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 Metrustry,[2] Kalliope Panoutsopoulou,[10] Yolande F M Ramos,[5] Gudmar 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 Christiaan Keurentjes,[27] 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] J Mark Wilkinson,[40] Antonio Gonzalez,[37] Nancy E Lane,[41] L Stefan Lohmander,[42,43] John Loughlin,[15] Andres Metspalu,[7,8] Andre G Uitterlinden,[3,20] Ingileif Jonsdottir,[4,24] Kari Stefansson,[4,24] P Eline Slagboom,[5,6] Eleftheria Zeggini,[10] Ingrid Meulenbelt,[5,6] John PA Ioannidis,[1,44] Tim D Spector,[2] Joyce B J van Meurs,[3] Ana M Valdes[2,18]

Handling editor Tore K Kvien

▶ Additional material is published online only. To view please visit the journal online (http://dx.doi.org/10.1136/annrheumdis-2012-203114).

For numbered affiliations see end of article.

Correspondence to
Dr Evangelos Evangelou, Department of Hygiene and Epidemiology, University of Ioannina Medical School, Ioannina 45110, Greece; vangelis@cc.uoi.gr
and Dr Ana Valdes, Department of Academic Rheumatology, University of Nottingham, Nottingham, NG5 1PB, UK; Ana.Valdes@nottingham.ac.uk

EE, HJK, US, EEN, EZ, IM, JPAI, TDS, JBvM and AMV contributed equally.

Received 12 December 2012
Accepted 26 July 2013
Published Online First 29 August 2013

To cite: Evangelou E, Kerkhof HJ, Styrkarsdottir U, *et al*. *Ann Rheum Dis* 2014;**73**:2130–2136.

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 $p=7.9\times10^{-9}$ and OR=1.28 (95% CI 1.18 to 1.39) in the combined analysis of discovery ($p=5.6\times10^{-8}$) and follow-up studies ($p=7.3\times10^{-4}$). 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 30%, $p=9.9\times10^{-7}$, OR=1.10) and rs3757837 at 7p13 (MAF 6%, $p=2.2\times10^{-6}$, 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 ($p<5\times10^{-8}$). A locus on chr7q22 near the orphan receptor *GPR22* derived from genome-wide association studies (GWAS)[8 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.[7] 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/SENP6*), 12p11 (*KLHDC5/PTHLH*) and 12q23 (*CHST11*) were found to be associated with total hip replacement (THR) or hip OA in European populations.[6] Finally, GWAS and functional studies revealed that the *DOT1L* gene on 19q13 is also associated with hip OA and cartilage thickness.[11 12]

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 the 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 $p<1\times10^{-6}$ 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 I^2metric.[13] Moreover, a random-effects (RE) model was applied.[14] A p-value of $<5\times10^{-8}$ 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.[15 16] 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

Study	N cases	N controls	λ	N SNPs	ROA/SOA	Controls	Genotyping platform	Imputation method
Discovery stage								
arcOGEN stage 1	1728	4896	1.058	2.454.242	ROA/SOA	Population	Illumina Human610 (cases)+Illumina 1.2M Duo (controls)	Impute
deCODE	1423	31 385	1.182	2.399.690	SOA	Osteoarthritis-free	Infinium HumanHap 300+humanCNV370	Impute
EGCUT	64	2531	0.994	2.242.156	ROA	Population	Illumina HumanCNV370 or HumanOmniExpress	Impute
GARP	106	1671	1.294	2.406.007	SOA/ROA	KL<2	Illumina Infinium HD Human660W-Quad	Impute
RSI	760	3233	1.009	2.450.385	ROA	KL<2	Illumina HumanHap550v3	MACH
RSII	159	1472	0.993	2.442.419	ROA	KL<2	Illumina HumanHap550-Duo	MACH
RSIII	41	1487	0.962	2.397.764	ROA	KL<2	Illumina Human660W-Quad	MACH
TwinsUK	68	228	0.993	2.358151	ROA	KL<2	Infinium HumanHap300	Impute
Total discovery	4349	46 903	1.028	2.567.279			–	–
Replication stage								
arcOGEN stage 2	1763	6157	1.058	–	ROA/SOA	–	Illumina Human610 (cases)+Illumina 1.2M Duo (controls)	Impute
SOF*	761	2376	1.008	–	ROA	–	Illumina Omni 1 array (1.1 million probes)	MACH
MrOS*	446	2837	1.002	–	ROA	–	Illumina Omni 1 array (1.1 million probes)	MACH
arcOGEN plus	223	1828	–	–	ROA/SOA	–	Illumina Human610 (cases)+Illumina 1.2M Duo (controls)	Impute
EGCUT	977	1131	–	–	ROA	–	Illumina HumanCNV370 or HumanOmniExpress	Impute
Greek TJR cases	93	361	–	–	SOA	–	Single base extension using SNaPshot Multiplex Kit (Applied Biosystems, Foster City, USA))	–
Paprika Study	602	2321	–	–	SOA	–	Sequenom (MassARRAY iPLEX Gold)	–
Spanish TJR cases	697	783	–	–	SOA	–	Single base extension using SNaPshot Multiplex Kit (Applied Biosystems, Foster City, USA))	–
Icelandic cases	857	1857	–	–	SOA	–	Centaurus (Nanogen)[37]	
Swedish MDC study	509	919	–	–	SOA	–	Centaurus (Nanogen)[37]	–
Total replication	6928	20 570					–	–
Total	11 277	67 473					–	–

EGCUT, Estonian Genome Center, University of Tartu; GWAS, genome-wide association studies; HD, high density; KL, Kellgren-Lawrence score; MDC, Malmo Diet Cancer study; MrOS, Osteroporotic Fractures in Men study; ROA, radiographic osteoarthritis; SOA, symptomatic osteoarthritis; SOF, Study of Osteoporotic Fractures; TJR, total joint replacement.

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))[17] on MACH or IMPUTE[18] 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, Osteroporotic 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), Paprika 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 provided values). The obtained raw probe-level data (overall mean normalised probe level value of measured genes in cartilage) were exported for analyses using Limma.[19] As implemented in Limma, a paired t test was used on all samples. There were two probes, approximately 2.2 kb apart, on the array used for *NCOA3*.

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 $p<1.0\times10^{-6}$ 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 11q25) 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 Association results for hip OA meta-analysis of the discovery and the replication studies

					Discovery stage			Replication stage		Combined			
SNP	Chr	Gene	A1	EAF (%)	OR (95% CI)	p Value	I^2 (%)	OR (95% CI)	p Value	OR (95% CI)	p Value	I^2 (%)	H&E RE p value
Combined analysis of both genders:													
rs6094710	20q13	NCOA3	A	4	1.40 (1.24 to 1.58)	5.6×10^{-8}	0	1.20 (1.08 to 1.34)	7.3×10^{-4}	1.28 (1.18 to 1.39)	7.9×10^{-9}	64	2.0×10^{-10}
rs1577792	6q14	HMGN3	A	39	1.14 (1.09 to 1.20)	7.3×10^{-8}	0	1.03 (0.99 to 1.07)	0.17	1.07 (1.04 to 1.11)	7.8×10^{-5}	50	9.6×10^{-6}
rs5009270	7q31	IFRD1	A	30	1.15 (1.09 to 1.21)	5.7×10^{-7}	21	1.05 (1.00 to 1.11)	0.061	1.10 (1.06 to 1.14)	9.0×10^{-7}	46	3.1×10^{-6}
rs10773046	12q24	DNAH10	G	45	1.13 (1.08 to 1.19)	9.7×10^{-7}	10	1.01 (0.96 to 1.06)	0.72	1.07 (1.03 to 1.11)	2.0×10^{-4}	64	1.9×10^{-6}
rs17610181	17q23	NACA2	A	14	1.20 (1.11 to 1.29)	1.9×10^{-6}	32	1.05 (0.99 to 1.12	0.12	1.12 (1.06 to 1.18)	1.3×10^{-5}	42	7.5×10^{-6}
Female-specific analysis:													
rs10878630	12q15	DYRK2	A	41	1.19 (1.11 to 1.26)	1.2×10^{-7}	25	0.98 (0.91 to 1.05)	0.52	1.09 (1.04 to 1.14)	3.9×10^{-4}	65	1.1×10^{-5}
Male-specific analysis:													
rs12551314	9q22	PHF2	A	12	1.30 (1.18 to 1.44)	6.3×10^{-7}	28	0.98 (0.86 to 1.11)	0.72	1.16 (1.07 to 1.25)	2.9×10^{-4}	66	4.3×10^{-5}
rs3757837	7p13	CAMK2B	C	6	1.46 (1.26 to 1.69)	8.3×10^{-7}	46	1.15 (1.00 to 1.32)	0.044	1.27 (1.15 to 1.41)	2.2×10^{-6}	78	7.7×10^{-10}

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 8787 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 1386/7087 and 1451/4956 cases/controls in the replication stage for rs12551314 and rs3757837 respectively.
A1, coded/effect allele; Chr, chromosome; EAF, effect allele frequency; Gene, nearest gene; HE RE P, Han & Eskin Random Effects P; I^2, 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 $p=5.8\times10^{-8}$ and 7.3×10^{-8}, respectively, with no observed heterogeneity ($I^2=0$). The SNP rs6094710 replicated with $p=7.3\times10^{-4}$ in the follow-up samples and reached GWS level with $p=9.3\times10^{-9}$ 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, OR=1.15) in the follow-up samples, but did not reach GWS in the overall analysis with $p=2.2\times10^{-6}$ and OR=1.27. Moderate heterogeneity was observed for this SNP in the discovery analysis ($I^2=46\%$). Heterogeneity was increased, in the overall analysis with all the studies combined ($I^2=78\%$), reflecting the further heterogeneity introduced by the replication data. We, therefore, also applied the Han and Eskin (RE) model, an approach that allows more heterogeneity in the data compared with traditional models.[14] Using this model rs3757837 showed a stronger association with $p=7.7\times10^{-10}$ in this analysis. The strength of the association of this approach was not substantially different from the fixed-effect model for any other SNP in our study (table 2). rs3757837 resides in intron 8 of the *CAMK2B* (calcium/calmodulin-dependent protein kinase II β) gene.

The rs5009270 SNP at 7q31 remained suggestively associated in the final analysis ($p<5\times10^{-6}$) with combined two-stage $p=9.9\times10^{-7}$ and summary OR 1.10. rs5009270 is located near the *IFRD1* (interferon-related) gene. The p with the RE approach was 3.1×10^{-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 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 $\lambda_s=5^{20}$ then the discovered signals of OA contribute 3% of the heritability in OA

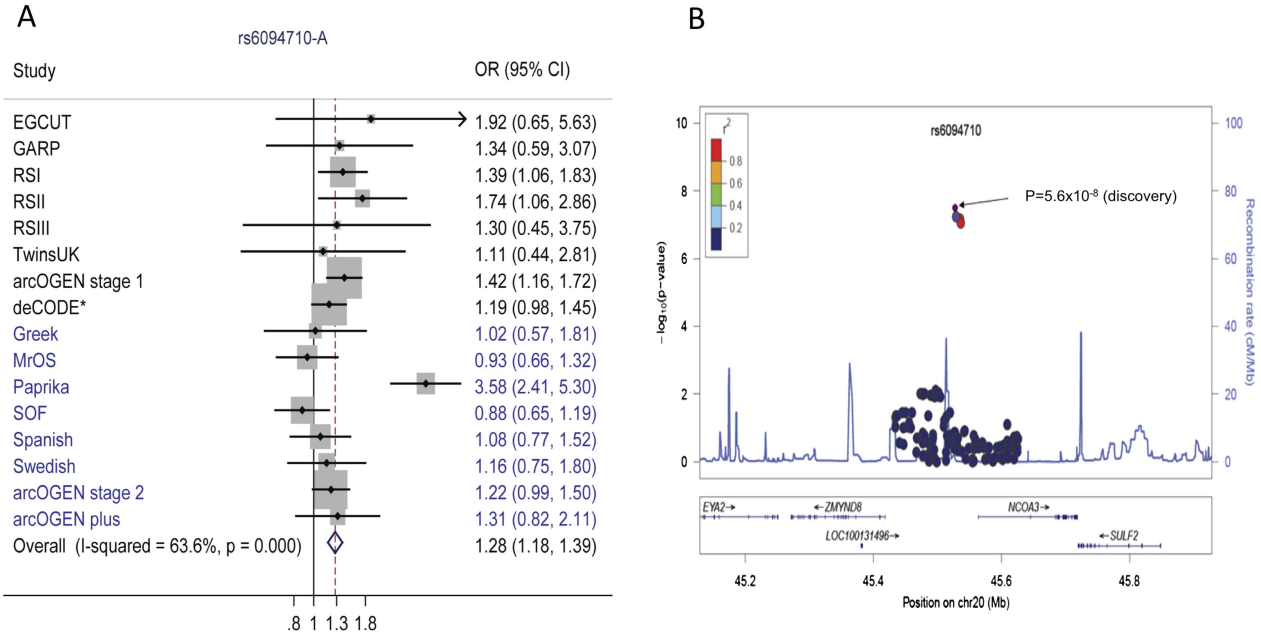


Figure 1 (A) Forest plot for meta-analysis of rs6094710. The blue diamond in the forest plot denotes the summary effect size and its edges the respective 95% CIs. Studies shaded in blue were included in the replication stage. *Discovery and replication estimate combined; and (B) regional plot for rs6094710 comprising directly genotyped and imputed SNPs. Case-control association results (−log10 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.

