blanco FJ, Englund M, *et al*. Call for standardized definitions of osteoarthritis and risk stratification for clinical trials and clinical use. *Osteoarthritis Cartilage* 2015;23:1233–41.
2 Loeser RF, Goldring SR, Scanzello CR, *et al*. Osteoarthritis: a disease of the joint as an organ. *Arthritis Rheum* 2012;64:1697–707.
3 Blanco FJ. "Osteoarthritis and atherosclerosis in joint disease". *Reumatol Clin (Engl Ed)* 2018;14:251–3.
4 Silverwood V, Blagojevic-Bucknall M, Jinks C, *et al*. Current evidence on risk factors for knee osteoarthritis in older adults: a systematic review and meta-analysis. *Osteoarthritis Cartilage* 2015;23:507–15.
5 Aubourg G, Rice SJ, Bruce-Wootton P, *et al*. Genetics of osteoarthritis. *Osteoarthritis Cartilage* 2022;30:636–49.
6 Emrani PS, Katz JN, Kessler CL, *et al*. Joint space narrowing and kellgren-lawrence progression in knee osteoarthritis: an analytic literature synthesis. *Osteoarthritis Cartilage* 2008;16:873–82.
7 Riddle DL, Stratford PW, Perera RA. The incident tibiofemoral osteoarthritis with rapid progression phenotype: development and validation of a prognostic prediction rule. *Osteoarthritis Cartilage* 2016;24:2100–7.
8 Driban JB, Harkey MS, Barbe MF, *et al*. Risk factors and the natural history of accelerated knee osteoarthritis: a narrative review. *BMC Musculoskelet Disord* 2020;21:332.
9 Driban JB, Stout AC, Lo GH, *et al*. Best performing definition of accelerated knee osteoarthritis: data from the osteoarthritis initiative. *Ther Adv Musculoskelet Dis* 2016;8:165–71.
10 Driban JB, Eaton CB, Lo GH, *et al*. Association of knee injuries with accelerated knee osteoarthritis progression: data from the osteoarthritis initiative. *Arthritis Care Res (Hoboken)* 2014;66:1673–9.
11 Felson D, Niu J, Sack B, *et al*. Progression of osteoarthritis as a state of inertia. *Ann Rheum Dis* 2013;72:924–9.
12 Lequesne M. Etiologic factors of rapid destructive coxarthrosis. *Presse Med* 1989;18:1231–3.
13 Mitrovic DR, Riera H. Synovial, articular cartilage and bone changes in rapidly destructive arthropathy (osteoarthritis) of the hip. *Rheumatol Int* 1992;12:17–22.
14 Yamamoto T, Schneider R, Iwamoto Y, *et al*. Rapid destruction of the hip joint in osteoarthritis. *Ann Rheum Dis* 2008;67:1783–4.
15 Blair-Levy JM, Watts CE, Fiorentino NM, *et al*. A type I collagen defect leads to rapidly progressive osteoarthritis in a mouse model. *Arthritis Rheum* 2008;58:1096–106.
16 Walker EA, Davis D, Mosher TJ. Rapidly progressive osteoarthritis: biomechanical considerations. *Magn Reson Imaging Clin N Am* 2011;19:283–94.
17 Mavrogenis AF, Flevas DA, Panagopoulos GN, *et al*. Rapid destructive arthritis of the hip revisited. *Eur J Orthop Surg Traumatol* 2015;25:1115–20.
18 Fang H, Liu X, Shen L, *et al*. Role of mtDNA haplogroups in the prevalence of knee osteoarthritis in a southern Chinese population. *Int J Mol Sci* 2014;15:2646–59.
19 Blanco FJ, Valdes AM, Rego-Pérez I. Mitochondrial DNA variation and the pathogenesis of osteoarthritis phenotypes. *Nat Rev Rheumatol* 2018;14:327–40.
20 Rego-Pérez I, Blanco FJ, Roemer FW, *et al*. Mitochondrial DNA haplogroups associated with MRI-detected structural damage in early knee osteoarthritis. *Osteoarthritis Cartilage* 2018;26:1562–9.
21 Koo BS, Song Y, Lee S, *et al*. Association of Asian mitochondrial DNA haplogroup B with new development of knee osteoarthritis in Koreans. *Int J Rheum Dis* 2019;22:411–6.

