Published in final edited form as:
Ann Rheum Dis. 2011 February ; 70(2): 349–355. doi:10.1136/ard.2010.132787.

Meta-analysis of genome-wide association studies confirms a susceptibility locus for knee osteoarthritis on chromosome 7q22

Evangelos Evangelou[1,*], **Ana M. Valdes**[2,*], **Hanneke J.M Kerkhof**[3,4,*], **Unnur Styrkarsdottir**[5,*], **YanYan Zhu**[6,*], **Ingrid Meulenbelt**[4,7,*], **Rik J. Lories**[8], **Fotini B. Karassa**[1], **Przemko Tylzanowski**[8], **Steffan D. Bos**[4,7], **arcOGEN consortium**, **Toru Akune**[9], **Nigel K. Arden**[10,11], **Andrew Carr**[12], **Kay Chapman**[12,13], **L. Adrienne Cupples**[6], **Jin Dai**[14], **Panos Deloukas**[15], **Michael Doherty**[16], **Sally Doherty**[16], **Gunnar Engstrom**[17], **Antonio Gonzalez**[18], **Bjarni V. Halldorsson**[5,19], **Christina L. Hammond**[20], **Deborah J. Hart**[2], **Hafdis Helgadottir**[5], **Albert Hofman**[21], **Shiro Ikegawa**[22], **Thorvaldur Ingvarsson**[23], **Qing Jiang**[14], **Helgi Jonsson**[24,25], **Jaakko Kaprio**[26,27,28], **Hiroshi Kawaguchi**[29], **Kalle Kisand**[30], **Margreet Kloppenburg**[31,32], **Urho M. Kujala**[33,34], **L. Stefan Lohmander**[35], **John Loughlin**[36], **Frank P. Luyten**[8], **Akihiko Mabuchi**[37], **Andrew McCaskie**[36,38], **Masahiro Nakajima**[22], **Peter M. Nilsson**[17], **Nao Nishida**[37], **William E.R. Ollier**[39], **Kalliope Panoutsopoulou**[40], **Tom van de Putte**[41], **Stuart H. Ralston**[42], **Fernando Rivadeneira**[3,21], **Janna Saarela**[26], **Stefan Schulte-Merker**[20], **P. Eline Slagboom**[4,7], **Akihiro Sudo**[43], **Agu Tamm**[44], **Ann Tamm**[45], **Gudmar Thorleifsson**[5], **Unnur Thorsteinsdottir**[5,25], **Aspasia Tsezou**[46], **Gillian A. Wallis**[47], **J. Mark Wilkinson**[48,49], **Noriko Yoshimura**[50], **Eleftheria Zeggini**[40,51], **Guangju Zhai**[2], **Feng Zhang**[2], **Ingileif Jonsdottir**[5,25], **Andre G. Uitterlinden**[3,4,21,‡], **David T Felson**[52,‡], **Joyce B. van Meurs**[3,4,‡], **Kari Stefansson**[5,25,‡], **John P.A. Ioannidis**[1,53,54,55,‡], and **Timothy D. Spector**[2,‡]

[1]Department of Hygiene and Epidemiology, University of Ioannina School of Medicine, Ioannina, Greece [2]Department of Twin Research and Genetic Epidemiology, St. Thomas' Hospital, King's College London, London, UK [3]Department of Internal Medicine, Erasmus Medical Center, Rotterdam, The Netherlands [4]The Netherlands Genomics Initiative-Sponsored Netherlands Consortium for Healthy Aging, Rotterdam, The Netherlands [5]deCODE Genetics, Reykjavik, Iceland [6]Department of Biostatistics, Boston University School of Public Health, Boston, MA, USA [7]Department of Molecular Epidemiology, Leiden University Medical Centre, Leiden, The Netherlands [8]Laboratory for Skeletal Development and Joint Disorders, Department of Musculoskeletal Sciences, Division of Rheumatology, Katholieke Universiteit Leuven, Leuven, Belgium [9]Department of Clinical Motor System Medicine, 22nd Century Medical and Research Center, Graduate School of Medicine, The University of Tokyo, Tokyo, Japan [10]MRC Epidemiology Resource Centre, Southampton, UK [11]Musculoskeletal Biomedical Research Unit, University of Oxford, Oxford, UK [12]Botnar Research Centre, University of Oxford,Nuffield Orthopaedic Centre, Oxford, UK [13]Wellcome Trust Centre for Human Genetics, University of Oxford, Oxford, UK [14]The Center of Diagnosis and Treatment for Joint Disease, Drum Tower Hospital, Nanjing University Medical School, Nanjing, China [15]Wellcome Trust Sanger Institute, UK [16]University of Nottingham, Academic Rheumatology, City Hospital, Nottingham, UK [17]Department of Clinical Sciences Malmo, Lund University, Sweden [18]Laboratorio Investigacion 10 and Rheumatology Service. Instituto Investigacion Sanitaria-Hospital Clinico Universitario de Santiago. Santiago de Compostela, Spain [19]Reykjavik University, Reykjavic, Iceland [20]Hubrecht Institute – KNAW & UMC, Utrecht, The Netherlands [21]Department of Epidemiology, Erasmus Medical Center, Rotterdam, The Netherlands [22]Laboratory for Bone and Joint Diseases, Center

Correspondence to jioannid@cc.uoi.gr; tim.spector@kcl.ac.uk.
*equal contribution
‡equal contribution