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

Reference	SNP	Gene	OR	EAF	λ_s	Expected GV explained (%)
Genome wide significant findings ($p<5\times10^{-8}$)						
This study	rs6094710	*NCOA3*	1.28	0.04	1.0029	0.46
arcOGEN[6]	rs6976	*GLT8D1*	1.12	0.37	1.0031	0.60
arcOGEN[6]	rs11177	*GNL3*	1.12	0.38	1.0031	0.60
arcOGEN[6]	rs4836732	*ASTN2*	1.20	0.47	1.0083	1.66
arcOGEN[6]	rs9350591	*FILIP1;SENP6*	1.18	0.11	1.0030	0.54
arcOGEN[6]	rs10492367	*KLHDC5;PTHLH*	1.14	0.19	1.0029	0.53
3arcOGEN[6]	rs835487	*CHST11*	1.13	0.34	1.0034	0.67
TreatOA[12]	rs12982744	*DOT1L*	1.17	0.38	1.0060	1.16
Suggestive findings ($5\times10^{-6}<p<5\times10^{-8}$)						
This study	rs5009270	*IFRD1*	1.10	0.30	1.0019	0.38
This study	rs3757837	*CAMK2B*	1.27	0.06	1.0040	0.64
arcOGEN[6]	rs12107036	*TP63*	1.21	0.52	1.0090	1.81
arcOGEN[6]	rs8044769	*FTO*	1.11	0.50	1.0027	0.54
arcOGEN[6]	rs10948172	*SUPT3H;CDC5L*	1.14	0.29	1.0037	0.71

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 $p<5\times10^{-6}$ 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 an amino acid change at position 218 in the protein (Arg>Cys). This amino acid change is predicted to have a benign and damaging effect on the protein by PolyPhen-2,[21] 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,[22] 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 25] the current *NCOA3* findings complement the previous outlined hypothesis that local T3 signalling may affect OA susceptibility.[26]

Another possible mechanism by which NCOA3 might be involved in cartilage homoeostasis 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 the common 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 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 *IFDRD1* 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
[1]Department of Hygiene and Epidemiology, University of Ioannina Medical School, Ioannina, Greece
[2]Department of Twin Research & Genetic Epidemiology, King's College London, London, UK
[3]Department of Internal Medicine, Erasmus Medical Center, Rotterdam, The Netherlands
[4]Department of Population Genetics, deCODE Genetics, Reykjavik, Iceland
[5]Department of Molecular Epidemiology, Leiden University Medical Center, Leiden, The Netherlands
[6]Netherlands Consortium for Healthy Ageing, The Netherlands
[7]Estonian Genome Center, University of Tartu, Tartu, Estonia
[8]Institute of Molecular and Cell Biology, University of Tartu, Tartu, Estonia
[9]California Pacific Medical Center Research Institute, San Francisco, USA
[10]Department of Human Genetics, Wellcome Trust Sanger Institute, Hinxton, UK
[11]NIHR Biomedical Research Unit and ARUK Centre of excellence for Sport, Exercise and Osteoarthritis, University of Oxford, Oxford, UK
[12]MRC Epidemiology Resource Centre, University of Southampton, Southampton, UK
[13]Worcestershire Royal Hospital, Worcestershire Acute Hospitals NHS Trust, Worcester, UK
[14]Centre of National Research on Disability and Rehabilitation Medicine, The University of Queensland, Brisbane, Australia
[15]Musculoskeletal Research Group, Institute of Cellular Medicine, Newcastle University, Newcastle upon Tyne, UK
[16]Wansbeck General Hospital, Northumbria Healthcare NHS Foundation Trust, Ashington, UK
[17]Rheumatology Division, Instituto de Investigación Biomédica-Hospital Universitario A Coruña, A Corunna, Spain
[18]Department of Academic Rheumatology, University of Nottingham, Nottingham, UK
[19]Department of Clinical Sciences Malmo, Lund University, Malmo, Sweden
[20]Department of Epidemiology, Erasmus Medical Center, Rotterdam, The Netherlands
[21]Department of Orthopedic Surgery, Akureyri Hospital, Akureyri, Iceland
[22]School of Health Sciences, University of Akureyri, Akureyri, Iceland
[23]Department of Medicine, The National University Hospital of Iceland, Reykjavik, Iceland
[24]Faculty of Medicine, University of Iceland, Reykjavik, Iceland
[25]Department of Public Health, University of Tartu, Tartu, Estonia
[26]Orthopedic Surgeons, Elva Hospital, Elva, Estonia
[27]Department of Orthopedics, Leiden University Medical Center, Leiden, The Netherlands
[28]Department of Rheumatology, Leiden University Medical Center, Leiden, The Netherlands
[29]Department of Quantitative Genetics, QIMR Berghofer Medical Research Institute, Brisbane, Australia
[30]Department of Genetic Epidemiology, QIMR Berghofer Medical Research Institute, Brisbane, Australia
[31]Department of Molecular Epidemiology, QIMR Berghofer Medical Research Institute, Brisbane, Australia
[32]Department of Epidemiology and Biostatistics, University of California, San Francisco, San Francisco, California, USA
[33]Centre for Integrated Genomic Medical Research, University of Manchester, Manchester, UK
[34]Worcestershire Acute Hospitals NHS Trust, Worcester, UK
[35]Institute of Genetics and Molecular Medicine, University of Edinburgh, Edinburgh, UK
[36]Department of Internal Medicine, Hospital U.M. Valdecilla-IFIMAV, University of Cantabria, Santander, Spain
[37]Laboratorio Investigacion 10 and Rheumatology Unit, Instituto de Investigacion Sanitaria—Hospital Clinico Universitario de Santiago, Santiago de Compostela, Spain
[38]Department of Biology, University of Thessaly, Medical School, Larissa, Greece
[39]Wellcome Trust Centre for Cell-Matrix Research, University of Manchester, Manchester, UK
[40]Department of Human Metabolism, University of Sheffield, Sheffield, UK
[41]Department of Medicine, University of California at Davis, Sacramento, USA
[42]Research Unit for Musculoskeletal Function and Physiotherapy, and Department of Orthopedics and Traumatology, University of Southern Denmark, Odense, Denmark
[43]Department of Clinical Sciences Lund, Lund University, Lund, Sweden
[44]Stanford 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, IJ, GT, HTH, UT and KS are employed by deCODE Genetics/Amgen. HJK is currently employed by Pfizer.

Ethics approval Meta-analysis of GWAS, different review boards approved the individual studies.

Provenance and peer review Not commissioned; externally peer reviewed.

REFERENCES

1 Lawrence RC, Felson DT, Helmick CG, *et al*. Estimates of the prevalence of arthritis and other rheumatic conditions in the United States. Part II. *Arthritis Rheum* 2008;58:26–35.
2 Valdes AM, Spector TD. The contribution of genes to osteoarthritis. *Med Clin North Am* 2009;93:45–66.
3 Spector TD, MacGregor AJ. Risk factors for osteoarthritis: genetics. *Osteoarthritis Cartilage* 2004;12(Suppl A):S39–44.
4 Jonsson H, Manolescu I, Stefansson SE, *et al*. The inheritance of hand osteoarthritis in Iceland. *Arthritis Rheum* 2003;48:391–5.
5 Kraus VB, Jordan JM, Doherty M, *et al*. The Genetics of Generalized Osteoarthritis (GOGO) study: study design and evaluation of osteoarthritis phenotypes. *Osteoarthritis Cartilage* 2007;15:120–7.
6 arcOGEN consosrtium; arcOGEN Collaborators. Identification of new susceptibility loci for osteoarthritis (arcOGEN): a genome-wide association study. *Lancet* 2012;380:815–23.
7 Day-Williams AG, Southam L, Panoutsopoulou K, *et al*. A variant in MCF2L is associated with osteoarthritis. *Am J Hum Genet* 2011;89:446–50.
8 Kerkhof HJ, Lories RJ, Meulenbelt I, *et al*. A genome-wide association study identifies an osteoarthritis susceptibility locus on chromosome 7q22. *Arthritis Rheum* 2010;62:499–510.
9 Evangelou E, Valdes AM, Kerkhof HJ, *et al*. Meta-analysis of genome-wide association studies confirms a susceptibility locus for knee osteoarthritis on chromosome 7q22. *Ann Rheum Dis* 2011;70:349–55.
10 Valdes AM, Evangelou E, Kerkhof HJ, *et al*. The GDF5 rs143383 polymorphism is associated with osteoarthritis of the knee with genome-wide statistical significance. *Ann Rheum Dis* 2011;70:873.
11 Castano Betancourt MC, Cailotto F, Kerkhof HJ, *et al*. Genome-wide association and functional studies identify the DOT1L gene to be involved in cartilage thickness and hip osteoarthritis. *Proc Natl Acad Sci USA* 2012;109:8218–23.
12 Evangelou E, Valdes AM, Castano-Betancourt MC, *et al*. The DOT1L rs12982744 polymorphism is associated with osteoarthritis of the hip with genome-wide statistical significance in males. *Ann Rheum Dis* 2013;72:1264–5.
13 Higgins JP, Thompson SG, Deeks JJ, *et al*. Measuring inconsistency in meta-analyses. *BMJ* 2003;327:557–60.
14 Han B, Eskin E. Random-effects model aimed at discovering associations in meta-analysis of genome-wide association studies. *Am J Hum Genet* 2011;88:586–98.
15 Lango Allen H, Estrada K, Lettre G, *et al*. Hundreds of variants clustered in genomic loci and biological pathways affect human height. *Nature* 2010;467:832–8.
16 Speliotes EK, Willer CJ, Berndt SI, *et al*. Association analyses of 249,796 individuals reveal 18 new loci associated with body mass index. *Nat Genet* 2010;42:937–48.
17 International HapMap C, Frazer KA, Ballinger DG, *et al*. A second generation human haplotype map of over 3.1 million SNPs. *Nature* 2007;449:851–61.
18 Marchini J, Howie B, Myers S, *et al*. A new multipoint method for genome-wide association studies by imputation of genotypes. *Nat Genet* 2007;39:906–13.
19 Smyth G. Linear models and empirical bayes methods for assessing differential expression in microarry experiments. *Stat Appl Genet Mol Biol* 2004;3:Article3.
20 Valdes AM, Spector TD. Genetic epidemiology of hip and knee osteoarthritis. Nature reviews. *Rheumatology* 2011;7:23–32.
21 Adzhubei IA, Schmidt S, Peshkin L, *et al*. A method and server for predicting damaging missense mutations. *Nat Methods* 2010;7:248–9.
22 Xu J, Liao L, Ning G, *et al*. The steroid receptor coactivator SRC-3 (p/CIP/RAC3/AIB1/ACTR/TRAM-1) is required for normal growth, puberty, female reproductive function, and mammary gland development. *Proc Natl Acad Sci USA* 2000;97:6379–84.
23 Nicholls JJ, Brassill MJ, Williams GR, *et al*. The skeletal consequences of thyrotoxicosis. *J Endocrinol* 2012;213:209–21.
24 Meulenbelt I, Bos SD, Chapman K, *et al*. Meta-analyses of genes modulating intracellular T3 bio-availability reveal a possible role for the DIO3 gene in osteoarthritis susceptibility. *Ann Rheum Dis* 2011;70:164–7.
25 Meulenbelt I, Min JL, Bos S, *et al*. Identification of DIO2 as a new susceptibility locus for symptomatic osteoarthritis. *Hum Mol Genet* 2008;17:1867–75.
26 Bos SD, Slagboom PE, Meulenbelt I. New insights into osteoarthritis: early developmental features of an ageing-related disease. *Curr Opin Rheumatol* 2008;20:553–9.
27 Arimura A, vn Peer M, Schroder AJ, *et al*. The transcriptional co-activator p/CIP (NCoA-3) is up-regulated by STAT6 and serves as a positive regulator of transcriptional activation by STAT6. *J Biol Chem* 2004;279:31105–12.
28 Darnell JE Jr, Kerr IM, Stark GR. Jak-STAT pathways and transcriptional activation in response to IFNs and other extracellular signaling proteins. *Science* 1994;264:1415–21.
29 Millward-Sadler SJ, Khan NS, Bracher MG, *et al*. Roles for the interleukin-4 receptor and associated JAK/STAT proteins in human articular chondrocyte mechanotransduction. *Osteoarthritis Cartilage* 2006;14:991–1001.
30 Shimazaki A, Wright MO, Elliot K, *et al*. Calcium/calmodulin-dependent protein kinase II in human articular chondrocytes. *Biorheology* 2006;43:223–33.
31 Huser CA, Davies ME. Calcium signaling leads to mitochondrial depolarization in impact-induced chondrocyte death in equine articular cartilage explants. *Arthritis Rheum* 2007;56:2322–34.
32 Sun HB. Mechanical loading, cartilage degradation, and arthritis. *Ann NY Acad Sci* 2010;1211:37–50.
33 Micheli L, Leonardi L, Conti F, *et al*. PC4/Tis7/IFRD1 stimulates skeletal muscle regeneration and is involved in myoblast differentiation as a regulator of MyoD and NF-kappaB. *J Biol Chem* 2011;286:5691–707.
34 Trudelle-Jackson E, Ferro E, Morrow JR Jr. Clinical implications for muscle strength differences in women of different age and racial groups: the WIN study. *J Womens Health Phys Therap* 2011;35:11–18.
35 Kerkhof HJ, Meulenbelt I, Akune T, *et al*. Recommendations for standardization and phenotype definitions in genetic studies of osteoarthritis: the TREAT-OA consortium. *Osteoarthritis Cartilage* 2011;19:254–64.
36 Genomes Project C, Abecasis GR, Auton A, *et al*. An integrated map of genetic variation from 1092 human genomes. *Nature* 2012;491:56–65.
37 Kutyavin IV, Milesi D, Belousov Y, *et al*. A novel endonuclease IV post-PCR genotyping system. *Nucleic Acids Res* 2006;34:e128.

