EXTENDED REPORT

A meta-analysis of genome-wide association studies identifies novel variants associated with osteoarthritis of the hip

Evangelos Evangelou,1,2 Hanneke J Kerkhof,3 Unnur Styrkarsdottir,4 Evangelia E Ntzani,1 Steffan D Bos,5,6 Tonu Esko,7,8 Daniel S Evans,9 Sarah Metreustry,2 Kalliope Panoutsopoulou,10 Yolande F M Ramos,5 Guðmar Thorleifsson,4 Konstantinos K Tsilidis,1 arcOGEN Consortium, Nigel Arden,11,12 Nadim Aslam,13 Nicholas Bellamy,14 Fraser Birrell,15,16 Francisco J Blanco,17 Andrew Carr,11 Kay Chapman,11 Aaron G Day-Williams,10 Panos Deloukas,10 Michael Doherty,18 Gunnar Engström,19 Hafdis T Helgadottir,4 Albert Hofman,20 Thorvaldur Ingvarsson,21,22 Helgi Jonsson,23,24 Aime Keis,25,26 J Christiana Keurentjes27 Margreet Kloppenburg,28 Penelope A Lind,29 Andrew McCaskie,15 Nicholas G Martin,30 Lili Milani,7 Grant W Montgomery,31 Rob G H H Nelissen,27 Michael C Nevitt,32 Peter M Nilsson,19 William ER Ollier,33 Neeta Parimi,9 Ashok Rai,34 Stuart H Ralston,35 Mike R Reed,16 Jose A Riancho,36 Fernando Rivadeneira,3,20 Cristina Rodriguez-Fontenla,37 Lorraine Southam,10 Unnur Thorsteinsdottir,4,23 Aspasia Tsezou,38 Gillian A Wallis,39 John PA Ioannidis,1,44 Tim D Spector,2 Joyce B J van Meurs,3 Ana M Valdes2,18

Handling editor Tore K Kvien

Abstract

Objectives Osteoarthritis (OA) is the most common form of arthritis with a clear genetic component. To identify novel loci associated with hip OA we performed a meta-analysis of genome-wide association studies (GWAS) on European subjects.

Methods We performed a two-stage meta-analysis on more than 78,000 participants. In stage 1, we synthesised data from eight GWAS whereas data from 10 centres were used for ‘in silico’ or ‘de novo’ replication. Besides the main analysis, a stratified by sex analysis was performed to detect possible sex-specific signals. Meta-analysis was performed using inverse-variance fixed effects models. A random effects approach was also used.

Results We accumulated 11,277 cases of radiographic and symptomatic hip OA. We prioritised eight single nucleotide polymorphism (SNPs) for follow-up in the discovery stage (4349 OA cases); five from the combined analysis, two male specific and one female specific. One locus, at 20q13, represented by rs6094710 (minor allele frequency (MAF) 4%) near the NCOA3 (nuclear receptor coactivator 3) gene, reached genome-wide significance level with P=7.9×10−9 and OR=1.28 (95% CI 1.18 to 1.39) in the combined analysis of discovery (P=5.6×10−8) and follow-up studies (P=7.3×10−4). We showed that this gene is expressed in articular cartilage and its expression was significantly reduced in OA-affected cartilage. Moreover, two loci remained suggestive associated; rs5009270 at 7q31 (MAF 6%, P=2.2×10−6; OR=1.27 in male specific analysis).

Conclusions Novel genetic loci for hip OA were found in this meta-analysis of GWAS.

Introduction

Osteoarthritis (OA) is the most common form of arthritis affecting 40% of people over the age of 70 years and is one of the most common disabling diseases observed worldwide.1 2 Its aetiology is multifactorial with a clear genetic component. Inheritance studies in twins and other family-based studies have assessed the estimated heritability for OA in the range of 40–65% depending on the joint site.3−5 The established OA loci have small ORs (range 1.10–1.20)6 7 and the genetic architecture of OA is likely to consist of multiple variants of similar magnitude.

During the last few years, extensive efforts have led to the identification of a number of OA susceptibility signals in European populations that have surpassed the genome-wide significance (GWS) level (p<5×10−8). A locus on chr7q22 near the orphan receptor GPR22 derived from genome-wide association studies (GWAS)9 9 and a variant in the GDF5 gene, originating from a candidate gene approach reached GWS for knee OA.10 An analysis using a 1000-genomes-project-based imputations identified a variant on chromosome 13q34 near the
MCF2L gene. A recent GWAS in UK subjects revealed eight more loci that increase risk for OA. From those eight loci, four signals in 9q33 (ASTN2), 6q14 (FILIP1/SENPH), 12p11 (KLHDC5/PTHLH) and 12q23 (CHST11) were found to be associated with total hip replacement (THR) or hip OA in European populations. Finally, GWAS and functional studies revealed that the DOTTIL gene on 19q13 is also associated with hip OA and cartilage thickness.

In this study a large-scale GWAS meta-analysis for hip OA was performed under the auspices of the Translational Research in Europe Applied Technologies in Osteoarthritis (TreatOA) consortium including eight sample sets, in the discovery stage. With a total of 4349 hip OA cases and 46,903 controls in the discovery stage, and a total of 11,277 cases and 67,473 controls, this is the largest study of hip OA to date.