Europe PMC Funders Author Manuscripts

for Genomic Medicine, RIKEN, Tokyo, Japan [23]FSA University Hospital, Institution of Health Science, University of Akureyri, Akureyri, Iceland [24]Department of Medicine, Landspitali University Hospital, Reykjavik, Iceland [25]Faculty of Medicine, University of Iceland, Reykjavik, Iceland [26]Institute for Molecular Medicine Finland (FIMM), University of Helsinki, Helsinki, Finland. [27]Department of Public Health, University of Helsinki, Helsinki, Finland [28]National Institute for Health and Welfare, Helsinki, Finland [29]Department of Orthopaedic Surgery, Sensory and Motor System Medicine, Graduate School of Medicine, The University of Tokyo, Tokyo, Japan [30]Immunology Group, Institute of General and Molecular Pathology, University of Tartu, Tartu, Estonia [31]Department of Rheumatology, Leiden University Medical Centre, Leiden, The Netherlands [32]Department of Clinical Epidemiology, Leiden University Medical Centre, Leiden, The Netherlands [33]Department of Health Sciences, University of Jyvaskyla, Jyvaskyla, Finland [34]ORTON Orthopaedic Hospital, ORTON Foundation, Helsinki, Finland [35]Department of Clinical Sciences Lund, Orthopedics, Lund University, Lund, Sweden [36]Institute of Cellular Medicine, Musculoskeletal Research Group, The Medical School, Newcastle University, Newcastle Upon Tyne, UK [37]Department of Human Genetics, International Health, Graduate School of Medicine, The University of Tokyo, Tokyo, Japan [38]The Newcastle Upon Tyne Hospitals NHS Trust, The Freeman Hospital, Newcastle, UK [39]Centre for Integrated Genomic Medical Research, The University of Manchester, Manchester, UK [40]Wellcome Trust Sanger Institute, Hinxton, UK [41]Tigenix, Leuven, Belgium [42]Institute of Genetics and Molecular Medicine, Western General Hospital, University of Edinburgh, Edinburgh, UK [43]Department of Orthopaedic Surgery, Faculty of Medicine, Mie University, Mie, Japan [44]Department of Internal Medicine, University of Tartu, Tartu, Estonia [45]Department of Sports Medicine and Rehabilitation, University of Tartu, Tartu, Estonia [46]Department of Biology, University of Thessaly Medical School, Larissa, Greece [47]Wellcome Trust Centre for Cell-Matrix Research, School of Translational Medicine, University of Manchester, Manchester, UK [48]Academic Unit of Bone Metabolism, Northern General Hospital, University of Sheffield, Sheffield, UK [49]NIHR Bone Biomedical Research Unit, Sheffield Teaching Hospitals NHS Foundation Trust, Sheffield, UK [50]Department of Joint Disease Research, 22nd Century Medical and Research Center, Graduate School of Medicine, The University of Tokyo, Tokyo, Japan [51]Wellcome Trust Centre for Human Genetics, University of Oxford, Oxford, UK [52]Clinical Epidemiology Unit, Boston University School of Medicine, Boston, MA, USA [53]Center for Genetic Epidemiology and Modelling, Institute for Clinical Research and Health Policy Studies, Tufts Medical Center, Tufts University School of Medicine, Boston, MA, USA [54]Biomedical Research Institute, Foundation for Research and Development-Hellas, Ioannina, Greece [55]Department of Epidemiology, Harvard School of Public Health, Boston, MA, USA

Abstract

Osteoarthritis (OA) is the most prevalent form of arthritis and accounts for substantial morbidity and disability, particularly in the elderly. It is characterized by changes in joint structure including degeneration of the articular cartilage and its etiology is multifactorial with a strong postulated genetic component. We performed a meta-analysis of four genome-wide association (GWA) studies of 2,371 knee OA cases and 35,909 controls in Caucasian populations. Replication of the top hits was attempted with data from additional ten replication datasets. With a cumulative sample size of 6,709 cases and 44,439 controls, we identified one genome-wide significant locus on chromosome 7q22 for knee OA (rs4730250, p-value=9.2×10^{-9}), thereby confirming its role as a susceptibility locus for OA. The associated signal is located within a large (500kb) linkage disequilibrium (LD) block that contains six genes; *PRKAR2B* (protein kinase, cAMP-dependent, regulatory, type II, beta), *HPB1* (HMG-box transcription factor 1), *COG5* (component of oligomeric golgi complex 5), *GPR22 (*G protein-coupled receptor 22), *DUS4L* (dihydrouridine synthase 4-like), and *BCAP29 (*the B-cell receptor-associated protein 29). Gene expression analyses of the (six) genes in primary cells derived from different joint tissues confirmed expression of all the genes in the joint environment.

Introduction

Osteoarthritis (OA) is the most prevalent form of chronic joint disease and accounts for substantial morbidity and disability, particularly among the elderly. It is characterized by loss of joint homeostasis. The articular cartilage cannot maintain its integrity and is progressively damaged, the subchondral bone envelope is thickened changing loads in the bone-cartilage biomechanical unit, the synovium shows signs of inflammation and bony spurs (osteophytes) appear at the edges of the bone. Its etiology is multifactorial with a significant genetic component as shown by twin and family studies[1, 2].

Many genetic variants have been considered as potential risk factors for OA but most of the reported associations are inconclusive or not replicated. Recently, a large-scale meta-analyses found evidence that the *GDF5* locus on chromosome 20 was associated with the increased risk of knee OA in Caucasians[3-6]. Other genome-wide data have reported an association with the *DVWA* gene in Asians but not Caucasians [7] and a *PTGS2* variant that replicated but did not reach genome-wide significance (GWS)[8]. Recently, a genome wide association (GWA) study identified a locus on chromosome 7q22 that shows an association with combined knee OA and/or hand OA phenotype[9].