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

Supplementary Material

Content

1. Supplementary Methods

a. Genotyping, data quality control and data imputation

To allow for meta-analysis across different marker sets, imputation of polymorphic HapMap European CEU SNPs was performed using MACH or IMPUTE [1]. Two research centres (Ioannina, Greece and Erasmus MC Rotterdam, the Netherlands) performed both the Quality Control (QC) and meta-analyses. A QC protocol was set up including validation of the results file format, reports for range of values and elimination of potential biases (i.e., extremely large beta's or SEs). Files were cross-validated between the two research centers after QC and after meta-analyses to check for inconsistencies. SNPs with a MAF <1%, imputation quality <0.30 (MACH) or <0.40 (IMPUTE) and beta's >4 or <-4 were excluded for further analysis.

b. Statistical analysis

The principal summary measure of association was the per-allele odds ratio (OR). We performed genomic control at the individual study level estimates; for each study, we recorded the inflation factor lambda for the study so as to adjust the standard error of the effect size (standard error is multiplied by the square root of lambda).

We summarized OR estimates using fixed-effects models [2]. Fixed-effects models assume that there is a common underlying effect and the variability observed is attributed to chance alone; random effects models acknowledge that true between-study heterogeneity exists, take into account the presence of heterogeneity into their calculations and, in the presence of heterogeneity, yield more conservative estimates. In the absence of heterogeneity, fixed- and random-effects models yield the same results. Fixed-effects models are more appropriate at the SNP discovery and prioritization stage and perform well at the replication stage. The presence of statistically significant heterogeneity was assessed by the Q statistic (significant at $p < 0.10$) and the extent of the observed heterogeneity was assessed by the I^2 (ranging from 0% to 100%) [3]. We also summarized OR estimates under a random-effects model proposed by Han and Eskin [4].

c. Study participants

The Rotterdam Study I, II, & III: The study population comprises men and women aged 55 years and older of the Rotterdam Study, which is a prospective population-based study on determinants of chronic disabling diseases. It consists of three sub-popualtions and the rationale and study design have been described previously [5]. The medical ethics committee of Erasmus University Medical School approved the study and written informed consent was obtained from each participant. Hip OA cases were defined as a KL grade≥2 or total hip replacement.

deCODE: a list of patients with OA of hip was obtained on the basis of patients' records at hospitals and health care centers in Iceland [6]. Controls were individuals with no external signs of OA in any joint who did not have a diagnosis of primary OA. The study was approved by the Data Protection Authority of Iceland and the National Bioethics Committee of Iceland. Informed consent was obtained from all participants.

TwinsUK: the study participants were white monozygotic and dizygotic twin pairs from the TwinsUK adult twin registry, a group used to study the heritability and genetics of age-related diseases [7]. These unselected twins were recruited from the general population through national media campaigns in the United Kingdom. Ethics approval was obtained from the Guy's and St. Thomas' Hospital Ethics Committee. Written informed consent was obtained from every participant.

The Genetics OsteoArthritis and Progression (GARP) study from Leiden, the Netherlands, consisted of 192 sibling pairs concordant for clinical and radiographically (K/L score) confirmed OA at two or more joint sites among hand, spine (cervical or lumbar), knee or hip [8], random controls (N=758) were partners of the offspring of the Leiden longevity study [9]. To comply with the discovery sample OA phenotypes for knee, hip and hand OA used were based on radiographic signs OA Written informed consent was obtained from each subject as approved by the ethical committees of the Leiden University Medical Center.

arcOGEN study

arcOGEN stage 1: The arcOGEN case samples were collected in two stages. The stage 1 samples comprised 1,728 hip cases from existing DNA collections from five United Kingdom locations within the arcOGEN consortium (London, Nottingham, Oxford, Sheffield, and Southampton). The detailed characteristics of these cases are described elsewhere [10]. Briefly, all were unrelated and of European origin, and all had primary OA of the hip of radiographic Kellgren-Lawrence (KL) grade ≥2, or clinical evidence of disease to a level requiring total joint replacement (TJR). The stage 1 study used 4,894 population-based UK controls from an early release of the Wellcome Trust Case Control Consortium 2 (WTCCC2) data which came from 2 distinct sources: the 1958 Birth Cohort [58BC] and the UK Blood Donor Service (UKBS) and were unrelated (www.wtccc.org.uk).

arcOGEN stage 2: The stage 2 cases (n=1,763 with hip OA) were collected prospectively as part of the arcOGEN study at nine locations across the UK (Edinburgh, London, Newcastle-Upon-Tyne, Nottingham, Oxford, Sheffield, Southampton, Wansbeck, and Worcester) [11]. The ascertainment criterion was primary OA that was severe enough for the individual to require joint replacement of the hip. All cases were unrelated and of European origin. The controls (n=6,157) were population-based, unrelated UK controls which came from five distinct sources: the 1958 Birth Cohort from the Type 1 Diabetes Genetics Consortium (T1DGC) study, the Avon Longitudinal Study of Parents

and Children (ALSPAC), the People of the British Isles (PoBI) study and additional controls from the 58BC and the UKBS from the WTCCC2 study that were not overlapping with those used in stage 1.

arcOGEN plus: The arcOGEN plus dataset (n=223 females with hip OA) comprises additional cases collected in stage 2 which were genotyped at a later stage. The ascertainment criterion was primary OA that was severe enough for the individual to require joint replacement of the hip. Controls (n=1,828) were unrelated, OA-free controls (females only) from the TwinsUK cohort which consist of twins ascertained to study the heritability and genetics of age-related diseases (www.twinsUK.ac.uk). Samples that overlapped with the TwinsUK dataset used in the discovery analysis were excluded from this study.

Estonian Genome Center, University of Tartu (EGCUT). The Estonian cohort is from the population based biobank of the Estonian Genome Project of University of Tartu. The whole project is conducted according to the Estonian Gene Research Act and all participants have signed the broad inform consent. The current cohort size is over 51,515, from 18 years of age and over, which reflects closely the age distribution of the adult Estonian population. Subjects were recruited randomly when visiting general practitioners (GPs) and hospitals. Each participant filled out a Computer Assisted Personal interview during 1-2 hours at doctors; office, including personal data. OA was diagnosed by a specialist as a clinical finding and was usually confirmed by a radiograph (KL grade>2). The OA cases for the current study had an ICD10 M16 and/or M17 diagnosis.

Greek case-control study: The individuals included in the study were of Greek origin living in the district of Thessalia in central Greece [12]. All of them had undergone a TKR/THR, meaning that all of them suffered from severe knee or hip OA, which is defined by a K/L grade >=2. None of the patients had evidence of arthritis due to another disease. All the controls had a K/L score of 0 and had undergone treatment for injuries or fractures. Patients with rheumatoid arthritis and other autoimmune diseases as well asmchondrodysplasias, infection-induced OA, and posttraumatic OA were not included in the study. The ethics committee of the Larissa University Hospital approved this study and all individuals gave their informed consent.

Spanish TJR cases: Patients were selected from consecutive patients, aged 55-75 years of age at time of the surgery, undergoing THR/TKR [13]. All patients were included if a rheumatologist considered them to suffer from severe primary OA. Exclusion criteria were inflammatory, infectious, traumatic or congenital joint pathology and lesions due to crystal deposition or osteonecrosis. Controls were recruited among subjects older than 55 years of age undergoing preoperative work-up for elective surgeries other than joint surgery and who did not show clinical manifestations of OA. This study was approved by the Ethical Committee for Clinical Research of Galicia and all cases and controls gave their written informed consent to participate.

Swedish Malmo Diet and Cancer (MDC) study: All men and women living in the city of Malmö in Sweden, who were born between 1923 and 1945 (men) or between 1923 and 1950 (women) were invited to participate in the Malmö Diet and Cancer (MDC) study. The screening examination was performed during 1991-1996. All participants (n=28449) were followed until first OA surgery, emigration from Sweden, death or December 31 2005, whichever came first. Hip osteoarthritis was defined as a first hip arthroplasty (procedures coded 8414, 8010, NFB09, NFB19, NFB29, NFB39, NFB49 and NFB99) in combination with a contemporaneous diagnosis of hip osteoarthritis (715 or M16 according to ICD-9 and ICD-10, respectively). Cases were matched (1:1) for age, gender and BMI, to MDC participants without THR in a nested case-control design.