METHODS

Study design and analysis plan

A two-stage design was used for the identification of potential associations. In the discovery stage imputed and directly genotyped autosomal single nucleotide polymorphism (SNPs) were assessed using a quality control (QC) procedure that is described in the online supplementary material (section 1.a). Briefly, we excluded SNPs based on low minor allele frequency (MAF <1%), low imputation quality, low call rate and deviation from Hardy-Weinberg equilibrium. Genomic control was applied to each study before meta-analysis. The effect estimates of each study where synthesised using an additive model. The variants that surpassed the p<1×10⁻⁶ threshold in the meta-analysis were selected for further follow-up. Besides the main analysis including all participants a separate analysis stratified by sex was performed to detect possible sex-specific signals. In silico and de novo replication was sought for the discovery signals in 10 additional studies. All the derived effects from the discovery and the replication stage were finally synthesised using inverse variance fixed-effect models and the between-study heterogeneity was assessed using the I² metric. Moreover, a random-effects (RE) model was applied. A p-value of <5×10⁻⁸ was considered GWS. The associations of the top findings at the discovery stage were also assessed when adjusted for other risk factors such as age, height and body mass index (BMI). We also examined the association of these markers with height and BMI in the large publicly available sample set of the Genetic Investigation of Anthropometric Traits (GIANT) consortium (133,000 individuals in height analysis and 123,000 individuals in BMI analysis. The detailed analysis plan is presented in the online supplementary material (section 1.b).

Study populations and phenotype definition

The studies included in the discovery and replication efforts are described in table 1 and more details are given in the online supplementary material (section 1.c). All studies had standardised definitions of the phenotypes. Specifically, the definition of the hip OA in the studies was either a radiographic Kellgren and Lawrence (K/L) grade of ≥2 or history of a THR surgery because of OA. THR subjects were excluded from the study if they had: other major arthropathy (eg, rheumatoid arthritis, ankylosing spondylitis); Paget’s disease affecting the pelvis or femur; THR due to hip trauma or avascular necrosis of the femoral head; or terminal illness. The control groups consisted of subjects who had no known affected joints. Population-based controls were used by the arcOGEN study.

Table 1 Studies included in the TreatOA GWAS meta-analysis

<table>
<thead>
<tr>
<th>Study</th>
<th>N cases</th>
<th>N controls</th>
<th>N SNPs</th>
<th>ROA/SOA Controls</th>
<th>Genotyping platform</th>
<th>Imputation method</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Discovery stage</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>arcOGEN stage 1</td>
<td>1728</td>
<td>4896</td>
<td>1.058</td>
<td>2.454.242</td>
<td>ROA/SAO Population</td>
<td>Illumina Human610 (cases)+Illumina 1.2M Duo (controls)</td>
</tr>
<tr>
<td>deCODE</td>
<td>1423</td>
<td>3138</td>
<td>1.182</td>
<td>2.399.690</td>
<td>SOA</td>
<td>Infinium HumanHap 300+HumanCNV370</td>
</tr>
<tr>
<td>EGCUT</td>
<td>64</td>
<td>2531</td>
<td>0.994</td>
<td>2.242.156</td>
<td>ROA Population</td>
<td>Illumina HumanCNV370 or HumanOmniExpress</td>
</tr>
<tr>
<td>GAPR</td>
<td>106</td>
<td>1671</td>
<td>1.294</td>
<td>2.406.007</td>
<td>ROA/ROA KL&lt;2</td>
<td>Illumina Infinium HD Human660W-Quad</td>
</tr>
<tr>
<td>RSI</td>
<td>760</td>
<td>3233</td>
<td>1.009</td>
<td>2.450.385</td>
<td>ROA KL&lt;2</td>
<td>Illumina HumanHap550v3</td>
</tr>
<tr>
<td>RSII</td>
<td>159</td>
<td>1472</td>
<td>0.993</td>
<td>2.442.419</td>
<td>ROA KL&lt;2</td>
<td>Illumina HumanHap550-Duo</td>
</tr>
<tr>
<td>RSIII</td>
<td>41</td>
<td>1487</td>
<td>0.962</td>
<td>2.397.764</td>
<td>ROA KL&lt;2</td>
<td>Illumina Human660W-Quad</td>
</tr>
<tr>
<td>TwinsUK</td>
<td>68</td>
<td>228</td>
<td>0.993</td>
<td>2.358151</td>
<td>ROA KL&lt;2</td>
<td>Illumina HumanHap300</td>
</tr>
<tr>
<td><strong>Total discovery</strong></td>
<td>4349</td>
<td>46,903</td>
<td>1.028</td>
<td>2.567.279</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Replication stage</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>arcOGEN stage 2</td>
<td>1763</td>
<td>6157</td>
<td>1.058</td>
<td>–</td>
<td>ROA/SAO –</td>
<td>Illumina Human610 (cases)+Illumina 1.2M Duo (controls)</td>
</tr>
<tr>
<td>SOF*</td>
<td>761</td>
<td>2376</td>
<td>1.008</td>
<td>–</td>
<td>ROA –</td>
<td>Illumina Omni 1 array (1.1 million probes)</td>
</tr>
<tr>
<td>MoOS*</td>
<td>446</td>
<td>2837</td>
<td>1.002</td>
<td>–</td>
<td>ROA –</td>
<td>Illumina Omni 1 array (1.1 million probes)</td>
</tr>
<tr>
<td>arcOGEN plus</td>
<td>223</td>
<td>1828</td>
<td>–</td>
<td>ROA/SAO –</td>
<td>Illumina Human610 (cases)+Illumina 1.2M Duo (controls)</td>
<td>Impute</td>
</tr>
<tr>
<td>EGCUT</td>
<td>977</td>
<td>1131</td>
<td>–</td>
<td>ROA –</td>
<td>Illumina HumanCNV370 or HumanOmniExpress</td>
<td>Impute</td>
</tr>
<tr>
<td>Greek TJR cases</td>
<td>93</td>
<td>361</td>
<td>–</td>
<td>SOA –</td>
<td>Single base extension using SNaPshot Multiplex Kit (Applied Biosystems, Foster City, USA)</td>
<td>–</td>
</tr>
<tr>
<td>Paprika Study</td>
<td>602</td>
<td>2321</td>
<td>–</td>
<td>SOA –</td>
<td>Sequenom (MassARRAY iPLEX Gold)</td>
<td>–</td>
</tr>
<tr>
<td>Spanish TJR cases</td>
<td>697</td>
<td>783</td>
<td>–</td>
<td>SOA –</td>
<td>Single base extension using SNaPshot Multiplex Kit (Applied Biosystems, Foster City, USA)</td>
<td>–</td>
</tr>
<tr>
<td>Icelandic cases</td>
<td>857</td>
<td>1857</td>
<td>–</td>
<td>SOA –</td>
<td>Centaurus (Nanogen)12</td>
<td>–</td>
</tr>
<tr>
<td>Swedish MDC study</td>
<td>509</td>
<td>919</td>
<td>–</td>
<td>SOA –</td>
<td>Centaurus (Nanogen)17</td>
<td>–</td>
</tr>
<tr>
<td><strong>Total replication</strong></td>
<td>6928</td>
<td>20,570</td>
<td>–</td>
<td>SOA –</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td>11,277</td>
<td>67,473</td>
<td>–</td>
<td>SOA –</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Genotyping and imputation