22 Zhao Z, Li Y, Wang M, *et al*. Mitochondrial DNA haplogroups participate in osteoarthritis: current evidence based on a meta-analysis. *Clin Rheumatol* 2020;39:1027–37.
23 Dobson PF, Dennis EP, Hipps D, *et al*. Mitochondrial dysfunction impairs osteogenesis, increases osteoclast activity, and accelerates age related bone loss. *Sci Rep* 2020;10:11643.
24 Scotece M, Rego-Pérez I, Lechuga-Vieco AV, *et al*. Mitochondrial DNA impact on joint damaged process in a conplastic mouse model after being surgically induced with osteoarthritis. *Sci Rep* 2021;11:9112.
25 Dalmao-Fernández A, Hermida-Gómez T, Lund J, *et al*. Mitochondrial DNA from osteoarthritic patients drives functional impairment of mitochondrial activity: a study on transmitochondrial cybrids. *Cytotherapy* 2021;23:399–410.
26 Cortés-Pereira E, Fernández-Tajes J, Fernández-Moreno M, *et al*. Differential association of mitochondrial DNA haplogroups J and H with the methylation status of articular cartilage: potential role in apoptosis and metabolic and developmental processes. *Arthritis Rheumatol* 2019;71:1191–200.
27 Fajardo RG, Fariña FO, Rey AM, *et al*. Relationship between the dynamics of telomere loss in peripheral blood leukocytes from knee osteoarthritis patients and mitochondrial DNA haplogroups. *J Rheumatol* 2021;48:1603–7.
28 Fernández-Moreno M, Soto-Hermida A, Oreiro N, *et al*. Mitochondrial haplogroups define two phenotypes of osteoarthritis. *Front Physiol* 2012;3:129.
29 Fernández-Moreno M, Rego-Pérez I, Blanco FJ. Is osteoarthritis a mitochondrial disease? What is the evidence. *Curr Opin Rheumatol* 2022;34:46–53.
30 Wesseling J, Dekker J, van den Berg WB, *et al*. Check (cohort hip and cohort knee): similarities and differences with the osteoarthritis initiative. *Ann Rheum Dis* 2009;68:1413–9.
31 Oreiro-Villar N, Raga AC, Rego-Pérez I, *et al*. PROCOAC (prospective cohort of A coruñA) description: Spanish prospective cohort to study osteoarthritis. *Reumatol Clin (Engl Ed)* 2022;18:100–4.
32 Andrews RM, Kubacka I, Chinnery PF, *et al*. Reanalysis and revision of the Cambridge reference sequence for human mitochondrial DNA. *Nat Genet* 1999;23:147.
33 Li H, Homer N. A survey of sequence alignment algorithms for next-generation sequencing. *Brief Bioinform* 2010;11:473–83.
34 Fernández-Moreno M, Soto-Hermida A, Vázquez-Mosquera ME, *et al*. Mitochondrial DNA haplogroups influence the risk of incident knee osteoarthritis in OAI and check cohorts. A meta-analysis and functional study. *Ann Rheum Dis* 2017;76:1114–22.
35 Bushel PR, Ward J, Burkholder A, *et al*. Mitochondrial-nuclear epistasis underlying phenotypic variation in breast cancer pathology. *Sci Rep* 2022;12:1393.
36 Boer CG, Hatzikotoulas K, Southam L, *et al*. Deciphering osteoarthritis genetics across 826,690 individuals from 9 populations. *Cell* 2021;184:6003–5.
37 Navaglia F, Basso D, Fogar P, *et al*. Mitochondrial DNA D-loop in pancreatic cancer: somatic mutations are epiphenomena while the germline 16519 T variant worsens metabolism and outcome. *Am J Clin Pathol* 2006;126:593–601.
38 Zaki EA, Freilinger T, Klopstock T, *et al*. Two common mitochondrial DNA polymorphisms are highly associated with migraine headache and cyclic vomiting syndrome. *Cephalalgia* 2009;29:719–28.
39 van Tilburg MAL, Zaki EA, Venkatesan T, *et al*. Irritable bowel syndrome may be associated with maternal inheritance and mitochondrial DNA control region sequence variants. *Dig Dis Sci* 2014;59:1392–7.
40 Wang Y, Huang X, Peng F, *et al*. Association of variants m.T16172C and m.T16519C in whole mtDNA sequences with high altitude pulmonary edema in Han Chinese lowlanders. *BMC Pulm Med* 2022;22:72.
41 Boles RG, Zaki EA, Kerr JR, *et al*. Increased prevalence of two mitochondrial DNA polymorphisms in functional disease: are we describing different parts of an energy-depleted elephant? *Mitochondrion* 2015;23:1–6.
42 Liao WQ, Pang Y, Yu CA, *et al*. Novel mutations of mitochondrial DNA associated with type 2 diabetes in Chinese Han population. *Tohoku J Exp Med* 2008;215:377–84.
43 Akouchekian M, Houshmand M, Hemati S, *et al*. High rate of mutation in mitochondrial DNA displacement loop region in human colorectal cancer. *Dis Colon Rectum* 2009;52:526–30.
44 Bai R-K, Leal SM, Covarrubias D, *et al*. Mitochondrial genetic background modifies breast cancer risk. *Cancer Res* 2007;67:4687–94.
45 Läer K, Vennemann M, Rothämel T, *et al*. Mitochondrial deoxyribonucleic acid may play a role in a subset of sudden infant death syndrome cases. *Acta Paediatr* 2014;103:775–9.
46 Murakami H, Ota A, Simojo H, *et al*. Polymorphisms in control region of mtDNA relates to individual differences in endurance capacity or trainability. *Jpn J Physiol* 2002;52:247–56.
47 Camilleri M, Carlson P, Zinsmeister AR, *et al*. Mitochondrial DNA and gastrointestinal motor and sensory functions in health and functional gastrointestinal disorders. *Am J Physiol Gastrointest Liver Physiol* 2009;296:G510–6.
48 Ruiz-Pesini E, Lott MT, Procaccio V, *et al*. An enhanced MITOMAP with a global mtDNA mutational phylogeny. *Nucleic Acids Res* 2007;35:D823–8.