In this study we have synthesized available data from four GWA studies under the auspices of the TreatOA (*Translational Research in Europe Applied Technologies for Osteoarthritis* [www.treatoa.eu]) consortium. A total of 2,371 knee OA cases were available for this first stage of the analysis. The most significant signals were further investigated in additional samples of European descent and SNPs that reached genome-wide significance (GWS) were further evaluated in Asian samples. .

Methods

Study design

A detailed description of all samples used in this study is provided in the Supplementary Material. We used a three-stage design for the identification of any potential associations between sequence variants and knee OA in populations of European ancestry. We first synthesized the available data from 4 GWA studies (deCODE, Rotterdam Study, Framingham, Twins UK) using inverse variance fixed effects models. The variants that reached the 2×10^{-5} level of significance were selected for further replication. These SNPs were followed-up in 8 additional European cohorts (arcOGEN, Greek, Spanish, Finnish, Nottingham, Chingford study, GARP, Estonian and Swedish). The SNPs that replicated in the follow-up samples were genotyped in additional 2 European samples (deCODE (Icelandic) and Swedish). One cohort provided *in silico* replication from an ongoing GWA study (arcOGEN, 12 SNPs were directly genotyped, and 6 were imputed), whereas *de novo* replication was performed in the other cohorts. Furthermore, the top hits were followed-up in Asian populations (Chinese and Japanese samples). The effect sizes from the meta-analysis of the GWA studies and the effect sizes from the replication effort were all combined to provide an overall estimate. We also synthesized the effect estimates of the European and Asian samples to provide a global summary effect estimate

Phenotype definitions

Study subjects with a radiographic Kellgren and Lawrence (K/L) grade≥ 2 [10] or total knee replacement (TKR) were included as cases in the analysis. When clinical criteria were considered (Greek, Spanish and GARP study groups) the American College of Rheumatology (ACR) classification criteria were used [11]. As controls, we considered

subjects who had no known affected joints among those assessed. For example in a cohort that assesses knee, hip and hand OA, controls were participants with no affected hip or hand joints for the knee OA analysis. Population-based controls were used for the arcOGEN study.

Genotyping and imputation

Samples from the GWA studies were genotyped using the Infinium HumanHap300 (Illumina) for deCODE and Twins UK samples, HumanHap550v3 Genotyping BeadCHip (Illumina) for the Rotterdam Study and the Affymetrix GeneChip® Human Mapping 500K for the Framingham cohort. The number of SNPs genotyped ranged from 314,075 to 500,510. Imputations were performed to increase the coverage. All the top SNPs studied had acceptable imputation quality. Finally, the genotyped and imputed SNPs that successfully passed the quality control criteria (n=2,335,627) were considered for the analyses. Detailed information on genotyping platform, quality control and imputation methods for each cohort are presented in Supplementary Table 3.

The replication samples for the Greek, Spanish, Finnish, Chingford and GARP studies were genotyped using the MassArray iPlex Gold from Sequenom. Replication genotyping was carried out by a genotyping contractor (Kbiosciences Ltd) using a competitive allele-specific PCR SNP genotyping system for the Nottingham and the Estonian cohort. The additional 622 Icelandic cases and the samples from the Swedish cohort were genotyped by deCODE genetics using the Centaurus (Nanogen) platform [12]. Detailed information on genotyping is provided in Supplement.

Statistical analysis

Association analysis—Each team performed an association testing per gender for knee OA under a per-allele model. The lambda inflation factor was calculated per gender-specific effect size using the genomic control method [13] and the standard errors were corrected by the square root of the lambda inflation factor ($SE_{\text{corrected}}=SE_{\text{observed}}x\sqrt{\lambda}$). Robust standard errors were estimated to adjust for the family relationships (Framingham study and GARP).

Meta-analysis—The effect size for each SNP (odds ratio per copy of minor allele as per HapMap) was calculated using inverse-variance fixed effects models [14], synthesizing all the sex-specific effect sizes and the corrected standard errors. We also performed analyses combining men and women. In family studies results from men and women combined were used in order to account for relatedness between females and males within families. Meta-analyses of the GWA studies were performed using the METAL (www.sph.umich.edu/csq/abecasis/metal) software. Between-study heterogeneity was tested using the Cochran's Q statistic, which is considered significant at $p<0.1$. The extent of inconsistency across studies was quantified using the I^2 metric which ranges from 0 to 100% [15]. Heterogeneity is considered low, moderate, high and very high for 0-24%, 25-49%, 50-74% and >75% respectively [16]. We also computed the 95% CI for the I^2 [17]. Furthermore, we repeated the calculation with random effects models for all SNPs that were further evaluated in replication datasets. Results are shown in Supplementary Table 4. Meta-analyses of the 18 top-hits were performed using Stata version 10.1.

Assessment of credibility—In order to assess the credibility of the top hit we calculated the Bayes factor under a spike and smear prior using as an alternative an average genetic effect corresponding to an OR of 1.2 and a conservative agnostic prior of 0.0001% [18].

Functional analysis

Two methodological approaches were used to investigate the functional role of genes identified by GWA studies. (A) By assessing their expression in primary human joint cells (synovial fibroblasts, chondrocytes and meniscal cells) and its change in response to the proinflammatory cytokines TNFα and IL1β as well as comparing their gene expression profiles during chondrocyte de-differentiation (3D pellet cultures of vs monolayer culture, see Supplement) and (B) by assessing their expression dynamics by wholemount in situ hybridization using 6h (shield), 10h (bud), 13h (5-9 somites) and 1, 2, 3 and 4 days old zebrafish (*Danio rerio*) embryos, in order to explore their role during embryogenesis (see Supplement).