Genotyping of GWAS genotyping was performed by each study following standard protocols and imputation was then carried out at the individual study level on the ~2.5 million SNPs from HapMap Phase 2 release 22 using genome build 36 (Utah residents with ancestry from northern and western Europe (CEPH))\(^7\) on MACH or IMPUTE\(^8\) software. Imputation quality scores for each SNP were obtained from IMPUTE and MACH statistics, as appropriate. An overview of all studies and the genotyping platforms and imputation method used is given in Table 1. Ten studies of European ancestral origin provided data for independent replication. Four datasets (arcOGEN stage 2, arcOGEN plus, Osteoporotic Fractures in Men study and the Study of Osteoporotic Fractures) provided 'in silico' replication whereas 'de novo' replication was performed in six other study groups (Icelandic, Swedish, Estonian Genome Center, University of Tartu (EGCUT), Puprikra study, Greek, Spanish). QC criteria for deviation from Hardy-Weinberg equilibrium, MAF inconsistencies with the discovery data and outliers were applied in the replication data before including all the available data in the final analysis.

Gene expression

Expression was determined by Illumina HT-12 V3 microarrays using standard methods using 47 000 probes corresponding to over 25 000 well-characterised genes. DNA was available from blood and cartilage. Using the Beadstudio software the intensity values were normalised using the 'rsn' option in the Lumi R-package. The corresponding signals increase exponentially with relative levels and units are light intensity (Illumina probe-level value of measured genes in cartilage) divided values). The obtained raw probe-level data (overall mean with relative levels and units are light intensity (Illumina probe-level value of measured genes in cartilage) divided values). The obtained raw probe-level data (overall mean with relative levels and units are light intensity (Illumina probe-level value of measured genes in cartilage) divided values). The obtained raw probe-level data (overall mean with relative levels and units are light intensity (Illumina probe-level value of measured genes in cartilage) divided values). The obtained raw probe-level data (overall mean with relative levels and units are light intensity (Illumina probe-level value of measured genes in cartilage) divided values). The obtained raw probe-level data (overall mean with relative levels and units are light intensity (Illumina probe-level value of measured genes in cartilage) divided values). The obtained raw probe-level data (overall mean with relative levels and units are light intensity (Illumina probe-level value of measured genes in cartilage) divided values).

Heritability of hip OA explained by genetic variants

We calculated the sibling recurrence risk and the expected genetic variance explained for hip OA hits that were identified previously and in this study as described in online supplementary material (section 1.e)

RESULTS

The final analysis included a total of 11 277 radiographic and symptomatic hip OA cases and 67 473 controls of European ancestry, with 4349 cases and 46 903 controls included in the discovery stage and 6928 cases and 20 570 controls in the follow-up effort (Table 1). In the sex-specific analyses 2045/20 823 male cases and controls and 2689/25 384 female cases and controls were analysed in the discovery. After QC, 2 567 279 SNPs were analysed. Low genomic inflation factor was observed for the first stage with \(\lambda = 1.028\). The results of the discovery stage were uncorrected for the overall inflation factor. Quantile-Quantile plots showed an excess of signals compared with what was expected by chance, indicating the presence of true association signals that could confer susceptibility to hip OA (see online supplementary figures S1–S3). For male-specific analysis (see online supplementary figure S3) there is an early deviation from the neutrality line, therefore positive signals should be treated with caution.