Osteoporotic Fractures in Men Study (MrOS): The Osteoporotic Fractures in Men Study (MrOS) is a multi-center prospective, longitudinal, observational study of risk factors for vertebral and all non-vertebral fractures in older men, and of the sequelae of fractures in men [14,15]. The original specific aims of the study include: (1) to define the skeletal determinants of fracture risk in older men, (2) to define lifestyle and medical factors related to fracture risk, (3) to establish the contribution of fall frequency to fracture risk in older men, (4) to determine to what extent androgen and estrogen concentrations influence fracture risk, (5) to examine the effects of fractures on quality of life, (6) to identify sex differences in the predictors and outcomes of fracture, (7) to collect and store serum, urine and DNA for future analyses as directed by emerging evidence in the fields of aging and skeletal health, and (8) define the extent to which bone mass/fracture risk and prostate diseases are linked. The MrOS study population consists of 5,994 community dwelling, ambulatory men aged 65 years or older from six communities in the United States (Birmingham, AL; Minneapolis, MN; Palo Alto, CA; Monongahela Valley near Pittsburgh, PA; Portland, OR; and San Diego, CA). Inclusion criteria were designed to provide a study cohort that is representative of the broad population of older men. The inclusion criteria were: (1) ability to walk without the assistance of another, (2) absence of bilateral hip replacements, (3) ability to provide self-reported data, (4) residence near a clinical site for the duration of the study, (5) absence of a medical condition that (in the judgment of the investigator) would result in imminent death, (6) ability to understand and sign an informed consent, and (7) 65 years or older. To qualify as an enrollee, the participant had to provide written informed consent, complete the self-administered questionnaire (SAQ), attend the clinic visit, and complete at least the anthropometric, DEXA, and vertebral X-ray procedures. The MrOS cohort recruited only men.

Study of Osteoporotic Fractures (SOF): The Study of Osteoporotic Fractures (SOF) is a prospective multicenter study of risk factors for vertebral and non vertebral fractures[16]. The cohort is comprised of 9704 community – dwelling women 65 years old or older recruited from populations-based listings in four U.S. areas: Baltimore, Maryland; Minneapolis, Minnesota; Portland, Oregon; and the Monongahela Valley, Pennsylvania. Women enrolled in the study were 99% Caucasian with African American women initially excluded from the study due to their low incidence of hip fractures. A cohort of AA women was recruited at the 6th Visit. The SOF participants were followed up every four months by postcard or telephone to ascertain the occurrence of falls, fractures and changes in address. To date, follow-up rates have exceeded 95% for vital status and fractures. All fractures are validated by x-ray reports or, in the case of most hip fractures, a review of pre-operative radiographs. The inclusion criteria were: 1) 65 years or older, (2) ability to walk without the assistance of another, (3) absence of bilateral hip replacements, (4) ability to provide self-reported data, (5) residence near a clinical site for the duration of the study, (6) absence of a medical condition that (in the judgment of the investigator) would result in imminent death, and (7) ability to understand and sign an informed consent. To qualify as an enrollee, the participant had to provide written informed consent, complete the self-administered questionnaire (SAQ), attend the clinic visit, and complete at least the anthropometric measures. The SOF study recruited only women

Paprika study: The Paprika study is performed at the Leiden University Medical Center (Dept. Orthopedics) and consists in a long-term follow-up study of patients that have undergone total joint replacement (TJR) at hip or knee [17-19] and has been approved by the medical ethical committee. Patients of Caucasian descent were included when they were diagnosed with primary osteoarthritis based on radiographs and the ACR rheumatology classification criteria (mean age males-hip: 66; years males-knee: 68 years; females-hip: 66 years; females-knee: 69 years). Patients

with secondary OA or requiring a revision were excluded in this study. Written consent was obtained from each participant.

Genotyping.

The Rotterdam Study I, II & III: Genotyping of the samples with the Illumina HumanHap550v3 Genotyping BeadChip was carried out at the Genetic Laboratory of the Department of Internal Medicine of Erasmus Medical Center, Rotterdam, the Netherlands. The Beadstudio GenCall algorithm was used for genotype calling and quality control procedures were as described previously [20]. The following quality control filters were applied: SNP call rate >= 95%, minor allele frequency >= 5%, p-value HWE >=1x10-6. After quality control 500,510 SNPs remained for association analyses. The intensity cluster plots were visually inspected for the top-hits of the Rotterdam Study and no abnormalities were discovered. Genomic inflation factors were calculated for all analyses and there was no evidence of population stratification with lambdas of 1.01 for hip- and hand-OA, 1.00 for knee-OA

deCODE: All samples were assayed with the Infinium HumanHap 300 or humanCNV370 SNP chips (Illumina), containing 317,503 tagging SNPs derived from phase I of the International HapMap project. All of the SNPs tested in this report passed quality filtering (a call rate >97%, a minor allele frequency >1%, not a significant distortion from HWE (p-value >10-7 on any of the three chip types used (humanHap300, humanHap300-duo and humanCNV370). Any samples with a yield <98% were excluded from the analysis. Imputation was done using the IMPUTE software [1]. The additional cases in the replication analysis were genotyped using the Centaurus (Nanogen) platform

TwinsUK: Samples were genotyped with the Infinium HumanHap 300 assay (Illumina, San Diego, USA) at the Duke University Genotyping Center (NC USA), Helsinki University (Finland) and the Wellcome Trust Sanger Institute. The Illuminus calling algorithm was used for genotype calling. After strict quality control criteria were applied as described in [20] there were 314075 SNPs available for analysis. Imputation was performed using the IMPUTE software (v0.2.0) [1]. At imputed loci, all genotypes with posterior probabilities < 0.9 were discarded and the imputed loci were filtered out using usual QC filters.

Genetics OsteoArthritis and Progression (GARP) Study: For the GARP study the genome wide scan was genotyped by Illumina Infinuum II HumanHap 55KL Beadchips and Illumina Infinium II HumanHap550-Duo BeadChips (Illumina, San Diego, USA), respectively. Genotypes from the SNPs from the HapMap phase II v21 were imputed using IMPUTE.

arcOGEN study: arcOGEN stage 1 and stage 2 cases were genotyped using Illumina Human 610-Quad BeadChips. The publically available controls used for stage 1 and for stage 2 were genotyped on a variety of platforms (Table 1) [10,11]. ArcOGEN plus cases were genotyped on the Illumina HumanOmniExpress platform. This study used TwinsUK disease-free controls which were genotyped on Illumina Human 610-Quad BeadChips. All datasets underwent QC at the sample and SNP level separately for each case and control cohorts as previously [10,11]. Briefly samples were excluded if their call rate was <97% and if they showed gender discrepancies (estimated from genotypic data against external information). Individuals were also excluded on the basis of excess genome-wide heterozygosity or homozygosity. We identified samples that were accidentally duplicated or closely-related by calculating genome-wide IBD (given IBS information) for pairs of individuals. Multidimensional scaling (MDS) was performed in conjunction with data from the three HapMap phase II populations in order to identify and exclude individuals of non-

European descent. SNPs were excluded from further analysis based on the following criteria: Call rate <95% if minor allele frequency (MAF)≥5% or call rate <99% if MAF<5%, HWE exact p values <0.0001 in cases or controls, and MAF <1%. Association analyses were carried out under the additive model. Imputation was carried out using IMPUTE and imputed genotypes were analysed taking under account the full genotype probability distribution.

Osteoporotic Fractures in Men Study (MrOS) and Study of Osteoporotic Fractures (SOF): The Illumina HumanOmni1_Quad_v1-0 B was used for whole-genome genotyping. Samples from SOF and MrOS were randomized to 96-well genotyping plates by sex and clinic site. Eighty-one samples were plated twice to assess reproducibility. Pairwise concordance was 100%. 119 replicates of samples from HapMap trios of CEU and YRI populations and singletons from CHB and JPT populations were genotyped alongside MrOS and SOF samples, and compared to published HapMap genotypes. Concordance was 99.7% for CEU and YRI samples and was 95.0-99.7% for CHB and JPT samples. Genotypes were called using a clustering algorithm in Illumina's BeadStudio software at the Broad Institute. Samples with call rates < 97% were excluded. SNPs with GenTrain scores <0.6, cluster separation scores <0.4, call rates <97%, or MAF <0.01 were excluded. Autosomal SNPs with HWE P-value $<10^{-4}$ were excluded. In addition, genotype clusters for SNPs on chrX, chrY, chrXY and chrMT were reviewed manually. For MrOS and SOF samples, 740,713 SNPs passed QC.Additional samples were excluded based on: (1) genotypic sex mismatch using X and Y chromosome probe intensities, (2) relatedness among genotyped samples using the kinship coefficient that estimates probability that alleles are identical-by-descent, and (3) gross chromosomal abnormalities detected using the LogR Ratio and B allele frequency. Among the 3924 SOF samples that underwent whole-genome genotyping, 3682 samples had acceptable call rates. Among these 3682 SOF samples, 4 were removed due to relatedness and 53 were removed due to gross chromosomal abnormalities, leaving 3625 SOF samples with whole genome genotyping data that passed QC. Among the 5506 MrOS samples that underwent whole-genome genotyping, 5189 samples had acceptable call rates. Among these 5189 MrOS samples, 1 was removed due to relatedness and 37 were removed due to gross chromosomal abnormalities, leaving 5151 MrOS samples with whole genome genotyping data that passed QC. SNPs and samples that passed QC filters underwent SNP genotype imputation using minimac. HapMap phase II release 22 build 36 consensus phased haplotypes from a combined panel of CEU, YRI, CHB, and JPT HapMap samples were used as a reference panel

Estonian Genome Center, University of Tartu (EGCUT). All samples were genotyped with Illumina HumanCNV370 or HumanOmniExpress (Illumina, San Diego, USA) according to the Illumina protocol in the Estonian Biocenter. Data quality control was perform with PLINK (SNP call rate>98%;sample call rate>95%; MAF>0.01; HWE $p>10^{-6}$; cryptic relatedness). Imputation was performed with IMPUTE v1.0 (CEU HapMap rel22 build 36) and association analyses were carried out with SNPTEST. Inflation factors for directly genotyped and imputed data were 1.01 and 1.01 respectively.

Paprika study: In the present work, genotyping of the Paprika study was performed using the Sequenom MassARRAY iPLEX Gold or Taq-Man SNP Genotyping assays following the manufacturer's instructions. All SNPs passed the following quality criteria: call rate >98% and p-value for Hardy-Weinberg equilibrium <10-4.

d. De novo genotyping for imputed SNPs rs6094710 and rs640070

Imputed SNPs rs6094710 and rs640070 had MAF <4% and even though they passed the imputation quality criteria set upfront therefore they were de novo genotyped to minimize the possibility of imputation errors. Random samples of three populations from TWINS UK (n=392), arcOGEN (n=1046) and Rotterdam (n=865) were used for the assessment. For rs6094710 the concordance was 97.7%, 98.9% and 99.1% respectively. Poor concordance was found for rs640070 (<60% in all cases) and therefore it was excluded from further consideration.

e. Heritability of the identified markers of hip OA

We searched Pubmed for variants that have been identified as susceptible for hip OA in European populations. Only articles in English were eligible. We retrieved the hits that were GWS ($P<5x10^{-8}$) or reported as suggestive signals by the authors of the studies. From each study we recorded the study, the eligible variant, the risk allele frequency and the OR. We then calculated the sibling recurrence risk ratio attributed to these markers by using the formula $\lambda_s = \left(1 + \frac{pq(\gamma-1)^2}{2(p+\gamma q)^2}\right)^2$ where q is the risk allele frequency, p=1-q, γ=genotype relative risk under the log-additive model. The expected genetic variance explained was calculated as described in Ju-Hyun P et al [21]

2. Supplementary figures

FIGURE LEGENDS

Figure S1: Manhattan plot for the combined analysis of the hip OA GWAs meta-analysis.

Figure S2: QQ plot for the combined analysis of the hip OA GWAs meta-analysis. The expected p-value is indicated by the solid line and the associated 95% confidence intervals are indicated by the blue area either side

Figure S3: QQ plot for the A) female-specific and B) male-specific analysis of the hip OA GWAs meta-analysis. The expected p-value is indicated by the solid line and the associated 95% confidence intervals are indicated by the blue area either side

Figures S4-S11: Forest plots for the 8 SNPs that were followed-up in the 2nd stage of the analysis. The blue diamond denotes the summary effect size and its edges the respective 95% confidence intervals. Studies shaded in blue were included in the replication stage. * Discovery and replication estimate combined.

