EXTENDED SUPPLEMENTARY METHODS

Part of the data used in the preparation of this manuscript was obtained from the Osteoarthritis Initiative (OAI) database, which is available for public access at http://www.oai.ucsf.edu/. Specific datasets used are the 0.2.2 clinical dataset, 22 enrollees' dataset and 0.7, 1.7, 3.6, 5.6, 6.4, 8.1 and 10.1 X-ray imaging datasets.

Incident knee OA study in subjects from the OAI and CHECK

OAI Participants

The OAI cohort has a public archive of data, biological samples and joint images collected over time from a clinically well characterized population of individuals comprising two subgroups, i) those with clinically significant knee OA who are at risk of disease progression (progression subcohort); and ii) individuals who are at high risk for developing clinically significant knee OA (incidence subcohort). Participants are followed for eight years for changes in the clinical status of the knee and other joints, including worsening and onset of symptoms and disabilities, worsening and onset of knee structural abnormalities, and changes in other imaging and biochemical markers of knee OA.

For this study, we included longitudinal data in terms of Kellgren and Lawrence (KL) grade of 2579 OAI participants of the incidence subcohort during a follow-up period of eight years.

All the clinical centers of the OAI 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 OAI participants signed an informed consent.

CHECK participants

From October 2002 to September 2005 a cohort was formed of 1002 participants with pain and/or stiffness of the knee and/or hip, which is to be followed prospectively for a period of at least ten years. The study was approved by the medical ethics committees of the 12 participating centers, and all participants gave their written informed consent before entering the study. Individuals were eligible if they had pain or stiffness of knee or hip, were 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.

For this study, we included longitudinal data in terms of KL grade of 635 CHECK participants that met the eligibility criteria for an incident knee OA study monitored for a period of eight years.

mtDNA haplogroups genotyping

A multiplex polymerase chain reaction (PCR) was performed to amplify six mtDNA fragments that contain the informative SNPs that characterize the most common Caucasian mtDNA haplogroups (H, UK, J and T), as well as the less common ones (<3.5%) pooled into the group “Others”. The resulting PCR fragments were purified and analyzed by Single Base Extension (SBE) assay to further visualize the informative SNPs after loading the purified SBE product into an ABI 3130XL genetic analyzer (Applied Biosystems, Foster City, CA, United States). The assigned mtDNA haplogroups were verified by direct sequencing, in 10% of the samples, of regions that include key SNPs described in phylotree (http://www.phylotree.org).

Statistical analysis

All the statistical analyses were performed using IBM-SPSS software, release 19 (IBM, Armonk, NY, USA) and R software v3.1.2 (The R Foundation for Statistical Computing). All comparisons were two-sided, with $P<0.05$ defined as statistically significant.

Because the status of a patient was prospectively evaluated at predefined intervals (baseline, 1, 2, 3, 4, 6 and 8 years), the precise date at which incident knee OA occurred could not be determined; it always occurred between visits. Only interval during which the conversion occurred was observed, therefore such information is said to be interval-censored. Thus, to avoid potential biases associated with the use of standard survival analysis in this context, interval-censored data analysis methods were used.

Turnbull´s extension of the Kaplan-Meier curve to interval-censored data was used to estimate the cumulative probability of incident knee OA over time (survivor function) according to the mtDNA haplogroups and mtDNA clusters. An extended Cox proportional hazard model using iterative convex minorant algorithm was used for multivariate analysis adjusting for the confounder effects of gender, age, body mass index (BMI), Western Ontario and McMaster Universities Arthritis Index (WOMAC) (total) as well as contralateral OA, all of them at baseline. Due to the difficulties in deriving the asymptotic behavior of statistic tests based on interval-censored data, statistical significance was tested by confidence intervals (CIs) for the hazard ratios (HR) by means of resampling methods. Therefore, CIs were obtained using the bootstrap methodology (1000 replicates) with improved percentile method.

The multivariate analysis of incident knee OA was performed by comparisons between mtDNA haplogroups considering the most common

haplogroup H as the reference group. Therefore, in order to introduce mtDNA haplogroups in the models, a dummy coding was used with the haplogroup H as the reference group. Since there was no interest in all possible pair-wise comparisons, no additional adjusting for multiple comparisons was done.

Meta-analysis

The meta-analysis performed in this work was developed according to the PRISMA guidelines. In order to find relevant studies to include in the meta-analysis we performed a computerized search strategy.

According to above, we identified relevant studies published in English or Chinese by a computerized search via two databases, Pubmed and Web of Science. In addition, non-published studies were also searched in the database of OpenSIGLE (System for Information on Grey literature in Europe http://www.opengrey.edu). The search strategy involved the use of the following key words: osteoarthritis and (mtDNA or mitochondrial DNA) and (haplogroup or haplotype or genotype or genetic predisposition or SNP or polymorphism or variant or genetic susceptibility or genetics or allele) and (incidence or incident OA).

We only selected studies that met the following inclusion criteria: i) evaluating the association between mtDNA haplogroups and the rate of incident OA over time; ii) with sufficient data provided to calculate hazard ratios (HR) with their corresponding 95% confidence interval (95% CI). On the contrary, those studies analyzing the correlation between mtDNA haplogroups and the prevalence or risk of OA as well as those studies analyzing other mtDNA mutations with incidence were excluded.

The data that have been extracted from each study were the following: i) name of the author, ii) year of publication, iii) sample size, iv) country, v) ethnicity, vi) age of subjects, vii) type of study, viii) type of OA, ix) incident knee OA criteria, x) genotyping method, xi) confounder variables, xii) conclusion of the study.