Following analysis of the discovery stage it was found that eight independent loci reached the prespecified threshold of \(p<1.0 \times 10^{-8}\) required for further replication; five from the combined analysis of sexes, two male-specific and one female-specific loci. The number of independent SNPs is larger compared with three estimated independent SNPs expected under the null for the main and the sex-specific analyses (binomial test \(p=0.012\)). Two of these signals (rs6094710 at 20q13 and rs640070 at 17q23) were imputed, had MAF <5% and moderate or large effect sizes were observed (OR >1.2). Therefore, we examined for possible imputation errors by comparing with de novo genotyping in a random sample of the RS-I and the Twins UK studies to exclude any chance of false positive findings. The obtained MAF estimates between the imputation and the genotyping efforts were not consistent for rs640070, which was therefore excluded from further consideration, minimising the chance of any false positive signals (see online supplementary material; section 1.d). Moreover, we included in the replication stage rs17610181 at 17q23, a SNP that was just below the desired threshold, but was previously shown to be associated with height.\(^15\) Therefore, eight independent SNPs were included in the replication phase. The combined effect sizes and p values for the eight signals are presented in Table 2

### Table 2

<table>
<thead>
<tr>
<th>SNP</th>
<th>Gene</th>
<th>Chr</th>
<th>A1 EAF (%)</th>
<th>OR (95% CI)</th>
<th>p Value</th>
<th>I² (%)</th>
<th>OR (95% CI)</th>
<th>p Value</th>
<th>OR (95% CI)</th>
<th>p Value</th>
<th>I² (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Combined for both genders:</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>rs6094710</td>
<td>NCOA3</td>
<td>20</td>
<td>1.40 (1.24 to 1.58)</td>
<td>5.6 × 10⁻⁸</td>
<td>0</td>
<td></td>
<td>1.20 (1.08 to 1.34)</td>
<td>7.3 × 10⁻⁶</td>
<td>1.28 (1.18 to 1.39)</td>
<td>7.9 × 10⁻⁹</td>
<td>64</td>
</tr>
<tr>
<td>rs1577792</td>
<td>HMGN3</td>
<td>6</td>
<td>1.14 (1.09 to 1.20)</td>
<td>7.3 × 10⁻⁸</td>
<td>0</td>
<td></td>
<td>1.03 (0.99 to 1.07)</td>
<td>0.17</td>
<td>1.07 (1.04 to 1.11)</td>
<td>7.8 × 10⁻⁵</td>
<td>50</td>
</tr>
<tr>
<td>rs5009270</td>
<td>IFRD1</td>
<td>1q</td>
<td>1.15 (1.09 to 1.21)</td>
<td>5.7 × 10⁻⁷</td>
<td>21</td>
<td></td>
<td>1.05 (1.00 to 1.11)</td>
<td>0.061</td>
<td>1.10 (1.06 to 1.14)</td>
<td>9.0 × 10⁻⁴</td>
<td>46</td>
</tr>
<tr>
<td>rs10773048</td>
<td>DNAXH1O</td>
<td>4</td>
<td>1.13 (1.08 to 1.19)</td>
<td>9.7 × 10⁻⁷</td>
<td>10</td>
<td></td>
<td>1.01 (0.96 to 1.06)</td>
<td>0.72</td>
<td>1.07 (1.03 to 1.11)</td>
<td>2.0 × 10⁻⁶</td>
<td>46</td>
</tr>
<tr>
<td>rs17610181</td>
<td>NACA2</td>
<td>17</td>
<td>1.20 (1.11 to 1.29)</td>
<td>1.9 × 10⁻⁶</td>
<td>32</td>
<td></td>
<td>1.05 (0.99 to 1.12)</td>
<td>0.12</td>
<td>1.12 (1.06 to 1.18)</td>
<td>1.3 × 10⁻⁵</td>
<td>42</td>
</tr>
<tr>
<td>Female-specific analysis:</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>rs10878630</td>
<td>DYRK2</td>
<td>41</td>
<td>1.19 (1.11 to 1.26)</td>
<td>1.2 × 10⁻⁷</td>
<td>25</td>
<td></td>
<td>0.98 (0.91 to 1.05)</td>
<td>0.52</td>
<td>1.09 (1.04 to 1.14)</td>
<td>3.9 × 10⁻⁴</td>
<td>65</td>
</tr>
<tr>
<td>Male-specific analysis:</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>rs12551314</td>
<td>PHF2</td>
<td>12</td>
<td>1.30 (1.18 to 1.44)</td>
<td>6.3 × 10⁻⁷</td>
<td>28</td>
<td></td>
<td>0.98 (0.86 to 1.11)</td>
<td>0.72</td>
<td>1.16 (1.07 to 1.25)</td>
<td>2.9 × 10⁻⁴</td>
<td>66</td>
</tr>
<tr>
<td>rs357837</td>
<td>CAMK2B</td>
<td>6</td>
<td>1.46 (1.26 to 1.69)</td>
<td>8.3 × 10⁻⁷</td>
<td>46</td>
<td></td>
<td>1.15 (1.00 to 1.32)</td>
<td>0.044</td>
<td>1.27 (1.15 to 1.41)</td>
<td>2.2 × 10⁻⁶</td>
<td>78</td>
</tr>
</tbody>
</table>

The discovery sample set comprised 4349 hip OA cases and 46 903 controls, and the replication samples were 6928 cases and 20 570 controls in the analysis of both genders. The female-specific analysis comprised 2689 hip OA cases and 25 384 controls in the discovery set and 2398 cases and 8778 controls in the replication stage. For the male-specific analysis there were 2045 hip OA cases and 20 823 controls in the discovery set and 13867/878 and 1451/4956 cases/controls in the replication stage for rs12551314 and rs357837 respectively.