Results

Meta-analysis of GWA studies and replication of top findings

The descriptive characteristics of the GWA studies used for the meta-analyses are from Iceland (deCODE), the Netherlands (Rotterdam study), USA (Framingham) and the UK (Twins UK). The characteristics of these studies are presented in Table 1 and Supplement. The 4 GWA datasets included a total of 2,371 cases and 35,909 controls. A quantile-quantile (QQ) plot, comparing the meta-analysis association results of the four studies to those expected by chance, showed an excess of SNP associations indicating a likely true association signal (Figure 1). The data analysis revealed the strongest association on chromosome 7q22 with a p-value of 5.06×10^{-8} for rs4730250 localized in dihydrouridine synthase 4-like gene (*DUS4L*) (Figure 2). Other associated signals in 7q22 gene cluster are in high linkage disequilibrium (LD) ($r^2>0.8$) with the top signal (Figure 2).

We selected for follow-up in replication samples all SNPs with a p-value $<2\times10^{-5}$ in the meta-analysis association results. A total of 18 SNPs from 10 chromosomal loci satisfied this criterion (Supplementary Table 1). However, as some of those SNPs are fully equivalent in the HapMap-CEU dataset a total of 11 non-identical SNPs were tested for replication. We analyzed these 11 SNPs for replication in 3,326 cases and 7691 controls from 8 European studies (Table 1 and Supplement). Two SNPs, rs4730250 and rs10953541 both located at 7q22, replicated nominally ($p<0.05$) in the combined analysis of the follow-up samples, with p-values of 6.3×10^{-4} and 8.3×10^{-3} respectively. The two SNPs, rs4730250 and rs10953541, were then further genotyped in two additional replication sets.

Both SNPs reached GWS in a meta-analysis of all European sample sets (the GWA datasets and the replication cohorts) (Table 2). We analyzed a total of 6,709 knee OA cases and 44,439 controls. SNP rs4730250 was genome wide significant GWS with a per-allele summary OR of 1.17 (95% CI: 1.11-1.24) and a p-value of 9.2×10^{-9}. The minor allele frequency was 0.17 in the combined dataset. Low heterogeneity was observed ($I^2=15\%$, 95% CI: 0-48%), which was not statistically significant (p=0.26 for Cochran's Q statistic) (Figure 3). No gender specific effects were seen. The summary estimates did not differ significantly in men and women (p-value=0.74, test of homogeneity) (Figure 3). Analysis where both sexes were analyzed together in all cohorts did not alter the results (OR=1.17 [95% CI: 1.07-1.27], p-value=4.1×10^{-8}). The summary effect sizes of all loci under study are presented in Table 2. The two significant SNPs at 7q22, rs4730250 and rs10953541, are highly correlated ($D'=1$, $r^2=0.63$ in HapMap-CEU) and are likely to represent the same underlying association signal as shown by conditional association analysis (Supplementary Table 2).

Age and BMI are considered to be significant risk factors for the development of knee OA [19-25]. We performed an analysis where the top hit was adjusted for these risk factors in

deCODE samples and the Rotterdam Study. The association of the top hit remained largely unchanged in analyses adjusted for BMI and age.

In order to assess the credibility of the associations of the two SNPs, we calculated the Bayes factor[18] under a spike and smear prior using an average genetic effect corresponding to an OR of 1.2 and a conservative agnostic prior (assuming no prior knowledge of the association) of 0.0001%. The posterior credibility of these associations was 98% and remained similarly high even with a small alternative effect size of 1.1.

We also tested if the observed signal at the 7q22 region was replicated in East Asian samples (Japanese and Chinese cohorts). The total number of knee OA cases and controls assessed was 1183 and 1245 respectively. The rs12535761 was used as a proxy for the rs4730250. The two SNPs are in strong LD (r^2=1, D′=1, in HapMap Asian samples). The finding was not replicated in the Asian samples with a summary effect size of 1.03 (95% CI: 0.85-1.25). A meta-analysis including both European and Asian samples with 7,892 cases and 45,684 controls yielded a global summary effect of 1.15 (95% CI: 1.10-1.22) with p-value=5.7×10^{-8} for rs4730250 with low heterogeneity (I^2=19%).

Functional analysis of genes in 7q22 cluster

The associated signal at 7q22 is located within a large (500kb) linkage disequilibrium (LD) block that contains six genes; *PRKAR2B* (protein kinase, cAMP-dependent, regulatory, type II, beta), *HPB1* (HMG-box transcription factor 1), *COG5* (component of oligomeric golgi complex 5), *GPR22* (G protein-coupled receptor 22), *DUS4L* (dihydrouridine synthase 4-like), and *BCAP29 (*the B-cell receptor-associated protein 29).

We performed additional experiments to get more information about the genes in the cluster and their potential role in joint biology and pathology. Analysis of the mRNA expression data in chondrocyte pellet indicates that BCAP29, COG5, DUS4L and HBP1 expression levels were higher than in monolayer cultures suggesting that they are expressed in an environment that more accurately recapitulates articular cartilage. In contrast no difference were seen for GPR22 and PRKAR2B mRNA expression. In a zebrafish model the expression of all genes was detectable from the shield stage onwards. Results are described in detail in Supplement

Discussion

This study provides further evidence for a knee OA signal localizing to the 7q22 cluster region and associated with knee OA. The statistical credibility and confidence of this evidence is very high based on the calculations of the Bayes factor. The same locus has been identified and proposed as an OA susceptibility locus from the Rotterdam Study for the prevalence and progression of OA [9]. Our study and the earlier Rotterdam Study do include overlapping populations. However, our study was specifically targeting the knee OA phenotype. Furthermore, an additional 3 European cohorts and two Asian populations were used for further replication. Our study utilizes the largest sample size in the genetics of knee OA research to date with almost 8000 knee OA cases analyzed.