49 Stoccoro A, Coppedè F. Mitochondrial DNA methylation and human diseases. *Int J Mol Sci* 2021;22:4594.
50 Bianchetti R, Lucchini G, Crosti P, *et al*. Dependence of mitochondrial protein synthesis initiation on formylation of the initiator methionyl-tRNAf. *J Biol Chem* 1977;252:2519–23.
51 Cai N, Gomez-Duran A, Yonova-Doing E, *et al*. Mitochondrial DNA variants modulate N-formylmethionine, proteostasis and risk of late-onset human diseases. *Nat Med* 2021;27:1564–75.
52 Quirós PM, Mottis A, Auwerx J. Mitonuclear communication in homeostasis and stress. *Nat Rev Mol Cell Biol* 2016;17:213–26.
53 Zhang R, Wang Y, Ye K, *et al*. Independent impacts of aging on mitochondrial DNA quantity and quality in humans. *BMC Genomics* 2017;18:890.
54 Gu F, Chauhan V, Kaur K, *et al*. Alterations in mitochondrial DNA copy number and the activities of electron transport chain complexes and pyruvate dehydrogenase in the frontal cortex from subjects with autism. *Transl Psychiatry* 2013;3:e299.
55 Guo W, Yang D, Xu H, *et al*. Mutations in the D-loop region and increased copy number of mitochondrial DNA in human laryngeal squamous cell carcinoma. *Mol Biol Rep* 2013;40:13–20.
56 Eirin A, Saad A, Tang H, *et al*. Urinary mitochondrial DNA copy number identifies chronic renal injury in hypertensive patients. *Hypertension* 2016;68:401–10.
57 Wang Y, Zhao X, Lotz M, *et al*. Mitochondrial biogenesis is impaired in osteoarthritis chondrocytes but reversible via peroxisome proliferator-activated receptor γ coactivator 1α. *Arthritis Rheumatol* 2015;67:2141–53.
58 Gomes LC, Di Benedetto G, Scorrano L. During autophagy mitochondria elongate, are spared from degradation and sustain cell viability. *Nat Cell Biol* 2011;13:589–98.
59 Zhao T, Huang X, Han L, *et al*. Central role of mitofusin 2 in autophagosome-lysosome fusion in cardiomyocytes. *J Biol Chem* 2012;287:23615–25.
60 Kim D, Song J, Kang Y, *et al*. Fis1 depletion in osteoarthritis impairs chondrocyte survival and peroxisomal and lysosomal function. *J Mol Med (Berl)* 2016;94:1373–84.
61 Henrotin YE, Bruckner P, Pujol J-PL. The role of reactive oxygen species in homeostasis and degradation of cartilage. *Osteoarthritis Cartilage* 2003;11:747–55.
62 Reed KN, Wilson G, Pearsall A, *et al*. The role of mitochondrial reactive oxygen species in cartilage matrix destruction. *Mol Cell Biochem* 2014;397:195–201.
63 Bolduc JA, Collins JA, Loeser RF. Reactive oxygen species, aging and articular cartilage homeostasis. *Free Radic Biol Med* 2019;132:73–82.
64 Ruiz-Romero C, Carreira V, Rego I, *et al*. Proteomic analysis of human osteoarthritic chondrocytes reveals protein changes in stress and glycolysis. *Proteomics* 2008;8:495–507.
65 Ko JY, Sun YC, Li WC, *et al*. Chaperonin 60 regulation of SOX9 ubiquitination mitigates the development of knee osteoarthritis. *J Mol Med (Berl)* 2016;94:755–69.
66 Dai J, Zhang Y, Chen D, *et al*. Glabridin inhibits osteoarthritis development by protecting chondrocytes against oxidative stress, apoptosis and promoting mtor mediated autophagy. *Life Sci* 2021;268:118992.
67 Yin M, Xu Y. The protective effects of etomidate against interleukin-1β (IL-1β) -induced oxidative stress, extracellular matrix alteration and cellular senescence in chondrocytes. *Bioengineered* 2022;13:985–94.
68 Robinson WH, Lepus CM, Wang Q, *et al*. Low-Grade inflammation as a key mediator of the pathogenesis of osteoarthritis. *Nat Rev Rheumatol* 2016;12:580–92.
69 Loeser RF, Collins JA, Diekman BO. Ageing and the pathogenesis of osteoarthritis. *Nat Rev Rheumatol* 2016;12:412–20.
70 Sanchez-Lopez E, Coras R, Torres A, *et al*. Synovial inflammation in osteoarthritis progression. *Nat Rev Rheumatol* 2022;18:258–75.
71 Wiegertjes R, van de Loo FAJ, Blaney Davidson EN. A roadmap to target interleukin-6 in osteoarthritis. *Rheumatology (Oxford)* 2020;59:2681–94.
72 Fernández-Moreno M, Soto-Hermida A, Vázquez-Mosquera ME, *et al*. A replication study and meta-analysis of mitochondrial DNA variants in the radiographic progression of knee osteoarthritis. *Rheumatology (Oxford)* 2017;56:263–70.
73 Latorre-Pellicer A, Moreno-Loshuertos R, Lechuga-Vieco AV, *et al*. Mitochondrial and nuclear DNA matching shapes metabolism and healthy ageing. *Nature* 2016;535:561–5.
74 Geurts J, Nasi S, Distel P, *et al*. Prematurely aging mitochondrial DNA mutator mice display subchondral osteopenia and chondrocyte hypertrophy without further osteoarthritis features. *Sci Rep* 2020;10:1296.
75 Lechuga-Vieco AV, Latorre-Pellicer A, Johnston IG, *et al*. Cell identity and nucleo-mitochondrial genetic context modulate OXPHOS performance and determine somatic heteroplasmy dynamics. *Sci Adv* 2020;6:eaba5345.
76 Conaghan PG, Kloppenburg M, Schett G, *et al*. Osteoarthritis research priorities: a report from a EULAR AD hoc expert Committee. *Ann Rheum Dis* 2014;73:1442–5.
77 van Helvoort EM, van Spil WE, Jansen MP, *et al*. Cohort profile: the applied public-private research enabling osteoarthritis clinical headway (IMI-APPROACH) study: a 2-year, European, cohort study to describe, validate and predict phenotypes of osteoarthritis using clinical, imaging and biochemical markers. *BMJ Open* 2020;10:e035101.