Figure S1

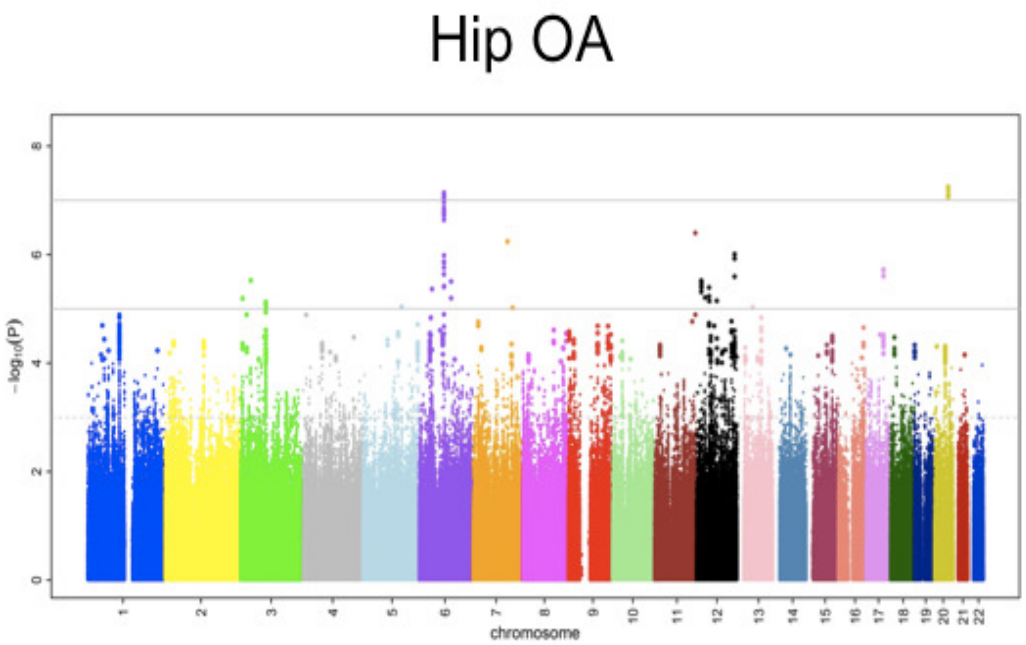


Figure S2

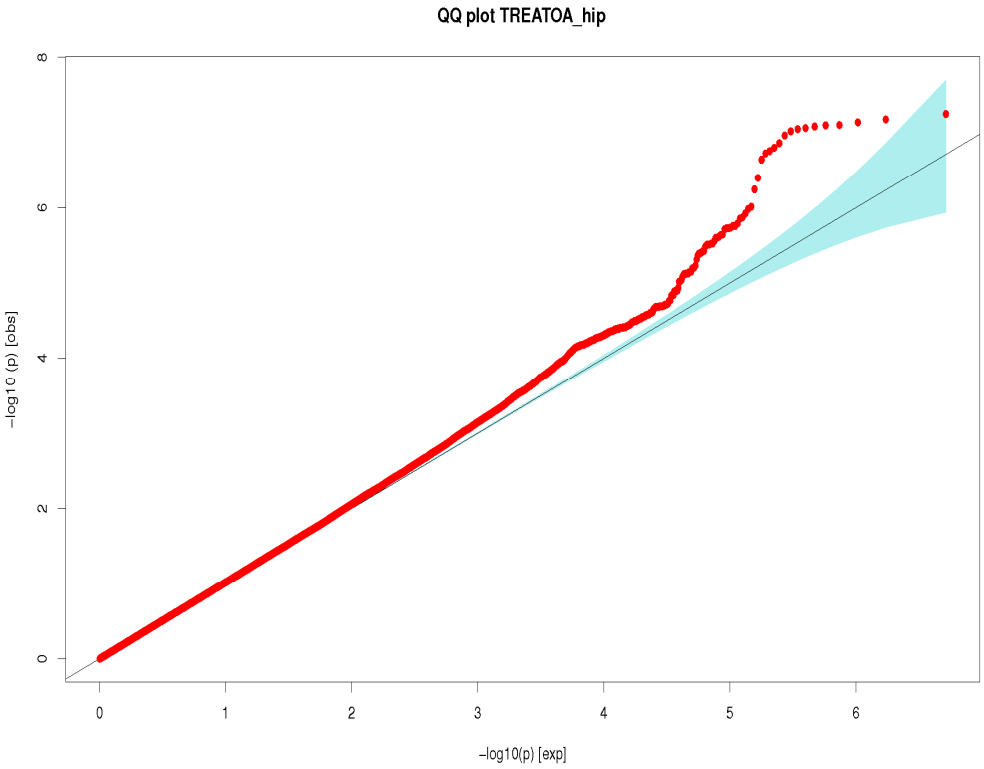


Figure S3

A)


QQ plot TREATOA_hip_females
−log10 (p) [obs]
−log10(p) [exp]
0
2
4
6
8
0
1
2
3
4
5
6

B)