The most significant hits identified by our study are located within a large (500kb) linkage disequilibrium (LD) block that contains six genes; *PRKAR2B*, *HPB1*, *COG5, GPR22*, *DUS4L* and *BCAP29*. The top hit rs4730250 is annotated in intron 3 of the *DUS4L* gene. Any of the genes at the 7q22 region may confer risk for knee OA, as the LD pattern across the region is high.

Europe PMC Funders Author Manuscripts

The functional analyses support the epidemiological findings but do not exclude any of the 6 candidate genes. Specifically, the zebrafish experiments show that both *COG5* and *DUS4L* are expressed in developing cartilage supporting the notion that either of these genes could have a biological function during chondrogenesis. The studies in the de-differentiation model of human chondrocytes (3D vs 2D culture) show that *BCAP29*, *COG5*, *DUS4L* and *HBP1* all have different expression patterns in 3D culture (chondro-like cells) than in 2D culture (de-differentiated cells) suggesting that these four genes may play a role in cartilage metabolism.

A major issue in the field of osteoarthritis is the definition of the disease phenotypes[4, 26]. Different criteria may introduce bias and dilute the effect. The cases in our study were defined either clinically by the presence of a knee replacement or radiographically using the Kellgren/Lawrence (K/L) system. The K/L system is however far from perfect and can be affected by differences in the position of the knee in which the radiographs were obtained, observer biases, interpretation of grading criteria and random error [27, 28]. Similarly there are no standard criteria for replacing knee joints. This may introduce heterogeneity and move the observed effects towards the unity, and so under-estimate the true strength of an association. In our study we synthesized data with a standardized definition of the phenotype, however small individual locus effects with ORs in the range of 1.1-1.2 as for other chronic diseases may well be plausible for knee OA, explaining the paucity of other significant hits despite the reasonable large-scale effort. These findings highlight that even larger collaborative studies and improved standardization of the phenotypes are needed to better understand and identify further genetic variants of OA.

Moreover, even though we were able to accumulate a large sample size, the power of the study to detect very small effect sizes in the range of 1.05-1.15 is inadequate. For example, identification of a GWS signal with an effect size of 1.15 and minor allele frequency of 20%, with 80% power would require almost 7000 additional knee OA cases.

Our results confirm that the 7q22 chromosomal region confers risk for knee OA, which along with our functional work implicates 6 possible genes. Further in depth genetic analysis of the locus, including deep-sequencing of the region and functional work including *in vitro* assays and animal models will be required to deepen our understanding of the underlying molecular pathways associated with the disease.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

We thank all arcOGEN participants for their contribution in this manuscript. arcOGEN is funded by a special purpose grant from the Arthritis Research Campaign (**arc**, 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. The samples used in arcOGEN derive from five centres in the UK: Nottingham, London, Oxford, Sheffield and Southampton. For Nottingham we acknowledge **arc** for funding the collection of the majority of cases. For London we thank the staff from the TwinsUK unit and the Chingford Study for patient ascertainment, we acknowledge financial support from **arc**, from the Wellcome Trust and 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. For Oxford we acknowledge funding support from the Collisson Foundation, the Botnar Foundation and the Jean Shanks Foundation for patient ascertainment, we acknowledge the NIHR for supporting the Biomedical Research Unit (BRU) at the University of Oxford, and we thank Bridget Watkins and Kim Clipsham for assistance in patient ascertainment. For Sheffield we

acknowledge the NIHR for supporting the Sheffield Bone BRU, the South Yorkshire Clinical Research Network for part-funding the Sheffield research nurse and for clerical support, the Royal College of Surgeons of England and the Cavendish Foundation. For Southampton we acknowledge the Wellcome Trust Clinical Research Facility at Southampton General Hospital and we thank Phillippa-Kate Battley and Elizabeth Arden for assistance with patient ascertainment, and Richard Keen and Anna Bara, principal investigator and trial manager for the **arc**-funded VIDEO study, respectively. 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 (JL and AMcC). We acknowledge 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 and we thank Kate Dixon, Kate Sherburn and Debbie Payne for their assistance. Genotyping was performed at the Wellcome Trust Sanger Institute and we thank Emma Gray, Sarah Edkins, Rhian Gwilliam, Suzannah Bumpstead and Cordelia Langford for their assistance. Analysis of the arcOGEN data was performed at the Wellcome Trust Centre for Human Genetics and at the Wellcome Trust Sanger Institute and we acknowledge the work of the arcOGEN analysis team members Nigel W. Rayner, Lorraine Southam, Guangju Zhai, Katherine S Elliott, Sarah E. Hunt, Hannah Blackburn, Simon C. Potter, Aaron Garth Day-Williams and Claude Beazley. EZ is supported by the Wellcome Trust (WT088885/Z/09/Z), LS is supported by the European Community Framework 7 large collaborative project grant TREAT-OA, KC is supported by a Botnar Fellowship and by the Wellcome Trust (WT079557MA), NWR is supported by the Wellcome Trust (WT079557MA), JMW is supported by the Higher Education Funding Council for England.

Rik Lories is the recipient of a postdoctoral fellowship from the Flanders Research Foundation (FWO Vlaanderen).