EXTENDED SUPPLEMENTARY METHODS

Cohort's description

The OAI cohort has a public archive of data, biological samples and images collected over time from a clinically well-characterized population. Participants from this cohort used in the present study were followed annually for 8 years to detect changes in the clinical status of their knees (and other joints), including worsening and onset of symptoms and disabilities, worsening and onset of knee structural abnormalities, as well as changes in other imaging and biochemical markers of knee OA.

The CHECK cohort was formed from October 2002 to September 2005. The cohort consists of participants with pain and/or stiffness of the knee and/or hip, which were followed prospectively. Individuals were eligible if they had pain or stiffness of knee or hip, aged between 45-65 years and had no yet consulted their physician for these symptoms, or the first consultation was within six months before entry.

The PROCOAC cohort is a longitudinal population-based cohort study on the determinants and prognosis of OA of the knees, hips and hands. The recruitment of patients started in 2006, and all patients are followed bi-annually. Patients were included attending to the following: i) patients from external consultations with hand pain and diagnosed with hand OA following the American College of Rheumatology (ACR) criteria; ii) patients with knee pain previously diagnosed with radiographic knee OA following ACR criteria; and iii) patients with hip pain previously diagnosed with radiographic hip OA following ACR criteria.

In-depth sequencing of mtDNA

Library preparation

Durán-Sotuela A, *et al. Ann Rheum Dis* 2023;0:1–11. doi: 10.1136/ard-2022-223570

Sequencing libraries were automatically generated using Precision ID DL8 Kit (ThermoFisher Scientific) for the Ion Chef system (ThermoFisher Scientific). This kit enables the preparation of Ion AmpliSeq libraries for the Precision ID mtDNA Whole Genome Panel (ThermoFisher Scientific). This panel consists of two primer pools, of 81 primer pairs each, that targets the entire human mitochondrial genome. Library preparation was done following the manufacturer´s recommended protocols. Briefly, gDNA samples were quantified using a Qubit device to reach a final concentration of 0.1ng of gDNA in a final volume of 15uL. The Ion Chef system is able to automatically prepare up to 8 libraries (samples) per run, resulting in a final pool of 700uL at 100pM in one tube. The maximum number of barcodes available using this technique is 32, so that we performed 4 independent runs on the Ion Chef to generate 4 pools of libraries (8 per pool). Finally, each set of (8x4) 32 libraries was pooled into a single tube in a final volume of 25ul and normalized to 30pM, ready for templating and sequencing.

Templating and sequencing

The generated libraries are ready for template preparation for clonal amplification on the Ion Chef system and sequencing on an Ion S5XL system. Once the libraries are normalized to 30pM, the templating process was performed automatically on the Ion Chef system following the manufacturer´s recommended protocols. Briefly, this process includes the joining of the libraries to the Ion Sphere Particles (ISP), in such a way that only one fragment will join one ISP, and the subsequent clonal amplification by emulsion PCR, by which each ISP will be covered by multiple copies of the same fragment. In addition, through an enrichment process, all empty ISPs are removed.

Data analysis

Ion S5XL reads were processed with Ion Torrent Suite™ software v 5.12 (ThermoFisher Scientific) for variant calling using an adapted pipeline for Ampliseq libraries using a low stringency mode in Torrent Variant Caller plug-in (TVC) (OCPv1i3 Panel - 530 - Chef - Somatic - Low Stringency). In order to confirm that results are not contingent to the presence of Nuclear Mitochondrial DNA Segments (NUMTs), a parallel pipeline was carried out, using the same steps but adding a previous clean-up step of potentially NUMT-derived reads using standalone IonTorrent tools from publicly available docker image (iontorrent/tsbuild). Clean-up consisted in re-aligning reads against both the reference nuclear genome (GRCh28, chromosomes 1-23) and the mitochondrial genome (revised Cambridge reference sequence (rCRS)).[35] For this approach, we used TMAP alignment optimization [36] with the same parameters used in the regular Ion Torrent Suite™ process in order to obtain reference-specific alignment score (AS) values to be compared. Those reads with a higher AS in the alignment against the nuclear genome reference than that obtained against the rCRS reference, have been tagged as potentially NUMT-derived and subsequently removed from the BAM file using Picard tools before variant calling using TVC.

Additionally, given the circularity of the mitochondrial genome, in order to avoid misalignments towards the ends of the rCRS reference, reason why the standard Ion Torrent Suite™ uses the PrecisionID_mtDNA_rCRS reference repeats the first 80 nucleotides of rCRS at the end of the sequence, the alignment is carried out again using a version of rCRS with an alternative cut point. In this alternative version of rCRS, the ends of the original reference sequence rCRS

are joined and the cut point is shifted to between nucleotides 8284 y 8285, which allows the confirmation of variants identified in positions less than 500 bps away from the ends of the original reference sequence.

All this process has been automated using the workflow management system snakemake within the same tsbuild docker container that provided the IonTorrent-specific tools (TMAP and TVC).

The resultant mitochondrial variants obtained after applying the above-described workflow were initially filtered by their population frequency. Thereby, those mitochondrial genetic variants with a minor allele frequency (MAF) <0.05 were discarded. In addition, a conservative heteroplasmy detection threshold was set to 20%. Even so, after automatic variant calling, manual verification was performed using Integrative Genomics Viewer (IGV) software v2.8.0.

Meta-analysis

In this work, we performed a subsequent meta-analysis of rapid progression of KOA including data from the American cohort (discovery cohort) and the European cohort (replication cohort). Although we followed PRISMA guidelines to perform the meta-analysis, the computerized search strategy did not find studies of rapid progression of KOA including the mtDNA variant described in this work, therefore, no additional studies were incorporated to the meta-analysis.

Both fixed and random effect models were considered for meta-analysis with binary outcome (rapid vs non-rapid progressors) data. Inverse variance method was used to calculate the fixed and random effects estimates. Estimation of adjusted ORs and their 95%CI was used as the effect size measurements for the association between m.16519C and the rapid progression of KOA. The I^2

index was computed to explore the heterogeneity. Meta-analysis results were presented on a forest plot graph. All the computations were carried out using R software v4.0.2, using the meta-package. Meta-analysis was planned to be performed on k=2 studies, with an estimated statistical power of 90.3% to detect as statistically significant an OR ≥1.5 associated to the presence of m.16519C, with a p-value <0.05 two-tailed significance level.