A1, coded allele; Chr, chromosome; EAF, effect allele frequency; Gene, nearest gene; HE RE P, Han & Eskin Random Effects P; \(\lambda\), measure of heterogeneity; OA, osteoarthritis.
and in online supplementary figures S4–S11. None of these variants associate with height or BMI in the large publicly available GIANT databases (see online supplementary table S2) nor did including these covariates, or age, in the analysis change the association of OA with these genetic variants (see online supplementary table S3).

Two SNPs, rs6094710 at 20q13 and rs1577792 at 6q14, were borderline GWS in the discovery stage with \( p = 5.8 \times 10^{-8} \) and \( 7.3 \times 10^{-8} \), respectively, with no observed heterogeneity \( (I^2 = 0) \). The SNP rs6094710 replicated with \( p = 7.3 \times 10^{-8} \) in the follow-up samples and reached GWS level with \( p = 9.3 \times 10^{-8} \) and OR=1.28 (95% CI 1.17 to 1.39) when we combined the discovery and the replication data, although large heterogeneity was observed \( I^2 = 64\% \) (figure 1). The SNP remained GWS after the second genomic control. rs6094710 is annotated near NCOA3 (nuclear receptor coactivator 3) gene. The OR of rs1577792 at 6q14 was close to unity in the replication effort with one study being significant in the opposite direction.

The rs3757837 SNP at 7p13 from the male-specific analysis replicated nominally \( (p = 0.044, \text{OR}=1.15) \) in the follow-up samples, but did not reach GWS in the overall analysis with replicated nominally \( (p = 0.044, \text{OR}=1.15) \) in the follow-up the Han and Eskin (RE) model, an approach that allows more introduced by the replication data. We, therefore, also applied with one study being significant in the opposite direction.

Using this model rs3757837 showed a stronger association with \( p = 2.2 \times 10^{-6} \) and \( 7.4; \text{range 6.3}^{-14.7} \). When we tested for differential expression of NCOA3 among the pairs of preserved and OA-affected cartilage, we observed a significantly lower expression of NCOA3 in the OA-affected cartilage \( (p = 0.0064) \).

Heritability of hip OA explained by genetic variants

Table 3 summarises genome-wide significant and suggestive signals of hip OA including this study. Based on these findings and if we consider a sibling recurrence ratio \( \lambda_s = 5\) then the discovered signals of OA contribute 3% of the heritability in OA.

The rs5009270 SNP at 7q31 remained suggestively associated in the final analysis \( (p = 5 \times 10^{-6}) \) with combined two-stage \( p = 9.9 \times 10^{-7} \) and summary OR 1.10. rs5009270 is located near the IFRD1 (interferon-related) gene. The p with the RE approach was \( 3.1 \times 10^{-6} \).

Gene expression data

To further investigate our findings for the top GWS hit at 20q13 we explored mRNA expression profile of the NCOA3 gene in articular cartilage, a highly relevant tissue for OA. Expression of the gene was examined by exploring a micro array mRNA expression dataset generated on Illumina V3 Human-12 chips in cartilage samples of 33 patients (13 men and 20 women of European descent aged 54 years to 80 years) that underwent joint replacement due to end-stage OA disease. Expression levels of NCOA3 in cartilage displaying symptoms of OA were compared with expression levels in cartilage that appeared macroscopically normal but isolated from the same joint (preserved cartilage). A moderate level of expression, as determined by the mean normalised probe level value, was observed for NCOA3 (mean level 8.49), which was above the observed average expression of genes in the articular cartilage (mean normalised probe level value of measured genes in cartilage was 7.4; range 6.6–14.9). Expression was also high in blood (6.96; range 6.3–14.7). When we tested for differential expression of NCOA3 among the pairs of preserved and OA-affected cartilage, we observed a significantly higher expression of NCOA3 in cartilage displaying symptoms of OA (\( p = 0.0064 \)).

Figure 1 (A) Forest plot for meta-analysis of rs6094710. The blue diamond in the forest plot denotes the summary effect size and its edges the respective 95% CIs. Studies shaded in blue were included in the replication stage. *Discovery and replication estimate combined; and (B) regional plot for rs6094710 comprising directly genotyped and imputed SNPs. Case-control association results \( (−\log_{10} P) \) in the discovery set are plotted against genomic position (National Center for Biotechnology Information build 36) for the stratum where the most significant meta-analysis p was observed. The colour reflects the correlation coefficients \( (r^2) \) of each genotyped SNP with the index SNP estimated using the CEU HapMap II panel.