ROAD (TA, HK, AM, NN, NY) acknowledge Katsushi Tokunaga, Shigeyuki Muraki, Hiroyuki Oka and Kozo Nakamura for scientific advice and data collection. We acknowledge funding support by Grants-in-Aid for Scientific Research (S19109007, B21390417) from the Japanese Ministry of Education, Culture, Sports, Science and Technology, H17-Men-eki-009 from the Ministry of Health, Labor and Welfare, and JOA-Subsidized Science Project Research 2006-1 from the Japanese Orthopaedic Association.

Grant Supporter: European Commission framework 7 programme grant 200800 TREAT-OA, NWO Investments (175.010.2005.011)

References

1. Zhai G, Hart DJ, Kato BS, MacGregor A, Spector TD. Genetic influence on the progression of radiographic knee osteoarthritis: a longitudinal twin study. Osteoarthritis Cartilage. Feb; 2007 15(2):222–225. [PubMed: 17045816]
2. Valdes AM, Spector TD. The contribution of genes to osteoarthritis. Rheum Dis Clin North Am. Aug; 2008 34(3):581–603. [PubMed: 18687274]
3. Chapman K, Takahashi A, Meulenbelt I, Watson C, Rodriguez-Lopez J, Egli R, et al. A meta-analysis of European and Asian cohorts reveals a global role of a functional SNP in the 5′ UTR of GDF5 with osteoarthritis susceptibility. Hum Mol Genet. May 15; 2008 17(10):1497–1504. [PubMed: 18299287]
4. Evangelou E, Chapman K, Meulenbelt I, Karassa FB, Loughlin J, Carr A, et al. Large-scale analysis of association between GDF5 and FRZB variants and osteoarthritis of the hip, knee, and hand. Arthritis Rheum. Jun; 2009 60(6):1710–1721. [PubMed: 19479880]
5. Valdes AM, Spector TD, Doherty S, Wheeler M, Hart DJ, Doherty M. Association of the DVWA and GDF5 polymorphisms with osteoarthritis in UK populations. Ann Rheum Dis. Dec 3.2008
6. Vaes RB, Rivadeneira F, Kerkhof JM, Hofman A, Pols HA, Uitterlinden AG, et al. Genetic variation in the GDF5 region is associated with osteoarthritis, height, hip axis length and fracture risk: the Rotterdam study. Ann Rheum Dis. Nov 24.2008
7. Meulenbelt I, Chapman K, Dieguez-Gonzalez R, Shi D, Tsezou A, Dai J, et al. Large replication study and meta-analyses of DVWA as an osteoarthritis susceptibility locus in European and Asian populations. Hum Mol Genet. Apr 15; 2009 18(8):1518–1523. [PubMed: 19181678]
8. Valdes AM, Loughlin J, Timms KM, van Meurs JJ, Southam L, Wilson SG, et al. Genome-wide association scan identifies a prostaglandin-endoperoxide synthase 2 variant involved in risk of knee osteoarthritis. Am J Hum Genet. Jun; 2008 82(6):1231–1240. [PubMed: 18471798]
9. Kerkhof HJLR, Meulenbelt I, Jonsdottir I, Valdes AM, Arp P, Ingvarsson T, et al. A Genome-Wide association study identifies a locus on chromosome 7q22 to influence susceptibility for osteoarthritis. Arthritis Rheum. 2010