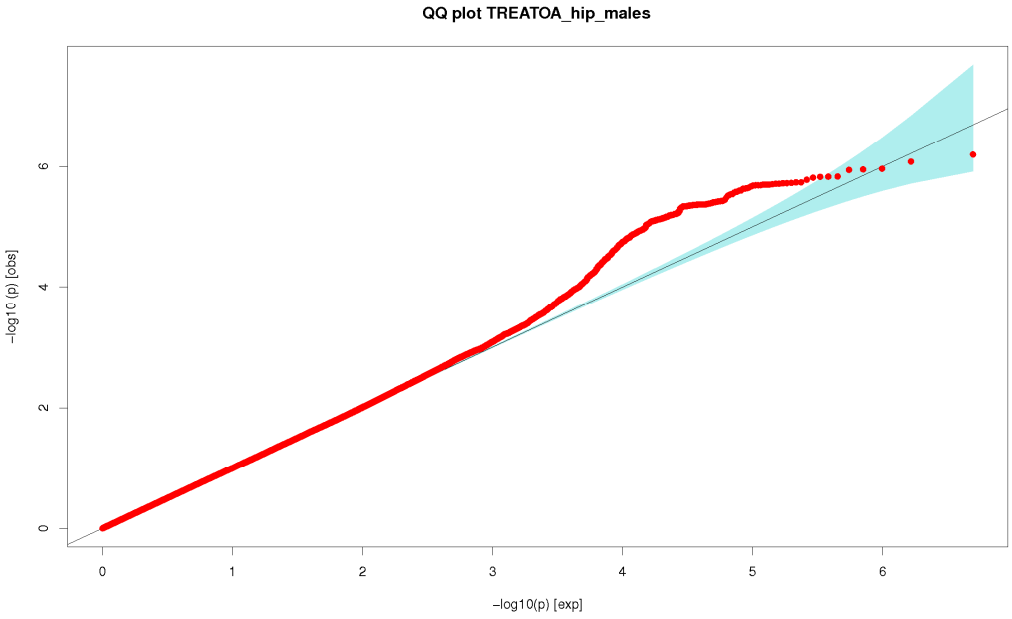

QQ plot TREATOA_hip_males
−log10 (p) [obs]
−log10(p) [exp]
0
2
4
6
0
1
2
3
4
5
6

Figure S4

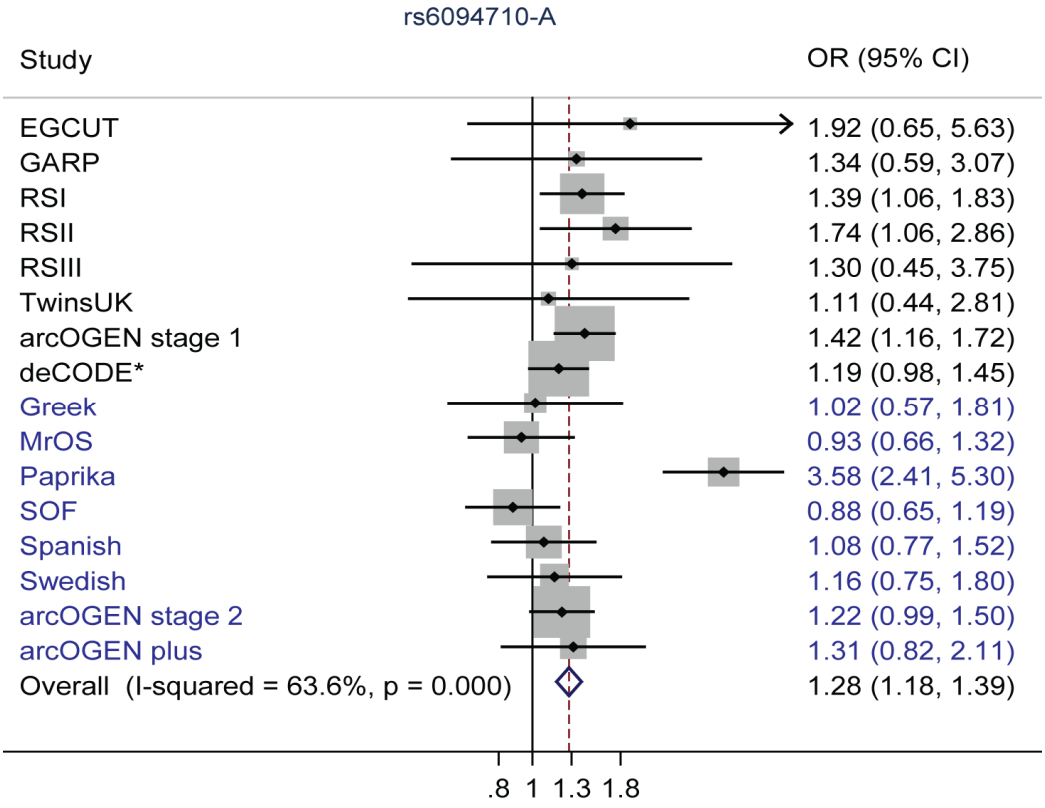


Figure S5

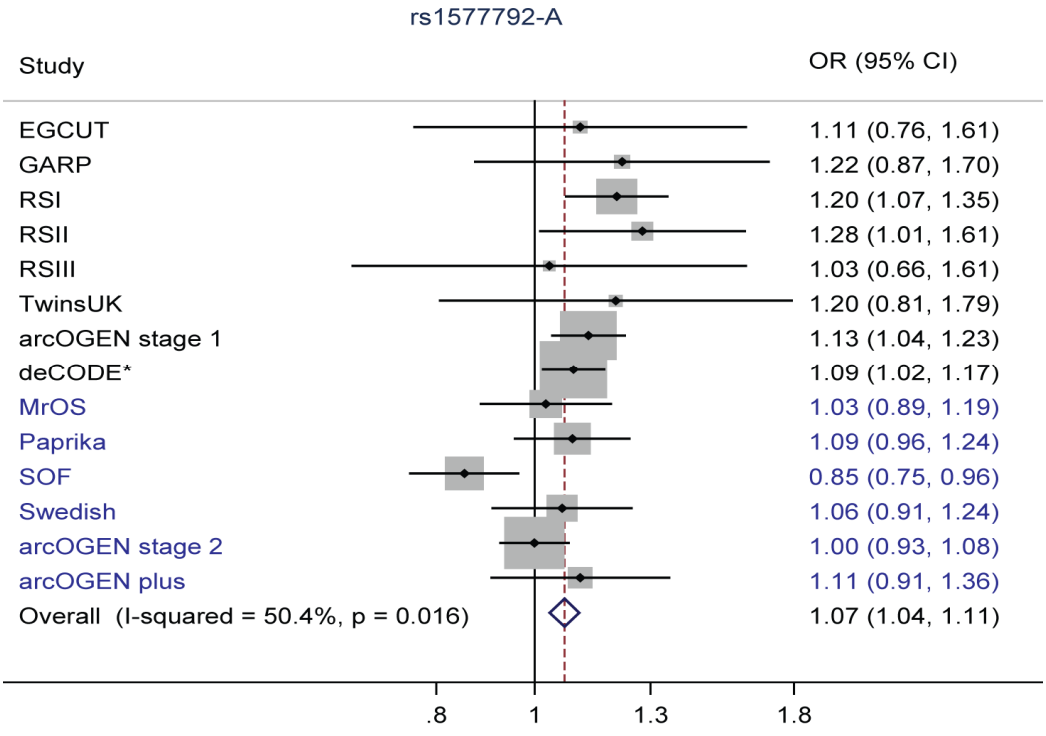


Figure S6

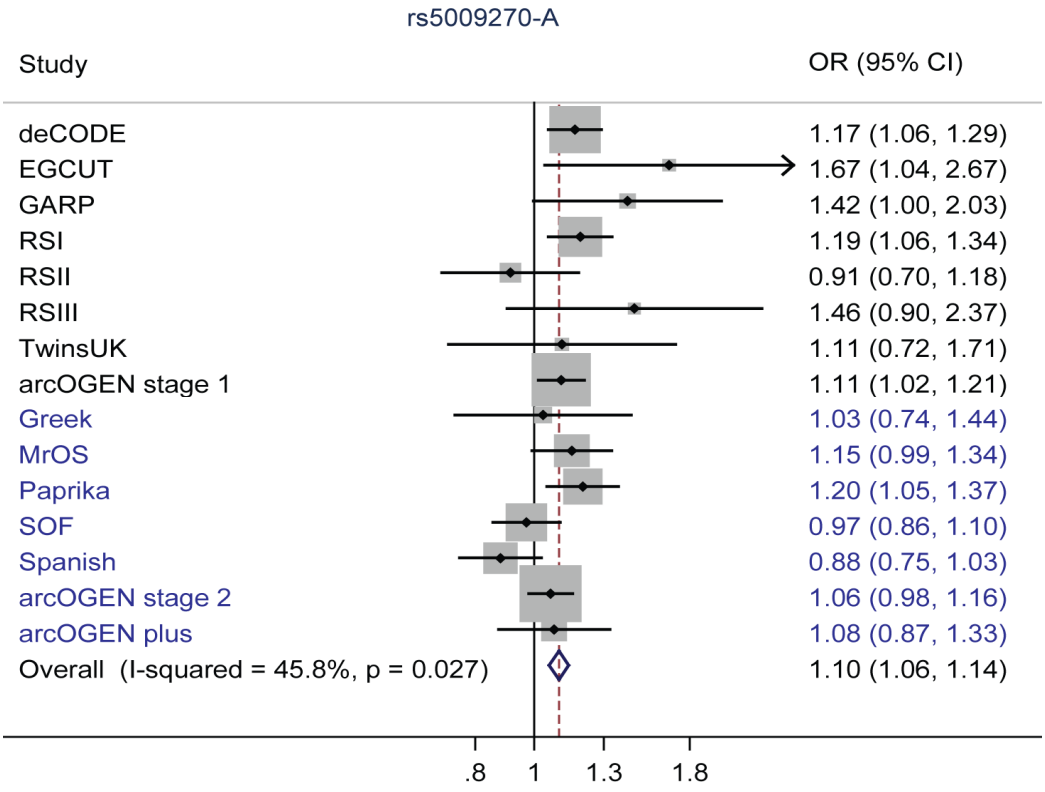


Figure S7

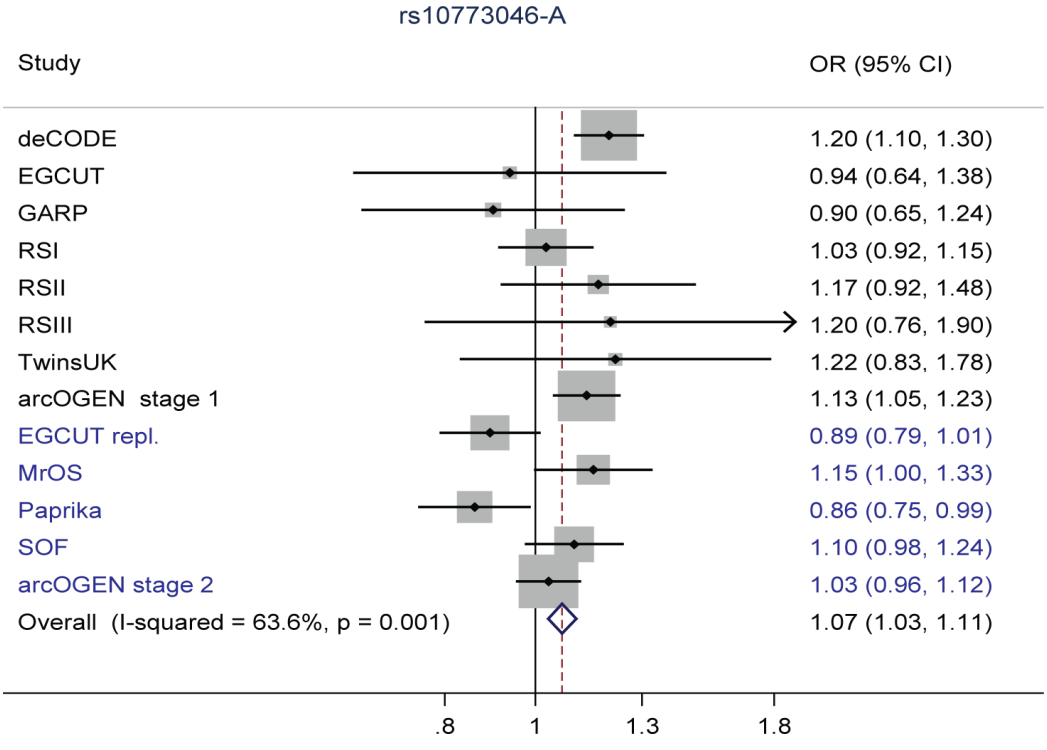


Figure S8

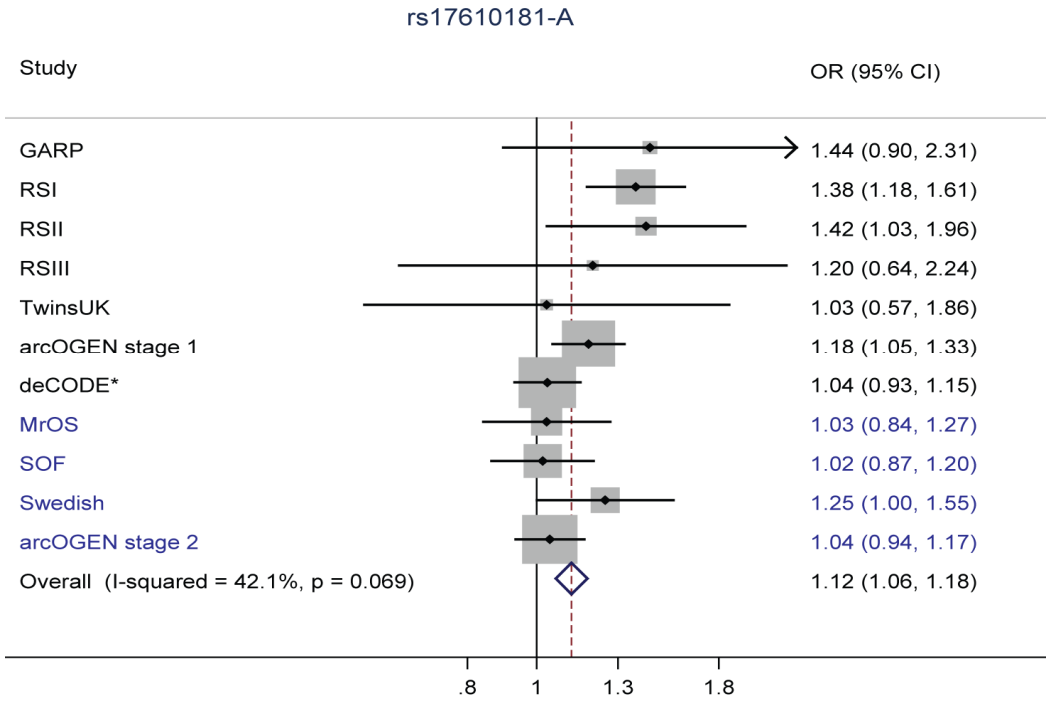


Figure S9

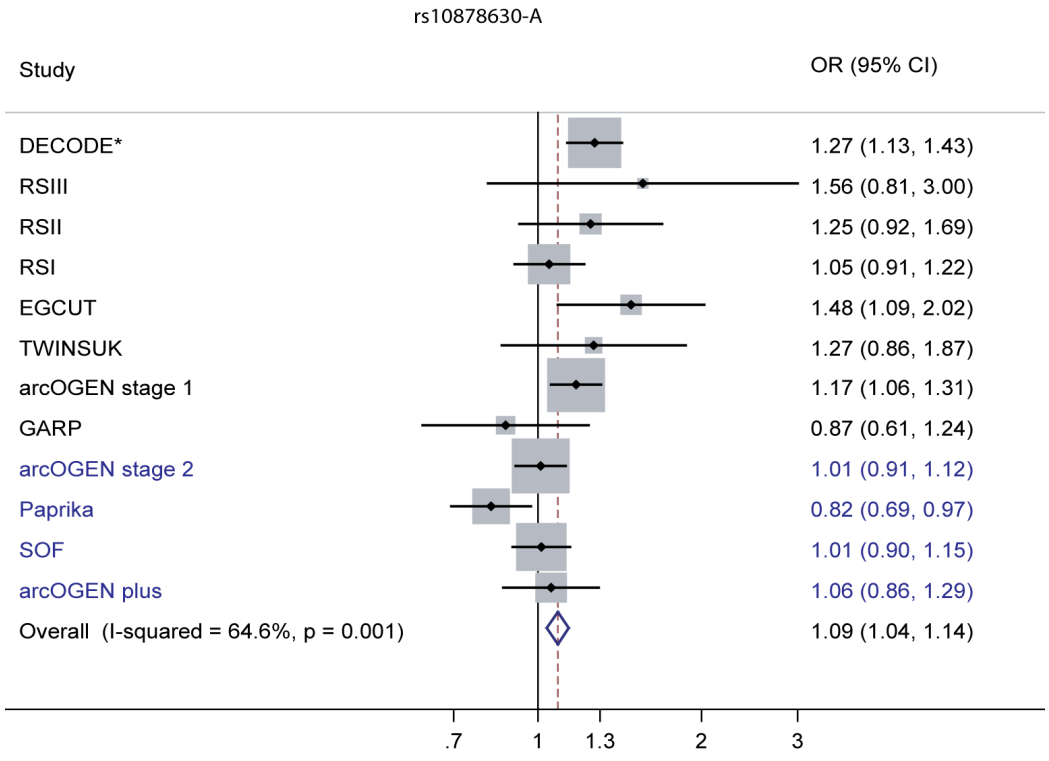


Figure S10

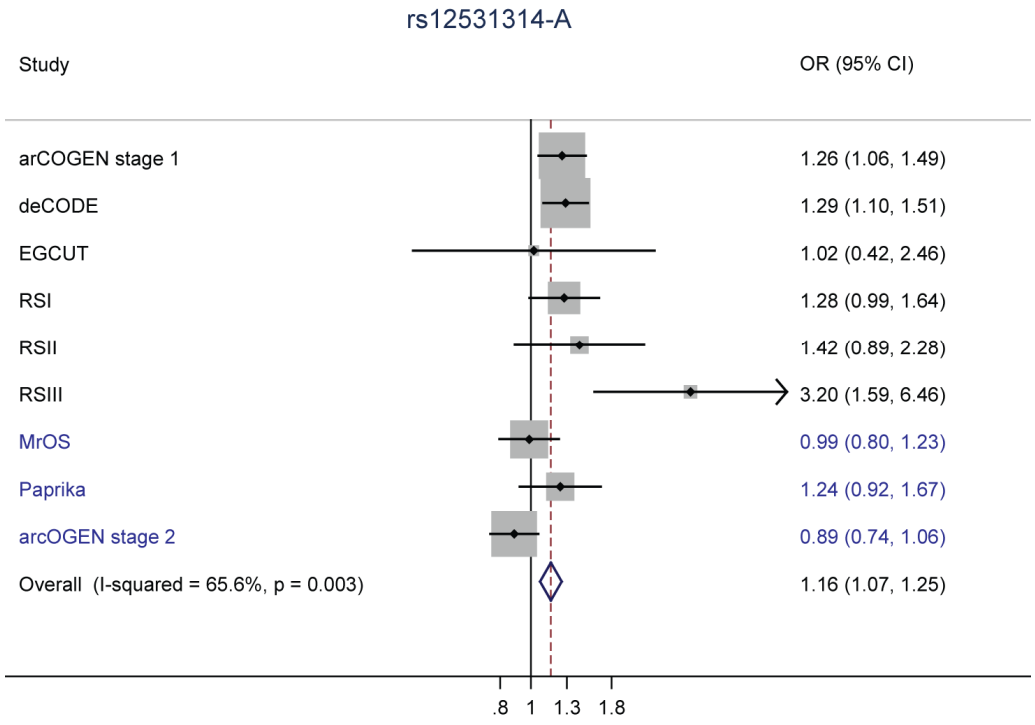
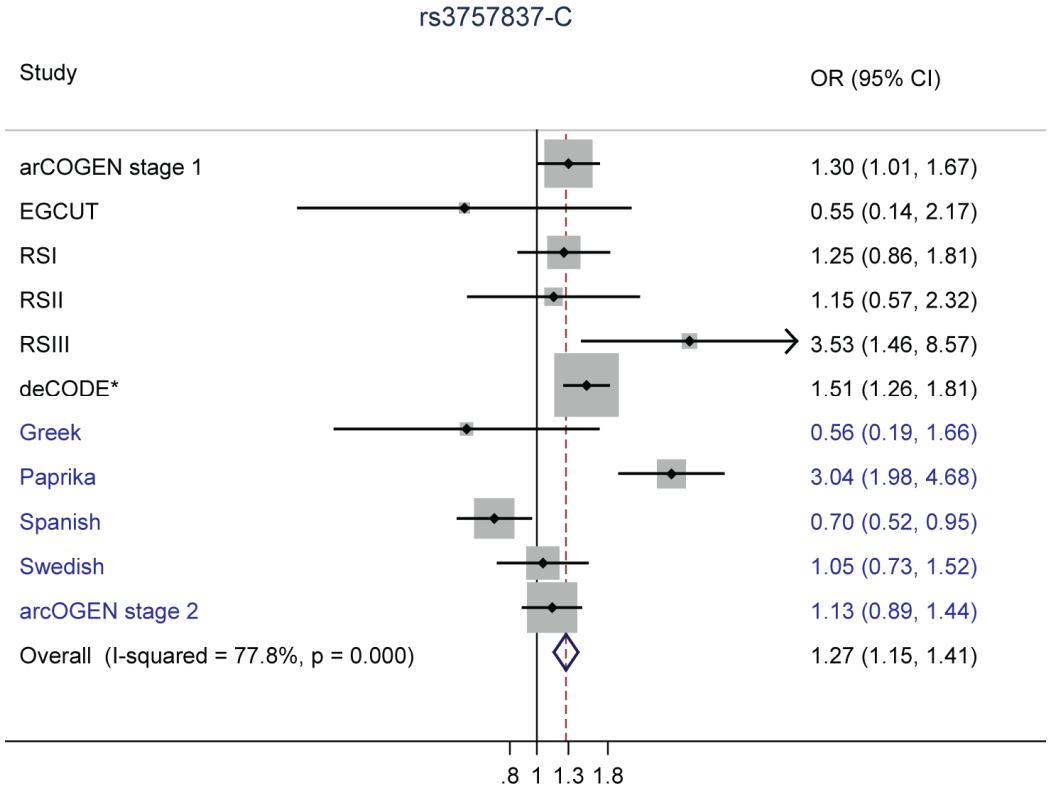


Figure S11



3. Supplementary Tables

Table S1. Summary statistics for cases and controls in the groups that were included in the discovery stage.

Study	N cases	Females (%)	Age Mean (SD)	BMI Mean (SD)	Height Mean (SD)	N controls	Females (%)	Age Mean (SD)	BMI Mean (SD)	Height Mean (SD)
Discovery										
arcOGEN	1728	64	65.8 (8.7)	28.1 (5.4)	165 (9.0)	4896	49.0	NA	NA	NA
deCODE	1423	55	69.7 (7.7)	26.8 (4.5)	169 (9.0)	31385	55	51.3 (21.7)	27.1 (5.3)	170 (9.0)
EGCUT	64	74	71.7 (13.2)	29.5 (4.6)	164.8 (9.2)	2531	56	47.4 (2.2)	25.7 (5.7)	164.3 (6.9)
GARP	106	82	60.1 (7.6)	26.8 (5.4)	168.0 (7.8)	1671	55	57.7 (1.4)	26.2 (5.5)	169.9 (9.3)
RSI	760	53	67.4 (7.7)	26.0 (3.5)	168.0 (9.3)	3233	51	66.9 (7.6)	25.8 (3.4)	168.3 (9.3)
RSII	159	52	64.0 (7.5)	27.0 (4.0)	169.1 (9.3)	1472	51	63.4 (6.9)	26.9 (4.0)	169.3 (9.3)
RSIII	41	56	55.7 (5.4)	27.3 (4.3)	171.2 (9.3)	1487	56	55.6 (5.4)	27.3 (4.3)	171.1 (9.4)
TwinsUK	68	100	56.2 (7.8)	26.0 (4.6)	161.0 (6.3)	228	100	49.0 (5.9)	24.3 (4.0)	162.4 (5.8)

Table S2. Association p-values of the prioritized SNPs before and after adjustments for age, BMI and height

SNP/Group	P Unadjusted	P Age-Adjusted	P BMI-Adjusted	P Height-Adjusted
rs6094710				
arcOGEN	0.0003398	NA	NA	NA
deCODE	0.013	0.028	0.029	0.026
EGCUT	0.226229	0.448463	0.236926	0.379193
GARP	0.42634	0.665	0.407	0.461
RSI	0.01931	0.0142	0.0157	0.01757
RSII	0.03662	0.04299	0.03607	0.02804
RSIII	0.6377	0.49	0.6325	0.499
TWINSUK	0.81883	0.9473	0.5693	0.5272
rs1577792				
arcOGEN	0.0026362	NA	NA	NA
deCODE	0.014	0.017	0.017	0.02
EGCUT	0.592781	0.437926	0.52881	0.526826
GARP	0.18416	0.474	0.223	0.254
RSI	0.001447	0.014	0.001551	0.001623
RSII	0.03931	0.07495	0.03963	0.03916
RSIII	0.8852	0.858	0.8509	0.7973