The random-effects model described by DerSimoniarn and Laird was used to calculate a summary statistic and its 95% confidence interval (*). Adjusted hazard ratio (HR) was used as the effect size measure for the association between mtDNA haplogroups and OA incidence. Meta-analysis results were presented on a forest plot graph. To explore heterogeneity, the I^2 index was computed. Meta-analysis was carried out using the R software program (version 3·2·2), using the meta package. A two-tailed P-value<0·05 was considered to be significant.

Functional studies using transmitochondrial cybrids

Cybrid cultures and culture conditions

The obtained cybrids were cultured until nearly confluent in Dulbecco´s Modified Eagle Medium (DMEM) 10% Fetal Bovine Serum (FBS) and penicillin (100 U/ml)/streptomycin (100 μg/ml) (P/S; Gibco). To carry out most of the functional studies, cybrids were plated in MW-6 ($8x10^4$ cells per well) during 48 hours in DMEM with 10% FBS.

In the case of platelet donors, informed consent and the agreement of the ethical committee from Galician Health Administration were obtained.

Cybrids stabilization and haplogroup genotyping

Total DNA from cultured cells was isolated with the QIamp DNA Mini kit (QIagen) following manufacturer recommendations. mtDNA copy number using

SYBR Green in a LightCycler 480 II system (Roche) was assessed to check the stabilization of the obtained cybrids. This measure was determined from the *Crossing Point* (CP) 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 haplogroups, these were assessed following the previously described method.

RNA isolation and gene expression assays

Cells from cybrid cultures (n=2 different individuals with the haplogroup H, two clones from each cybrid, and n=2 different individuals with the haplogroup J, two clones from each cybrid) 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. The results were analyzed with qBase+ software v2.5 (Biogazelle) using the Ribosomal Protein L13a (RPL13A) as housekeeping gene.

Flux assay measurements

Extracellular acidification rate (ECAR) (largely result of glycolysis) and oxygen consumption (OCR) (indicator of mitochondrial respiration) was determined by direct measurement in a SeaHorse XFp Extracellular Flux Analyzer instrument (Seahorse Bioscience, Agilent Technologies). $2x10^4$ cells per well were seeded in a pyruvate-free medium 24 hours prior to the assay in XF cell culture microplates and incubated at 37°C with 5 % CO_2. The next day,

cells were washed with assay medium before pre-incubated without CO_2 for 1 hour and ECAR and OCR was determined following the manufacturer's instructions.

1. *Flux assay parameters*

Glycolysis [(Maximum rate measurement before Oligomycin injection)–(Last rate measurement before Glucose injection)],

Glycolytic Capacity [(Maximum rate measurement after Oligomycin injection)–(Last rate measurement before Glucose injection)].

Basal Respiration [(Last rate measurement before first injection)–(Non-Mitochondrial Respiration Rate)].

ATP Production [(Last rate measurement before Oligomycin injection)–(Minimum rate measurement after Oligomycin injection)].

Non-mitochondrial Respiration (Minimum rate measurement after Rotenone/Antimycin-A injection)

Mitochondrial reactive oxygen species (ROS) production assay

Once the medium was depleted, cells were incubated for 1 hour prior to incubation with Dihydrorhodamine 123 (DHR123) during 30 min at 37ºC in darkness at a final concentration of 10 µM to measure the concentration of mitochondrial peroxide and peroxynitrite. Besides, mitochondrial superoxide anion was evaluated with MitoSox Red (Thermo Fisher Scientific) at a final concentration of 5 µM in Hank´s Balanced Sal Solution (HBSS, Sigma) during 15 min at 37ºC in darkness.

Cells were harvested by trypsin and resuspended 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 (AU).

Oxidative stress response assay. Viability assay

The viability assay was measured in $2x10^3$ cells using CellTiter 96® Aqueous Assay (Promega). Cells were grown in MW-96 for 24 hours in DMEM 10% FBS. Then, the medium was replaced by DMEM without serum and maintained in culture during 24 hours. Next, 300 µM H_2O_2 was added during 30 minutes and finally 20 µl 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 4 hours at 37°C followed by measurement of absorbance at 490nm in a microplate reader (NanoQuant Infinite M200, Tecan)

Apoptosis assay

The cells were cultured in presence of staurosporine at 0.2 µM during 2 hours to induce apoptosis and were subsequently washed and resuspended in 1 X Annexin-binding buffer; then 5 µl of the Annexin V-FITC and 5 µl of Propidium Iodide (PI) (ImmunoStep) were added to each 100µl of cell suspension. Cells were incubated at room temperature for 15 minutes in darkness. After incubation, 400 µl of 1X Annexin-binding buffer was added prior to analyze by flow cytometry a total of $1x10^4$ cells per assay. Data obtained were analyzed with CellQuest software (Becton Dickinson). The analysis of apoptosis was performed by taking into account staining cells simultaneously with Anexine V–FITC and PI, allowing the discrimination of intact cells (Anexine-V-FITC and PI negative), early apoptotic state (Anexine-V-FITC positive and PI

negative) and late apoptosis state (Anexine-V-FITC and IP positives). Results were expressed as percentage of positive cells to each dye.

In addition, the basal expression of the mitochondrial apoptotic-related genes B-cell CLL/lymphoma 2-Like 13 (BCL2L13) and BCL2 binding component 3 (BBC3) in H and J cybrids was also quantified.

Statistical analysis

Results were expressed as the mean of three independent experiments (mean±SD) using two cybrids (two J cybrids and two H cybrids) from two individuals (two different J individuals and two different H individuals) and two clones from each cybrid. Statistically significant differences between the two groups were determined with t-test. P-values below 0.05 were considered significant.