10. Kellgren JH, Lawrence JS. Radiological assessment of osteo-arthrosis. Ann Rheum Dis. Dec; 1957 16(4):494–502. [PubMed: 13498604]
11. Altman R, Asch E, Bloch D, Bole G, Borenstein D, Brandt K, et al. Development of criteria for the classification and reporting of osteoarthritis. Classification of osteoarthritis of the knee. Diagnostic and Therapeutic Criteria Committee of the American Rheumatism Association. Arthritis Rheum. Aug; 1986 29(8):1039–1049. [PubMed: 3741515]
12. Kutyavin IV, Milesi D, Belousov Y, Podyminogin M, Vorobiev A, Gorn V, et al. A novel endonuclease IV post-PCR genotyping system. Nucleic Acids Res. 2006; 34(19):e128. [PubMed: 17012270]
13. Devlin B, Roeder K. Genomic control for association studies. Biometrics. Dec; 1999 55(4):997–1004. [PubMed: 11315092]
14. Lau J, Ioannidis JP, Schmid CH. Quantitative synthesis in systematic reviews. Ann Intern Med. Nov 1; 1997 127(9):820–826. [PubMed: 9382404]
15. Higgins JP, Thompson SG. Quantifying heterogeneity in a meta-analysis. Stat Med. Jun 15; 2002 21(11):1539–1558. [PubMed: 12111919]
16. Higgins JP, Thompson SG, Deeks JJ, Altman DG. Measuring inconsistency in meta-analyses. BMJ. Sep 6; 2003 327(7414):557–560. [PubMed: 12958120]
17. Ioannidis JP, Patsopoulos NA, Evangelou E. Uncertainty in heterogeneity estimates in meta-analyses. BMJ. Nov 3; 2007 335(7626):914–916. [PubMed: 17974687]
18. Ioannidis JP. Calibration of credibility of agnostic genome-wide associations. Am J Med Genet B Neuropsychiatr Genet. Sep 5; 2008 147B(6):964–972. [PubMed: 18361430]
19. Manek NJ, Hart D, Spector TD, MacGregor AJ. The association of body mass index and osteoarthritis of the knee joint: an examination of genetic and environmental influences. Arthritis Rheum. Apr; 2003 48(4):1024–1029. [PubMed: 12687544]
20. Felson DT, Lawrence RC, Dieppe PA, Hirsch R, Helmick CG, Jordan JM, et al. Osteoarthritis: new insights. Part 1: the disease and its risk factors. Ann Intern Med. Oct 17; 2000 133(8):635–646. [PubMed: 11033593]
21. Felson DT, Zhang Y, Hannan MT, Naimark A, Weissman B, Aliabadi P, et al. Risk factors for incident radiographic knee osteoarthritis in the elderly: the Framingham Study. Arthritis Rheum. Apr; 1997 40(4):728–733. [PubMed: 9125257]
22. Blagojevic M, Jinks C, Jeffery A, Jordan KP. Risk factors for onset of osteoarthritis of the knee in older adults: a systematic review and meta-analysis. Osteoarthritis Cartilage. Jan; 18(1):24–33. [PubMed: 19751691]
23. Niu J, Zhang YQ, Torner J, Nevitt M, Lewis CE, Aliabadi P, et al. Is obesity a risk factor for progressive radiographic knee osteoarthritis? Arthritis Rheum. Mar 15; 2009 61(3):329–335. [PubMed: 19248122]
24. Toivanen AT, Heliovaara M, Impivaara O, Arokoski JP, Knekt P, Lauren H, et al. Obesity, physically demanding work and traumatic knee injury are major risk factors for knee osteoarthritis--a population-based study with a follow-up of 22 years. Rheumatology (Oxford). Feb; 49(2):308–314. [PubMed: 19946021]
25. Lohmander LS, Gerhardsson de Verdier M, Rollof J, Nilsson PM, Engstrom G. Incidence of severe knee and hip osteoarthritis in relation to different measures of body mass: a population-based prospective cohort study. Ann Rheum Dis. Apr; 2009 68(4):490–496. [PubMed: 18467514]
26. Kerkhof JM, Uitterlinden AG, Valdes AM, Hart DJ, Rivadeneira F, Jhamai M, et al. Radiographic osteoarthritis at three joint sites and FRZB, LRP5, and LRP6 polymorphisms in two population-based cohorts. Osteoarthritis Cartilage. Oct; 2008 16(10):1141–1149. [PubMed: 18406176]
27. Hart DJ, Spector TD. The classification and assessment of osteoarthritis. Baillieres Clin Rheumatol. May; 1995 9(2):407–432. [PubMed: 7656348]
28. Schiphof D, Boers M, Bierma-Zeinstra SM. Differences in descriptions of Kellgren and Lawrence grades of knee osteoarthritis. Ann Rheum Dis. Jul; 2008 67(7):1034–1036. [PubMed: 18198197]

Figure 1.
Q-Q plot of the expected vs observed distribution of p-values.

Europe PMC Funders Author Manuscripts

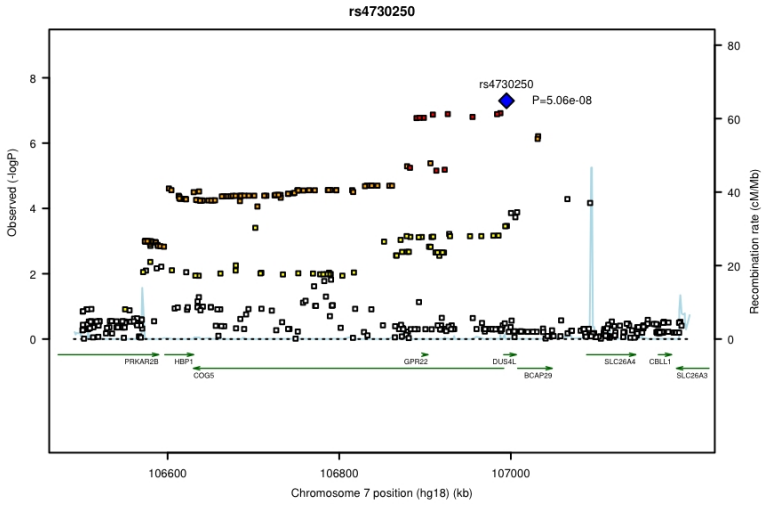

Figure 2.

Regional association plot of rs4730250. Statistical significance of the associated SNPs are illustrated on $-\log_{10}$ scale. The p-value of the rs4730250 and the other 10 selected SNPs are based on the meta-analysis of all datasets (both GWA studies and replication studies). P-values for the rest of the SNPs are based on the meta-analysis of the GWA studies. The sentinel SNP is shown in blue. The correlation of the sentinel SNP is shown on a scale from minimal (gray) to maximal (red). SNPs in red have $r^2 \geq 0.8$ with the sentinel SNP and SNPs in orange have $r^2 \geq 0.5$. Chromosome positions are based on HapMap release 22 build 36.