Single Base Extension (SBE) assay

The assignment of the mtDNA variant with a significantly different frequency between rapid and non-rapid progressors, obtained in a subset of patients from the discovery cohort, was carried out in the rest of the samples using the SBE technique on a SeqStudio genetic analyzer system (ThermoFisher Scientific). Briefly, an initial polymerase chain reaction (PCR) was performed to amplify the mtDNA fragment containing the informative SNP. The resulting PCR fragment was purified with ExoSap-It (ThermoFisher Scientific) and subsequently subjected to the SBE reaction following manufacturer´s recommendations. Purified SBE product using Fast-AP (ThermoFisher Scientific) was loaded into a SeqStudio genetic analyzer to visualize the SNP. Primers used for this assay are available upon request.

This technique was also used to confirm the selected mtDNA variant m.16519T>C in a 20% of randomly selected samples form the initial subset from the discovery cohort of the OAI.

Functional studies using transmitochondrial cybrids

RNA isolation and gene expression assays

Cells from cybrid cultures were pelleted and the RNA was isolated using Trizol ® (Thermo Fisher Scientific). Isolated RNA was quantified using NanoDrop

ND1000 (Thermo Scientific) and subsequently reverse transcribed (1µg) using SuperScript VILO Master Mix ® (Thermo Fisher Scientific) following manufacturer recommendations.

Quantitative real time PCR experiments were performed on a LightCycler 480 II system (Roche) using Universal Probe Library (UPL) probes to assess the gene expression of genes related to: i) mitochondrial biogenesis (peroxisome proliferator-activated receptor gamma, coactivator 1 alpha (*PPARGC1A*) and transcription factor A mitochondrial (*TFAM*)); and ii) mitochondrial fusion (mitofusin 2 (*MNF2*)) and fission (fission mitochondrial 1 (*FIS1*)) processes. The results were analyzed with qBase+ software v2.5 (Biogazelle) using the Glyceraldehyde-3-Phosphate Dehydrogenase (*GAPDH*) as housekeeping gene. Primer and probe sequences are available upon request.

RNA-seq

A whole transcriptome assay by next generation sequencing (NGS) techniques was performed in cybrids with and without the mtSNV m.16519C. Isolated RNA was quantified using a Qubit 4 flourometer (ThermoFisher). DNAse-treated isolated RNA from cybrids was retro-transcribed using SuperScript™ VILO™ cDNA Synthesis Kit (Thermofisher) using manufacturer recommendations. Then, 10ng of cDNA were subsequently used to automatically generate libraries using the Ion AmpliSeq™ Transcriptome Human Gene Expression Panel, Chef-Ready Kit (ThermoFisher). After the templating process, templated ISPs were automatically loaded onto one Ion 550 chip and placed on an Ion S5XL for sequencing, and subsequently post-processed with the ampliSeqRNA plug-in of the Torrent Suite™ software v 5.12 (ThermoFisher Scientific).

CHP files, normalized by reads per million (RPM), were automatically generated and imported into the Transcriptome Analysis Console (TAC) software v4.0 (ThermoFisher) to perform the differential expression analyses between cybrids with and without the mtSNV m.16519C. Only those genes with a Fold change (Fc) between groups >±2 and a corrected (FDR) p-value <0.05 were considered as differentially expressed.

Finally, to explore the biological meaning of the differentially expressed genes, we loaded both the list of up-regulated and down-regulated genes associated with the presence of the mtSNV m.16519C into the online Enrichr web-based tool (https://maayanlab.cloud/Enrichr/).

Subsequent quantitative real time PCR validation of Interleuquin 6 (*IL6*) and prostaglandin-endoperoxide synthase 2 (*PTGS2*) genes was performed on a LightCycler 480 II system (Roche) using dual labeled Taqman probes (ThermoFisher Scientific), as well as GAPDH as reference gene

Mitochondrial reactive oxygen species (ROS) production assay

Mitochondrial superoxide anion was evaluated with MitoSox Red (Thermo Fisher Scientific) at a final concentration of 5μM in Hank´s Balanced Sal Solution (HBSS, Sigma) during 15 min at 37ºC in darkness.

Cells were harvested by trypsin and re-suspended in saline solution prior to be analyzed by flow cytometry. A density of $1x10^4$ cells per assay were measured by flow cytometry and data were analyzed with CellQuest software (Becton Dickinson). Results were expressed as median of fluorescence intensity.

Oxidative stress response assay

To evaluate the oxidative stress response under H_2O_2 incubation, the CellTiter 96® Aqueous Assay kit (Promega) was used. $2x10^3$ cells were grown in

MW-96 for 24h in DMEM 10% FBS. Then, the medium was replaced by DMEM without serum, and maintained in culture during 24h. Next, 200µM H_2O_2 was added during 30 min and finally 20µl of kit reagents of warm solution (3-(4,5-dimethythiazol-2-yl)-5-(3-carboxymethoxyphenyl)-2-(4-sulfophenyl)-2H-tetrazolium (MTS) + Phenazine methosulfate (PMS)) were also added to all wells and incubated 4h at 37°C followed by measurement of absorbance at 490nm in a microplate reader (NanoQuant Infinite M200, Tecan).