Ann Rheum Dis: first published as 10.1136/annrheumdis-2012-203114 on 29 August 2013. Downloaded from http://ard.bmj.com/ on April 7, 2022 by guest. Protected by copyright.
clinical and epidemiological research

Table 3  Summary of the genome-wide significant and suggestive SNPs for hip OA and their estimated heritability

<table>
<thead>
<tr>
<th>Reference</th>
<th>SNP</th>
<th>Gene</th>
<th>OR</th>
<th>EAF</th>
<th>$\lambda_s$</th>
<th>Expected GV explained (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Genome wide significant findings (p&lt;5x10$^{-8}$)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>This study</td>
<td>rs6094710</td>
<td>NCOA3</td>
<td>1.28</td>
<td>0.04</td>
<td>10029</td>
<td>0.46</td>
</tr>
<tr>
<td>arcOGEN6</td>
<td>rs6976</td>
<td>GLT8D1</td>
<td>1.12</td>
<td>0.37</td>
<td>10031</td>
<td>0.60</td>
</tr>
<tr>
<td>arcOGEN6</td>
<td>rs11177</td>
<td>GNl3</td>
<td>1.12</td>
<td>0.38</td>
<td>10031</td>
<td>0.60</td>
</tr>
<tr>
<td>arcOGEN6</td>
<td>rs4836732</td>
<td>ASTN2</td>
<td>1.20</td>
<td>0.47</td>
<td>10083</td>
<td>1.66</td>
</tr>
<tr>
<td>arcOGEN6</td>
<td>rs9350591</td>
<td>FLP1;SPEP6</td>
<td>1.18</td>
<td>0.11</td>
<td>10030</td>
<td>0.54</td>
</tr>
<tr>
<td>arcOGEN6</td>
<td>rs10492367</td>
<td>KHDSC;PTLH</td>
<td>1.14</td>
<td>0.19</td>
<td>10029</td>
<td>0.53</td>
</tr>
<tr>
<td>3arcOGEN6</td>
<td>rs385487</td>
<td>CHST11</td>
<td>1.13</td>
<td>0.34</td>
<td>10034</td>
<td>0.67</td>
</tr>
<tr>
<td>TreatOA12</td>
<td>rs12982744</td>
<td>DOT1L</td>
<td>1.17</td>
<td>0.38</td>
<td>10060</td>
<td>1.16</td>
</tr>
<tr>
<td>Suggestive findings (5x10$^{-6}$&lt;p&lt;5x10$^{-8}$)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>This study</td>
<td>rs5009270</td>
<td>IFRD1</td>
<td>1.10</td>
<td>0.30</td>
<td>10019</td>
<td>0.38</td>
</tr>
<tr>
<td>This study</td>
<td>rs3757837</td>
<td>CAMK2B</td>
<td>1.27</td>
<td>0.06</td>
<td>10040</td>
<td>0.64</td>
</tr>
<tr>
<td>arcOGEN6</td>
<td>rs12107036</td>
<td>TP63</td>
<td>1.21</td>
<td>0.52</td>
<td>10090</td>
<td>1.81</td>
</tr>
<tr>
<td>arcOGEN6</td>
<td>rs8044769</td>
<td>FTO</td>
<td>1.11</td>
<td>0.50</td>
<td>10027</td>
<td>0.54</td>
</tr>
<tr>
<td>arcOGEN6</td>
<td>rs10948172</td>
<td>SUPT3H;CDC5L</td>
<td>1.14</td>
<td>0.29</td>
<td>10037</td>
<td>0.71</td>
</tr>
</tbody>
</table>

EAF, effect allele frequency; GV, genetic variance; OA, osteoarthritis.

when all hits are considered and 2.1% if only GWS are evaluated.

Discussion

In this report we attempted to further clarify the genetic architecture of the genetic background of hip OA by using the largest sample-size for hip OA to-date of more than 78 000 genotyped individuals under a GWAS framework. During the discovery phase, we identified eight signals that qualified for further independent replication. Of these, the signal at chromosome 20q13 near NCOA3 gene was found to be GWS and two other loci were suggestively significant at a $p<5\times10^{-8}$ level in the joint analysis of the discovery and replication stages. Adjusted analyses of the prioritised signals revealed that these markers were not associated with body size. These genetic risk factors contribute to our knowledge base in the field of the susceptibility for hip OA by conferring a medium OA risk.

The top signal identified in this meta-analysis was rs6094710, a variant that is annotated on chromosome 20q13 near the NCOA3 gene, increasing the risk for hip OA for the carriers of the A allele by almost 30%. Furthermore, we showed that the identified NCOA3 gene was expressed in articular cartilage and its expression was significantly reduced in OA affected cartilage, further supporting a role of the NCOA3 gene signal in OA disease process. Interestingly, rs6094710 is in complete linkage disequilibrium with rs6094752 ($r^2=1$) which is a missense SNP leading to a functional change of the protein by PolyPhen-2, dependent on the variant protein. The functional consequences of SNP rs6094752 are unknown, and further research is therefore needed to unravel the biological mechanism of this amino acid change in relation to OA.