TWINSUK	0.36496	0.9349	0.332	0.2965
rs5009270				
arcOGEN	0.01702	NA	NA	NA
deCODE	0.0024	0.0023	0.015	0.014
EGCUT	0.0323308	0.0276716	0.0386288	0.0366348
GARP	0.027003	0.041	0.062	0.053
RSI	0.004403	0.002941	0.004449	0.004
RSII	0.4895	0.508	0.479	0.4734
RSIII	0.1292	0.2191	0.1215	0.3647
TWINSUK	0.63171	0.7448	0.8074	0.8516
rs10773046				
arcOGEN	0.0015225	NA	NA	NA
deCODE	0.0000194	0.0000125	0.00029	0.00036
EGCUT	0.756363	0.989485	0.77615	0.831619
GARP	0.47104	0.274	0.367	0.404
RSI	0.6551	0.5028	0.6319	0.6769
RSII	0.1994	0.3065	0.1791	0.1978
RSIII	0.4235	0.1293	0.4174	0.1927
TWINSUK	0.30965	0.4033	0.3059	0.1895
rs17610181				
arcOGEN	0.0034746	NA	NA	NA
deCODE	0.57	0.92	0.3	0.33
EGCUT	NA	0.0549912	0.146748	0.168055
GARP	0.084506	0.085	0.089	0.103
RSI	0.000087	0.0001067	0.000216	0.1851
RSII	0.03551	0.03559	0.03916	0.0302
RSIII	0.5777	0.6116	0.1849	0.1398
TWINSUK	0.91205	0.6116	0.7059	0.6352
rs10878630				

arcOGEN	0.0282	NA	NA	NA
deCODE	0.00014	0.00029	0.0000429	0.0000365
EGCUT	0.0125402	0.0722638	0.0460731	0.0821358
GARP	0.4427	0.182	0.431	0.451
RSI	0.5189	0.5768	0.4333	0.5768
RSII	0.1499	0.1465	0.1564	0.1465
RSIII	0.1882	0.1864	0.1849	0.1854
TWINSUK	0.23472	0.4759	0.2289	0.1637
rs12551314				
arcOGEN	0.0096	NA	NA	NA
deCODE	0.0028	0.0026	0.00044	0.00089
EGCUT	0.962784	0.562756	0.539139	0.659855
GARP	NA	0.705	0.271	0.259
RSI	0.06476	0.05551	0.248	0.06039
RSII	0.1507	0.1674	0.135	0.1449
RSIII	0.001947	0.0009252	0.001961	0.001217
TWINSUK	NA	NA	NA	NA
rs3757837				
arcOGEN	0.03802	NA	NA	NA
deCODE	0.0000122	0.0000234	0.0000243	0.0000381
EGCUT	0.38955	0.203123	0.104614	0.562756
GARP	NA	0.197	0.256	0.296
RSI	0.2497	0.1901	0.2607	0.242
RSII	0.7047	0.8979	0.7049	0.7067
RSIII	0.01069	0.00433	0.005202	0.005141
TWINSUK	NA	NA	NA	NA

Table S3. Association p-values of the prioritized SNPs in publicly available databases of height and BMI

		Height		BMI	
Marker	Locus	P	# individuals	P	# individuals
rs6094710	20q13	0.092	131389	0.96	123206
rs1577792	6q14	0.66	133766	0.046	123861
rs5009270	7q31	0.27	127727	0.63	119547
rs10773046	12q24	2.1E-04	133828	0.42	123863
rs17610181	17q23	0.70	132978	0.13	123864
rs10878630	12q15	0.55	133647	1.00	123718
rs12551314	9q22	0.30	133762	0.65	123866
rs3757837	7p13	0.46	116897	0.81	116638