To examine for oxidized proteins relevant to mitochondrial stress, we analyzed post-translational modifications (PTMs) by liquid chromatography-tandem mass spectrometry (LC-MS/MS) in two clones from each type of cybrids. The analysis was performed following a classic workflow of nanoliquid chromatography coupled to a high resolution TIMS TOF Pro mass spectrometer (Bruker). Then, mass spectrometry raw files were processed with PEAKSXpro 10-6 software (Bioinformatics solutions Inc.) against the UniProtKB/Swiss-Prot human database (release 2021_02) previously filtered by mitochondrial proteins. For the quantification of PTMs, only mitochondrial oxidized proteins with an AScore >20 (p-value <0.01) were considered (Beausoleil SA, Villén J, Gerber SA, et al. A probability-based approach for high-throughput protein phosphorylation analysis and site localization. Nat Biotechnol 2006;24:1285–92).

Sample preparation and LC-MS/MS analysis

Equal amounts of two transmitochondrial cybrids (509 and 513) from two different clones were reduced with 10 mM dithiothreitol for 1 h at 37 °C, and subsequently alkylated with 50 mM iodoacetamide for 45 minutes at room temperature in the dark. Samples were digested with sequencing grade modified trypsin (Promega) at 1:40 enzyme-to-substrate ratio. After 16 h of digestion at 37

°C, samples were acidified with 10% trifluoroacetic acid to ~pH 3. The digested peptides were desalted using in-house made stage tips (3M Empore SPE-C18 disk,47 mm, Sigma Aldrich) and finally samples were dried under speed-vacuum (Thermo, USA). The dried eluates were re-constituted in 0.1% FA (formic acid) for direct LC-MS. The peptide mixture (200ng) was loaded in a nanoElute (Bruker Daltonics) nano-flow LC was coupled to a high-resolution TIMS-QTOF (timsTOF Pro, Bruker Daltonics) with a CaptiveSpray ion source (Bruker Daltonics). Technical replicates for each sample were injected. Liquid chromatography was performed at 50 °C and with a constant flow of 400 nL/min on a reversed-phase column (15 cm * 75 m i.d.) with a pulled emitter tip, packed with 1.9 m C18-coated porous silica beads (Dr. Maisch, Ammerbuch-Entringen, Germany). Chromatographic separation was carried out using a linear gradient of 5-35% buffer B (100% ACN and 0.1% FA) over 50 min. After ESI ionization, peptides were analyzed in data-dependent mode with Parallel Accumulation–Serial Fragmentation (DDA-PASEF) enabled.

Analysis of mass spectrometry data

Mass spectrometry raw files were processed with PEAKSXpro 10.6 build 20201221 (Bioinformatics solutions Inc.). The MS/MS spectra were matched to in silico derived fragment mass values of tryptic peptides against an in-house mitochondrial database (1136 entries) generated by filtering the UniProtKB/Swiss-Prot human database according to MitoCarta 3.0 database (MitoCarta3.0: An Inventory of Mammalian Mitochondrial Proteins and Pathways | Broad Institute). Search parameters were: Parent Mass Error Tolerance: 15.0 ppm; Fragment Mass Error Tolerance: 0.02 Da; Enzyme: Trypsin; Variable Modifications: Acetylation (Protein N-term): 42.01; Deamidation (NQ): 0.98;

Oxidation (M): 15.99; Oxidation (HW): 15.99; Oxidation or Hydroxylation (C): 15.99; Oxidation or Hydroxylation: 15.99; Persulfide: 31.97 and 27 more. Fixed Modification: Carbamidomethylation: 57.02. PEAKS PTM module allows screening more than 30 known PTM types and supports quantitative PTM analysis.

Mitochondrial membrane potential (ΔΨm)

We analyzed the mitochondrial membrane potential (ΔΨm) using the fluorescent Mito Probe JC-1 in cybrids 509 (m.16519T) and 513 (m.16519C) by flow cytometry. Mito Probe JC-1 is a lipophilic cation that exists as a monomer at low values of ΔΨm (green fluorescence), while it forms aggregates at high ΔΨm (red fluorescence).

For this procedure, cybrids cells were seeded at 8×104 cells per well in 12-well plates in DMEM 10% FBS; when the cell confluence arrived near 80%, the medium was changed to DMEM serum-free and treated with the protonophore carbonylcyanide-p-trifluoromethoxyphenylhydrazone (FCCP) at 500 μM during 3h as depolarizer inductor. Cybrids were collected by trypsinization and then incubated with 2μM of Mito Probe JC-1 for 30 min at 37°C in 5% CO2, sedimented, washed in PBS and analyzed by flow cytometry using CytoFLEX (Beckman Coulter). Mean fluorescence intensity values of each channel, expressed as arbitrary units, were obtained for all experiments. In each experiment, at least 5000 events were analyzed. The percentage of depolarization in a cell population was calculated by dividing the values of red fluorescence by the values of green fluorescence. Therefore, the ratio aggregates:monomers (red:green fluorescence) in relation to the FCCP ratio corresponding to each cybrid was subsequently used to present the data.

Autophagy determination

Autophagy was evaluated through the quantification of Microtubule-associated protein 1A/1B-light chain 3 (LC3) (*I and II*), Sequestosome-1 (p62/SQSTM1) and phospho-ribosomal protein S6 (p-rpS6) by Western blotting (WB).