The NCOA3 gene is a nuclear receptor coactivator that directly binds nuclear receptors and stimulates the transcriptional activities in a hormone-dependent fashion. In this signalling process, NCOA3 recruits histone acetyltransferases and methyltransferases for chromatin remodelling and facilitating downstream gene transcription. NCOA3 is involved in the coactivation of different nuclear receptors, such as for steroids, retinoids, thyroid hormone, vitamin D3 and prostanoids. Many of these hormones have been implicated in skeletal metabolism and OA, which makes NCOA3 a compelling causal candidate gene. Previously, NCOA3 knockout mice were generated through homologous recombination in embryonic stem cells.22 These mice showed growth retardation and reduced adult body size, but the molecular mechanism responsible for this growth retardation remains largely unknown. In addition, female mice exhibited abnormal development and function of their reproductive system and oestrogen levels were significantly lower in the knockout mice compared with the wild type, possibly indicating involvement of NCOA3 in steroid regulation.

NCOA3 could be also implicated through regulation of the target tissue responses to thyroid hormone (T3).23 Since intracellular T3 is tightly regulated by deiodinase, iodothyronine, type-2 and deiodinase, iodothyronine, type-3 encoded by the DIO2 and DIO3 genes, respectively, that were previously recognised as OA susceptibility genes,24 the current NCOA3 findings complement the previous outlined hypothesis that local T3 signalling may affect OA susceptibility.25

Another possible mechanism by which NCOA3 might be involved in cartilage homeostasis is through transcriptional regulation in mechanotransduction. NCOA3 is upregulated by the signal transducer and activator of transcription 6 (STAT6) in naïve splenic B cells from BALB/c and serves as a positive regulator of transcriptional activation by STAT6.27 STAT6 is known to be an important signal transducer of interleukin (IL)-4 receptor α chain and mediates IL-4 and IL-13-induced responses.28 Chondrocytes from normal and OA cartilage signal through a type II IL-4R in human articular chondrocyte mechanotransduction. This signalling is via a STAT6-independent pathway. Differences in IL-4 signalling are likely due to crosstalk between integrin and cytokine signalling pathways.29 Therefore, NCOA3 may be related to cartilage function and molecular signalling and transcriptional regulation in mechanotransduction.

The male-specific locus on 7p13, represented by rs3757837, showed considerable heterogeneity between studies. The signal was strongly supported by the RE model. Unlike the conservative traditional RE methods this new method has been shown to achieve higher statistical power when heterogeneity exists, allowing for new discoveries in the field of genetic epidemiology.14 rs3757837 is located in CAMK2B gene, which belongs to the calcium/calmodulin-regulated kinase (CaMKII) subfamily.
There is evidence that CaMKII-signalling may be important in onset and progression of OA.\(^{30}\)\(^{31}\) This pathway has been described as central to the molecular events that regulate chondrocyte responses to mechanical stimulation and in particular, to the upstream effect of IL-4. This would be linked to the STAT6 pathway, which in turn is regulated by NCOA3. Thus these two genetic signals, NCOA3 and CAMK2B, both point to pathways related to cartilage mechanotransduction, suggesting that genetic defects in this pathway may be central to the degeneration of cartilage that takes place in hip OA. OA is a disease affecting articular cartilage and the underlying bone, resulting from many biological and mechanical interacting factors which change the extracellular matrix and cells and lead to increasing levels of cartilage degeneration. Joint tissues are exquisitely sensitive to their mechanical environment, and mechanical loading may be the most important external factor regulating the development and long-term maintenance of joint tissues.\(^{32}\) Finally, rs5009270 resides near IFDDR1 which codes an interferon-related developmental regulator and has been implicated in skeletal muscle regeneration.\(^{33}\) Reduction in muscle strength is strongly associated with functional decline, and individuals with lower quadriceps strength adjusted for body weight are more likely to develop OA.\(^{34}\)

Our study has certain limitations. In our discovery stage, two out of the eight independent loci had no heterogeneity and in five signals the heterogeneity was moderate or low. However, when we included the replication data, large heterogeneity was introduced for all SNPs. Conflicting results in the replication data, besides chance, could be explained by inconsistent definitions of the OA phenotypes but also from the different population structure that can introduce heterogeneity. The TreatOA consortium has addressed the need of standardisation of OA phenotypes\(^{35}\) and this effort may have diminished the observed heterogeneity in the discovery stage. Efforts including more data using common and stringent QC criteria and standardisation methods should substantially improve the power of GWAS to identify novel findings in the near future. The main limitation of the expression study is that there are few individuals included, in particular only two carriers of the rs6094710 variant.

In conclusion, novel loci involved in hip OA were discovered through a large-scale meta-analysis of GWAS. The exact underlying mechanism leading to a higher risk of OA remains to be elucidated by functional experiments. It is evident from this work and other recent studies that deciphering the architecture of the genetics of OA requires major large-scale efforts, and in this regard calls for international worldwide collaborations are not fruitless. In the near future emphasis should be given to the enhancement of the total sample size, the adoption of stringent and standardised definitions of the phenotypes and the application of imputation-based meta-analysis using the panel of the 1000 Genomes Project.\(^{36}\) In addition, linking results from genetic association studies to for example genome-wide RNA expression data might further improve our understanding. Eventually, large-scale studies with whole-genome sequencing will be needed to target the heritability caused by less common variants.