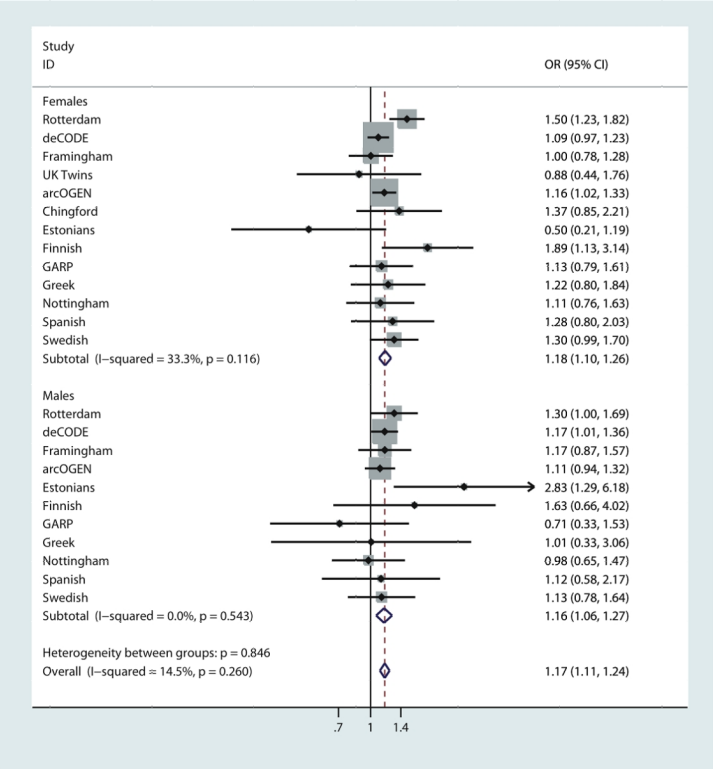

Figure 3.

a) Forest plot of study-specific estimates (black boxes) and summary OR estimates and 95% confidence intervals (95% CIs) (diamond) for the association between the rs4730250 SNP and knee osteoarthritis.

Europe PMC Funders Author Manuscripts

Table 1

Characteristics of the studies included in the analysis

Team	Knee OA Cases/ Controls	Platform used	Age mean (range)	BMI mean (range)	Females (%)	Knee OA definition	Control definition
GWA studies							
deCODE*	1033/32482	Infinium HapMap 300	69(19-99)	26(14-60)	58%	TKR	Health care records
Framingham	419/1674	Affymetrix GeneChip®	64(29-93)	26(14-54)	56%	Radiographic	Radiographic
Rotterdam	868/1464	Illumina HapMap550v 3	67(55-94)	26(16-56)	59%	Radiographic	Radiographic
TwinsUK	51/289	Infinium HapMap 300	54(37-76)	25(15-51)	100%	Radiographic	Radiographic
Replication cohorts stage 1							
arcOGEN	1643/4894	Illumina 610 Quad	NA	NA	71%	Radiographic /clinical	General population
Chingford[(a)]	64/236	NP	63 (54-77)	26 (17-43)	100%	Radiographic	Radiographic
Finnish	112/210	NP	67 (51-74)	29 (20-42)	75%	TKR	Population-based
Greek	368/606	NP	61(20-90)	26(17-34)	72%	Clinical	Clinical
GARP	161/758	NP	60(30-79)	27(19-47)	63%	Radiographic /clinical	Radiographic/clinical
Spanish	262/294	NP	66(32-94)	31(18-53)		TKR/clinical	Clinical
Nottingham[(b)]	647/237	NP	66 (40-97)	27 (15-51)	53%	TKR	Radiographic and clinical
Estonian	69/456	NP	47 (32-60)	28(15-47)	69%	Radiographic	Radiographic
Replication cohorts-Stage 2							
deCODE	622/32482[(c)]	Illumina and Centaurus (Nanogen)	77 (40-99)	29 (19-49)	63%	TKR	Population-based
Swedish	390/839	NP	62 (46-73)	29 (18-51)	63%	TKR+concomitant clinical & radiographic diagnosis of OA	General population without TKR

NP: Not pertinent; TKR: Total knee replacement; THR: total hip replacement

[(a)] Numbers excluding the samples already included in the arcOGEN study.

[(b)] Numbers excluding the samples already included in the arcOGEN study.

[(c)] same controls as for discovery cohort.

Europe PMC Funders Author Manuscripts

Table 2

Summary odds ratios and 95% confidence intervals of SNPs in the analysis including all European descent data.

SNP rs number	Minor (risk)allele	Chr	Position	Gene	MAF	OR (95% CI) Fixed effects	p-value	I2 (95% CI)	Cochran's Q
rs4730250	G	7	106994931	DUS4L	0.17	1.17 (1.11-1.24)	9.17×10^{-9}	15(0-49)	0.26
rs10953541	T	7	107031781	BCAP29	0.24	1.17 (1.10-1.23)	$3.90.\times10^{-8}$	19 (0-54)	0.23
rs3749132	A	2	68907001	ARHGAP25	0.07	1.17 (1.05-1.30)	4.08×10^{-3}	47 (0-74)	0.04
rs886827	C	7	42285581	GLI3	0.27	1.07 (0.99-1.16)	0.089	65 (43-80)	0.001
rs1886695	G	20	33643949	CPNE1	0.16	0.89 (0.84-0.95)	1.76×10^{-4}	42 (2-66)	0.02
rs10071956	T	5	173093290	Intergenic	0.38	1.12 (1.06-1.19)	5.05×10^{-5}	15 (0-53)	0.29
rs6816070	G	4	16089455	LDB2	0.42	0.91 (0.86-0.95)	1.34×10^{-4}	0 (0-54)	0.46
rs661924	T	10	21353562	NEBL	0.39	1.11 (1.05-1.17)	1.82×10^{-4}	30 (0-67)	0.18
rs436354	G	5	783271	ZDHC11	0.17	1.19 (1.01-1.30)	1.79×10^{-2}	41(2-63)	0.06
rs1994104	T	12	83040643	intergenic	0.13	0.88 (0.80-0.96)	3.13×10^{-3}	46 (2-70)	0.02
rs9857056	G	3	181698548	intergenic	0.12	1.11 (1.02-1.20)	1.65×10^{-2}	72 (43-87)	0.001

MAF: minor allele frequency; Minor allele is the OR allele