Total protein was obtained from 6 x 10^5 cells lysed using 200µl of lysis buffer (6M urea/2% SDS buffer) and then were sonicated at 4 ºC. The homogenate was centrifuged and the supernatant was recovered. The protein concentrations in the supernatant were determined by a Bicinchoninic Acid reagent assay (BCA) using pierce BCA Protein Assay kit (Pierce Biotechnology, IL, USA). The samples were then adjusted to equal concentration (10µg) before resolution on 4–20% gels (Invitrogen). Protein was transferred onto nitrocellulose membranes (Invitrogen) and then blocked for 1h at room temperature (rt) with BSA 5% for LC3, p62/SQSTM1 and p-rpS6, or with 5% dry milk in the case of α-tubulin in Tris buffered saline–Tween (TBST). The membranes were incubated with antibodies for LC3-II (1:1000, Cell Signaling Technology, Beverly, MA, Cat. Number #3868), p62/SQSTM1 (1:2000, Cell Signaling Technology, Cat. Number #88588), p-rpS6 (1:2000, Cell Signaling Technology, Cat. Number #4858) and α-tubulin (1:5000, Sigma-Aldrich, Cat. Number #T9026). The membranes were then incubated overnight at 4 ºC with horseradish peroxidase (HRP)–conjugated anti-rabbit IgG or anti-mouse IgG (1:1000 and 1:10000 respectively. Sigma-Aldrich Cat Numbers #NA934 and #NA931, respectively) for 1h at rt. Then, the membranes were washed 3 times with TBST and developed using an enhanced chemiluminescent substrate (Pierce Biotechnology). Cells were incubated with 30µM chloroquine (Sigma-Aldrich) for 16h as positive control for LC3 and

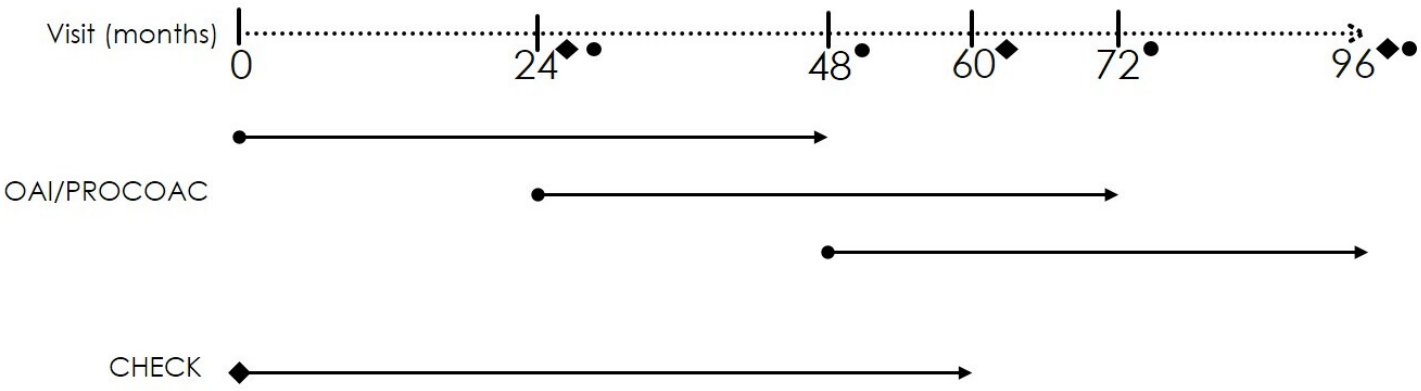
p62/SQSTM1 accumulation and, in the case of p-rpS6, cells were treated with 10μM Rapamycin (Calbichem, Darstadt, Germany) during 16h. The intensity of the bands and the WB quantification were assessed by using Amershan Imagen 600 software.

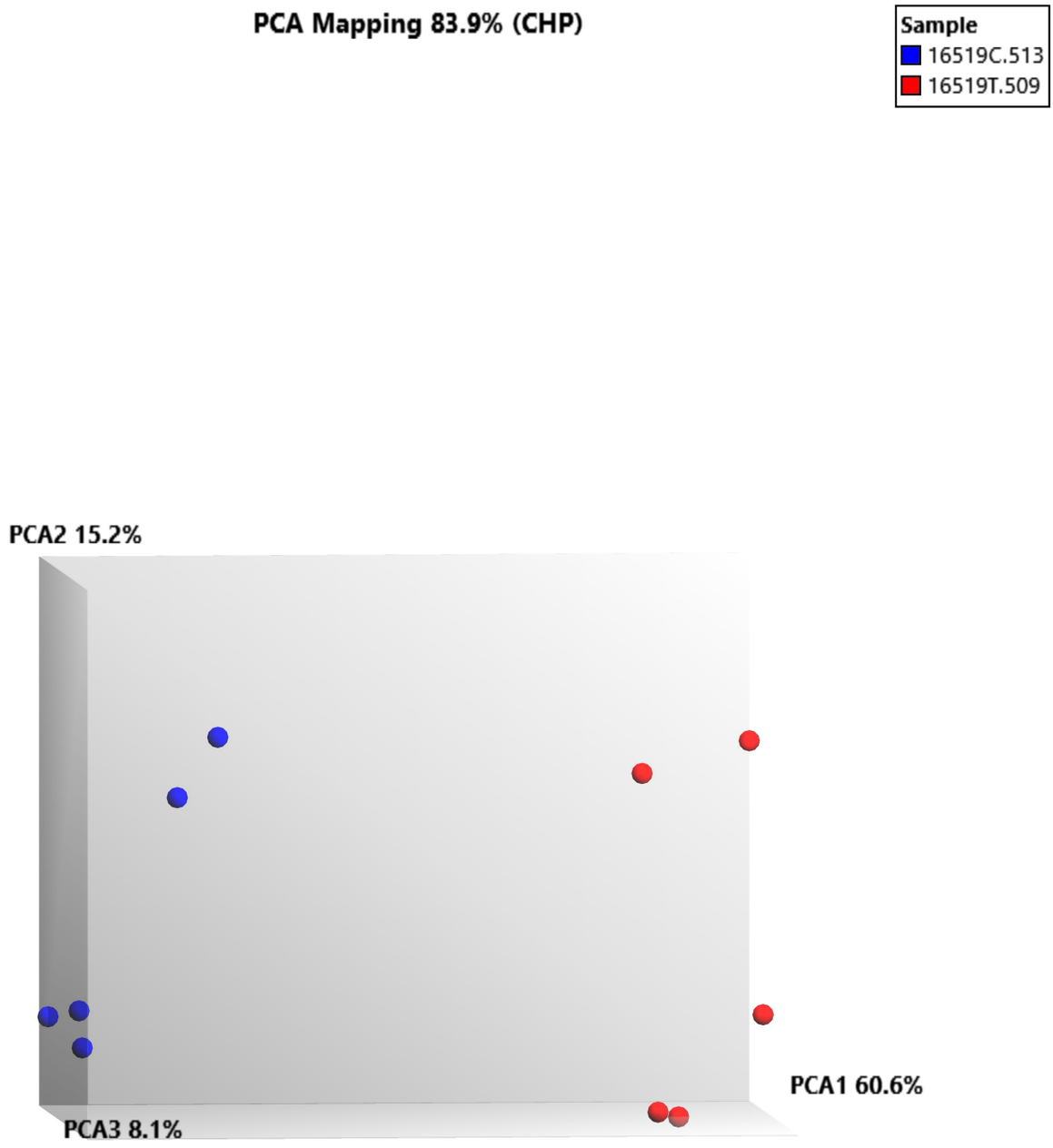
Statistical analysis of cybrids experiments