**Author affiliations**

1Department of Hygiene and Epidemiology, University of Ioannina Medical School, Ioannina, Greece
2Department of Twin Research & Genetic Epidemiology, King’s College London, London, UK
3Department of Internal Medicine, Erasmus Medical Center, Rotterdam, The Netherlands
4Department of Population Genetics, deCODE Genetics, Reykjavik, Iceland
5Department of Molecular Epidemiology, Leiden University Medical Center, Leiden, The Netherlands
6Netherlands Consortium for Healthy Ageing, The Netherlands
7Estonian Genome Center, University of Tartu, Tartu, Estonia
8Institute of Molecular and Cell Biology, University of Tartu, Tartu, Estonia
9California Pacific Medical Center Research Institute, San Francisco, USA
10Department of Human Genetics, Wellcome Trust Sanger Institute, Hinxton, UK
11NIHR Biomedical Research Unit and ARUK Centre of excellence for Sport, Exercise and Osteoarthritis, University of Oxford, Oxford, UK
12MRC Epidemiology Resource Centre, University of Southampton, Southampton, UK
13Worcestershire Royal Hospital, Worcestershire Acute Hospitals NHS Trust, Worcester, UK
14Centre of National Research on Disability and Rehabilitation Medicine, The University of Queensland, Brisbane, Australia
15Muscloskeletal Research Group, Institute of Cellular Medicine, Newcastle University, Newcastle upon Tyne, UK
16Wansbeck General Hospital, Northumbria Healthcare NHS Foundation Trust, Ashington, UK
17Laboratory Division, Instituto de Investigación Biomédica-Hospital Universitario A Coruña, A Coruña, Spain
18Department of Academic Rheumatology, University of Nottingham, Nottingham, UK
19Department of Clinical Sciences Malmo, Lund University, Malmo, Sweden
20Department of Epidemiology, Erasmus Medical Center, Rotterdam, The Netherlands
21Department of Orthopedic Surgery, Akureyri Hospital, Akureyri, Iceland
22School of Health Sciences, University of Akureyri, Akureyri, Iceland
23Department of Medicine, The National University Hospital of Iceland, Reykjavik, Iceland
24Faculty of Medicine, University of Iceland, Reykjavik, Iceland
25Department of Public Health, University of Tartu, Tartu, Estonia
26Orthopedic Surgeons, Elva Hospital, Elva, Estonia
27Department of Orthopedics, Leiden University Medical Center, Leiden, The Netherlands
28Department of Rheumatology, Leiden University Medical Center, Leiden, The Netherlands
29Department of Quantitative Genetics, QIMR Berghofer Medical Research Institute, Brisbane, Australia
30Department of Genetic Epidemiology, QIMR Berghofer Medical Research Institute, Brisbane, Australia
31Department of Molecular Epidemiology, QIMR Berghofer Medical Research Institute, Brisbane, Australia
32Department of Epidemiology and Biostatistics, University of California, San Francisco, San Francisco, California, USA
33Centre for Integrated Genomic Medical Research, University of Manchester, Manchester, UK
34Worcestershire Acute Hospitals NHS Trust, Worcester, UK
35Institute of Genetics and Molecular Medicine, University of Edinburgh, Edinburgh, UK
36Department of Internal Medicine, Hospital U.M. Valdecilla-RIFMA, University of Cantabria, Santander, Spain
37Laboratorio Investigacion 10 and Rheumatology Unit, Instituto de Investigacion Sanitaria-Hospital Clinico Universitario de Santiago, Santiago de Compostela, Spain
38Department of Biology, University of Thessaly, Medical School, Larissa, Greece
39Wellcome Trust Centre for Cell-Matrix Research, University of Manchester, Manchester, UK
40Department of Human Metabolism, University of Sheffield, Sheffield, UK
41Department of Medicine, University of California at Davis, Sacramento, USA
42Research Unit for Musculoskeletal Function and Physiotherapy, and Department of Orthopedics and Traumatology, University of Southern Denmark, Odense, Denmark
43Department of Clinical Sciences Lund, Lund University, Lund, Sweden
44Stanford Prevention Research Center, Stanford University School of Medicine, Stanford, USA

**Correction notice** This article has been corrected since it was published Online First. The affiliations for P Eline Slagboom have been corrected and affiliations 1 and 2 have been amended.

**Acknowledgements** The authors are very grateful to all study participants, the staff from all studies and the participating physicians and pharmacists. Acknowledgements for each study are given in online supplementary material (section 4).

**Contributors** EE, HJK, US, IM, JPAI, TDS, JBvM and AMV conceived and designed the study. EE, HJK, EEN and KKT analysed the data. EE, HJK, US, EEN, EZ, IM, JPAI, TDS, JBvM and AMV wrote the first draft of the manuscript. All authors interpreted the results, critically commented and approved the final version of the manuscript.

**Funding** This project was supported through Coordination Theme 1 (Health) of the European Community’s FP7 (grant 200800) TreatOA. The work was also supported by EU FP7 Small-scale focused research collaborative projects EurHEALTHAging (grant 277849).

**Competing interests** US, II, GT, HTH, UT and KS are employed by deCODE Genetics/Amgen. HJK is currently employed by Pfizer.
Provenance and peer review Not commissioned; externally peer reviewed.

Open Access This is an Open Access article distributed in accordance with the Creative Commons Attribution Non Commercial (CC BY-NC 3.0) license, which permits others to distribute, remix, adapt, build upon this work non-commercially, and license their derivative works on different terms, provided the original work is properly cited and the use is non-commercial. See: http://creativecommons.org/licenses/by-nc/3.0/

REFERENCES