Acknowledgements

The authors are very grateful to all study participants, the staff from all studies and the participating physicians and pharmacists. TreatOA is funded by the European Commission framework 7 program (grant 200800). The authors thank all arcOGEN participants for their contribution to this manuscript. arcOGEN (http://www.arcogen.org.uk/) was funded by a special purpose grant from Arthritis Research UK (grant 18030). This study used genotype data from population controls that was generated by the Wellcome Trust Case Control Consortium 2 (http://www.wtccc.org.uk), funded by The Wellcome Trust (grant 083948). The population controls were from the 1958 British Birth Cohort collection funded by the Medical Research Council (grant G0000934) and The Wellcome Trust (grant 068545) and from the UK Blood Services Collection of Common Controls funded by The Wellcome Trust. This research utilizes resources provided by the Type 1 Diabetes Genetics Consortium, a collaborative clinical study sponsored by the National Institute of Diabetes and Digestive and Kidney Diseases (NIDDK), National Institute of Allergy and Infectious Diseases (NIAID), National Human Genome Research Institute (NHGRI), National Institute of Child Health and Human Development (NICHD), and Juvenile Diabetes Research Foundation International (JDRF) and supported by U01 DK062418. The UK samples used in arcOGEN derive from nine centres; Nottingham, London, Oxford, Sheffield, Southampton, Newcastle, Wansbeck, Worcester and Edinburgh. For Nottingham we acknowledge Arthritis Research UK for funding the collection of the majority of extant cases (stage 1). For London we thank the staff from the TwinsUK unit and the Chingford Study for stage 1 patient ascertainment, we acknowledge financial support from Arthritis Research UK, from the Wellcome Trust, from the Department of Health via the National Institute for Health Research (NIHR) comprehensive Biomedical Research Centre (BRC) award to Guy's & St Thomas' NHS Foundation Trust in partnership with King's College London, and from the European Commission framework 7 programme TREAT-OA (grant 200800). For Oxford we acknowledge funding support from the Collisson Foundation, the Botnar Foundation and the Jean Shanks Foundation for stage 1 patient ascertainment and we acknowledge the NIHR for supporting the Biomedical Research Unit (BRU) at the University of Oxford. For Sheffield we acknowledge the NIHR for supporting the Sheffield Musculoskeletal BRU collection site, the South Yorkshire Clinical Research Network for part- funding the Sheffield research nurses and for clerical support, the Royal College of Surgeons of England and the Cavendish Hip Foundation. For Southampton we acknowledge the Wellcome Trust Clinical Research Facility at Southampton General Hospital and we thank Anna Bara, trial manager for the Arthritis Research UK-funded VIDEO study. For Edinburgh we acknowledge the contribution of Gayle Scott and Helen Watters for recruitment and ascertainment of stage 2 cases. For Newcastle we acknowledge the Freeman Hospital arthroplasty surgeons Nigel Brewster, David Deehan, Craig Gerrand, Andy Gray, Munawar Hashmi, James Holland and David Weir for assistance in stage 2 patient ascertainment, we thank Karen Bettinson for her assistance, we acknowledge the support of the UK NIHR BRC for Ageing and Age-related disease award to the Newcastle upon Tyne Hospitals NHS Foundation Trust and the support of the Northumberland, Tyne and Wear CLRN. For Wansbeck we acknowledge the Northumbria arthroplastyconsulants Mr Partington, Mr Muller, Mr Emmerson, Mr Prasad, Mr Kramer, Mr Jones, Mr Murty, Mr Asaad, Mr Leitch, Mr Pratt, Mr Townshend and Miss Van Kampen for accrual, and the research nurse Donna Gray for patient ascertainment. For Worcester we acknowledge the support of the West Midlands (south) CLRN. We acknowledge stage 1 and 2 sample management undertaken by the UK DNA Banking Network funded by the Medical Research Council at CIGMR - the Centre for Integrated Genomic Medical Research, University of Manchester. Genotyping was performed at the Wellcome Trust Sanger Institute and we thank Emma Gray, Sarah Edkins and Cordelia Langford for their assistance. We thank Vanessa Cox of the MRC Epidemiology Resource Centre, Southampton, for database and computing services for the arcOGEN stage 2 patient questionnaires. The Study of Osteoporotic Fractures (SOF) is supported by National Institutes of Health funding. The National Institute on Aging (NIA) provides support under the following grant numbers: AG05407, AR35582, AG05394, AR35584, AR35583, R01 AG005407, R01 AG027576-22, 2 R01 AG005394-22A1, R01 AR052000-01 A1, 2K24-AR04884, AR043052 and 2 R01 AG027574-22A1. The Osteoporotic Fractures in Men (MrOS) Study is supported by National Institutes of Health funding. The following institutes provide support: the National Institute of Arthritis and Musculoskeletal and Skin Diseases (NIAMS), the National Institute on Aging (NIA), the National Center for Research Resources (NCRR), and NIH Roadmap for Medical Research under the following grant numbers: U01 AR45580, U01 AR45614, U01 AR45632, U01 AR45647, U01 AR45654, U01 AR45583, U01 AG18197, U01-AG027810, R01 AR052000-01 A1, 2K24-AR04884, AR043052 and UL1 RR024140. The Leiden University Medical Centre, the Dutch Arthritis Association and Pfizer Inc., Groton, CT, USA support the GARP study, whilst genotypic work was supported by the Netherlands Organization of Scientific Research (MW 904-61-095, 911-03-016, 917 66344 and 911-03-012), Leiden University Medical Centre, and by the “Centre of Medical System Biology” and the “Netherlands Consortium of Healthy Aging” in the framework of the Netherlands Genomics Initiative (NGI). Furthermore, the research leading to these results has received funding from the Dutch Arthritis Association (DAA 2010_017) and the European Union's Seventh Framework Programme (FP7/2007-2011) under grant agreement n° 259679.TwinsUK acknowledges financial support from the Wellcome Trust, the Department of Health via the National Institute for Health Research (NIHR) comprehensive Biomedical Research Centre award to Guy's & St Thomas' NHS Foundation Trust in partnership with King's College London and Arthritis Research Campaign. The deCODE Genetics OA project

is funded in part by the European Commission framework 7 programme TREAT-OA (grant 200800). The Rotterdam Study is supported by the Netherlands Organisation of Scientific Research NWO Investments (nr. 175.010.2005.011, 911-03-012), the Research Institute for Diseases in the Elderly (014-93-015; RIDE2), the Netherlands Genomics Initiative (NGI)/Netherlands Organisation for Scientific Research (NWO) (project nr. 050-060-810) and the Erasmus Medical Center and Erasmus University, Rotterdam. EGCUT received financing by FP7 grants (201413, 245536), grant from Estonian Government SF0180142s08, from the EU through the European Regional Development Fund, OPENGENE, in the frame of Centre of Excellence in Genomics and Univ. of Tartu grant SP1GVARENG. We acknowledge EGCUT and Estonian Biocentre personnel, especially Mr V. Soo. The Study of Osteoporotic Fractures (SOF) is supported by National Institutes of Health funding. The National Institute on Aging (NIA) provides support under the following grant numbers: R01 AG005407, R01 AR35582, R01 AR35583, R01 AR35584, R01 AG005394, R01 AG027574, and R01 AG027576. The Osteoporotic Fractures in Men (MrOS) Study is supported by National Institutes of Health funding. The following institutes provide support: the National Institute of Arthritis and Musculoskeletal and Skin Diseases (NIAMS), the National Institute on Aging (NIA), the National Center for Research Resources (NCRR), and NIH Roadmap for Medical Research under the following grant numbers: U01 AR45580, U01 AR45614, U01 AR45632, U01 AR45647, U01 AR45654, U01 AR45583, U01 AG18197, U01-AG027810, and UL1 RR024140. The National Institute of Arthritis and Musculoskeletal and Skin Diseases (NIAMS) provides funding for the MrOS ancillary study 'Replication of candidate gene associations and bone strength phenotype in MrOS' under the grant number R01-AR051124. The National Institute of Arthritis and Musculoskeletal and Skin Diseases (NIAMS) provides funding for the MrOS ancillary study 'GWAS in MrOS and SOF' under the grant number RC2ARO58973. Spanish and Greek sample genotyping was done with funding from grant 11/01048 and by RETICS Program, RD08/0075 (RIER) of the Instituto de Salud Carlos III (Spain) that are partially financed by the European Regional Development Fund of the European Union. AGD-W, LS, KP and EZ are supported by the Wellcome Trust (grant 098051)

References

1. Marchini J, Howie B, Myers S, McVean G, Donnelly P (2007) A new multipoint method for genome-wide association studies by imputation of genotypes. Nat Genet 39: 906-913.
2. Kavvoura FK, Ioannidis JP (2008) Methods for meta-analysis in genetic association studies: a review of their potential and pitfalls. Hum Genet 123: 1-14.
3. Higgins JP, Thompson SG, Deeks JJ, Altman DG (2003) Measuring inconsistency in meta-analyses. BMJ 327: 557-560.
4. Han B, Eskin E (2011) Random-effects model aimed at discovering associations in meta-analysis of genome-wide association studies. Am J Hum Genet 88: 586-598.
5. Hofman A, Breteler MM, van Duijn CM, Krestin GP, Pols HA, et al. (2007) The Rotterdam Study: objectives and design update. Eur J Epidemiol 22: 819-829.
6. Stefansson SE, Jonsson H, Ingvarsson T, Manolescu I, Jonsson HH, et al. (2003) Genomewide scan for hand osteoarthritis: a novel mutation in matrilin-3. Am J Hum Genet 72: 1448-1459.
7. Spector TD, Williams FM (2006) The UK Adult Twin Registry (TwinsUK). Twin Res Hum Genet 9: 899-906.
8. Riyazi N, Meulenbelt I, Kroon HM, Ronday KH, Hellio le Graverand MP, et al. (2005) Evidence for familial aggregation of hand, hip, and spine but not knee osteoarthritis in siblings with multiple joint involvement: the GARP study. Ann Rheum Dis 64: 438-443.
9. Heijmans BT, Beekman M, Houwing-Duistermaat JJ, Cobain MR, Powell J, et al. (2006) Lipoprotein particle profiles mark familial and sporadic human longevity. PLoS Med 3: e495.
10. Panoutsopoulou K, Southam L, Elliott KS, Wrayner N, Zhai G, et al. (2011) Insights into the genetic architecture of osteoarthritis from stage 1 of the arcOGEN study. Ann Rheum Dis 70: 864-867.
11. arcOGEN Consortium and arcOGEN Collaborators (2012) Identification of new susceptibility loci for osteoarthritis (arcOGEN): a genome-wide association study. Lancet 380: 815-823.
12. Fytili P, Giannatou E, Papanikolaou V, Stripeli F, Karachalios T, et al. (2005) Association of repeat polymorphisms in the estrogen receptors alpha, beta, and androgen receptor genes with knee osteoarthritis. Clin Genet 68: 268-277.

13. Rodriguez-Lopez J, Pombo-Suarez M, Liz M, Gomez-Reino JJ, Gonzalez A (2006) Lack of association of a variable number of aspartic acid residues in the asporin gene with osteoarthritis susceptibility: case-control studies in Spanish Caucasians. Arthritis Res Ther 8: R55.
14. Blank JB, Cawthon PM, Carrion-Petersen ML, Harper L, Johnson JP, et al. (2005) Overview of recruitment for the osteoporotic fractures in men study (MrOS). Contemp Clin Trials 26: 557-568.
15. Orwoll E, Blank JB, Barrett-Connor E, Cauley J, Cummings S, et al. (2005) Design and baseline characteristics of the osteoporotic fractures in men (MrOS) study--a large observational study of the determinants of fracture in older men. Contemp Clin Trials 26: 569-585.
16. Cummings SR, Nevitt MC, Browner WS, Stone K, Fox KM, et al. (1995) Risk factors for hip fracture in white women. Study of Osteoporotic Fractures Research Group. N Engl J Med 332: 767-773.
17. Keurentjes JC, Blane D, Bartley M, Keurentjes JJ, Fiocco M, et al. (2013) Socio-economic position has no effect on improvement in health-related quality of life and patient satisfaction in total hip and knee replacement: a cohort study. PLoS One 8: e56785.
18. Keurentjes JC, Fiocco M, So-Osman C, Onstenk R, Koopman-Van Gemert AWMM et al. (2013) Patients with severe radiographic osteoarthritis have a better prognosis in psysical functioning after hip and knee replacement; A cohort-study. Plos One 8: e59500
19. Keurentjes JC, Fiocco M, Nelissen RG (2013) Willingness to undergo surgery again validated clinically important differences in health-related quality of life after total hip replacement or total knee replacement surgery. J Clin Epidemiol. pii: S0895-4356(13)00171-6. doi: 10.1016/j.jclinepi.2013.04.010.
20. Richards JB, Rivadeneira F, Inouye M, Pastinen TM, Soranzo N, et al. (2008) Bone mineral density, osteoporosis, and osteoporotic fractures: a genome-wide association study. Lancet 371: 1505-1512.
21. Ju-Hyun P, Wacholder S, Gail MH, Peters U, Jacobs KB et al. (2010) Estimation of effect size distributions from genome-wide association studies and implications for furture discoveries. Nat Genet 42:570-575