Results were expressed as the mean of different independent measures, with a minimum of three observations per clone (mean±SEM), using two cybrids (one H cybrid harboring the mtDNA variant m.16519C and one H cybrid harboring the mtDNA reference m.16519T) and two clones from each cybrid. Statistically significant differences between the two groups were determined using Mann-Whitney non-parametric U test. P-values below 0.05 were considered significant.

Preliminary determination of serum levels of IL6

We compared the baseline serum levels of IL6 between (N=26) rapid and (N=200) non-rapid progressors belonging to the OAI. Millipore's MILLIPLEX MAP High Sensitivity Human Cytokine multiplex kit was used to measure serum concentrations of IL6 following the manufacturer's instructions. Serum samples and standards were assayed in duplicate and measured using the FlexMap3D instrument (Luminex, Corp.). We adjusted for batch in all analyses. The limit of detection was 0.10 pg/mL. Intra-assay coefficient of variation for the serum inflammatory marker studied was less than 5%; inter-assay coefficient of variation was less than 19%. Obtained mean and median measures were categorized based on the median and statistically analyzed following a logistic regression model adjusting for the confounder variables of age, gender, BMI, contralateral knee OA and WOMAC pain at baseline.

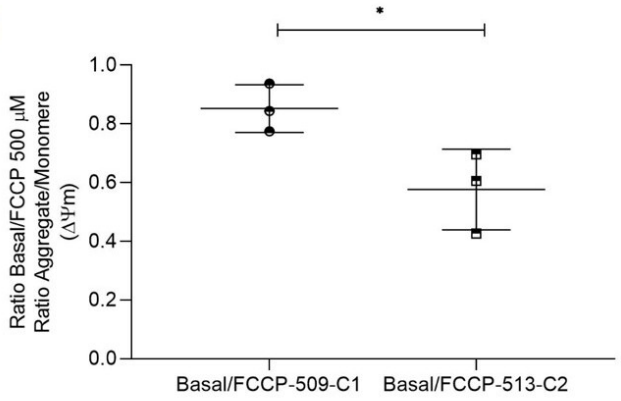

Visit (months)
0
24
48
60
72
96
OAI/PROCOAC
CHECK

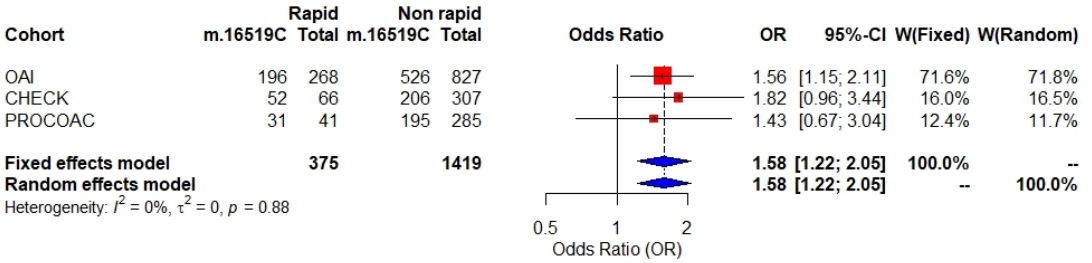

PCA Mapping 83.9% (CHP)
Sample
16519C.513
16519T.509
PCA2 15.2%
PCA1 60.6%
PCA3 8.1%

A


513
509
log2(ratio)
P61604_CH10 – M32d
P10809_CH60 – N482d
P38646_GRP75 – N149d
P30048_PRDX3 – N150d
4.0
0.0
-4.0

B


*
Ratio Basal/FCCP 500 µM
Ratio Aggregate/Monomere
(ΔΨm)
1.0
0.8
0.6
0.4
0.2
0.0
Basal/FCCP-509-C1
Basal/FCCP-513-C2


Cohort
Rapid m.16519C Total
Non rapid m.16519C Total
Odds Ratio
OR 95%-CI W(Fixed) W(Random)
OAI 196 268 526 827 1.56 [1.15; 2.11] 71.6% 71.8%
CHECK 52 66 206 307 1.82 [0.96; 3.44] 16.0% 16.5%
PROCOAC 31 41 195 285 1.43 [0.67; 3.04] 12.4% 11.7%
Fixed effects model 375 1419 1.58 [1.22; 2.05] 100.0% --
Random effects model 1.58 [1.22; 2.05] -- 100.0%
Heterogeneity: I^2 = 0%, τ^2 = 0, p = 0.88
0.5
1
2
Odds Ratio (OR)