NOVEL GENETIC VARIANTS ASSOCIATED WITH LUMBAR DISC DEGENERATION IN NORTHERN EUROPEANS: A META-ANALYSIS OF 4,600 SUBJECTS

Williams FMK, Bansal AT, Menni C, Popham M, van Meurs J, Uitterlinden A, Hofman A, Suri P, Sambrook PN, Meulenbelt I, Hunter DJ, MacGregor AJ, Spector TD

Dept Twin Research and Genetic Epidemiology, King's College London, St Thomas' Hospital, London UK

Acclarogen Ltd, St John's Innovation Centre, Cambridge, UK

Erasmus

Framingham Heart Study

University of Sydney and Royal North Shore Hospital, Sydney Australia

GARP

Work funded by the Wellcome Trust, Arthritis Research UK

ABSTRACT

To date, the agnostic search of the genome by genome-wide association (GWA) to identify common variants associated with lumbar disc degeneration (LDD) has not been fruitful. Likely reasons for this include lack of clarity over what constitutes a 'case' as well as inadequate sample sizes, both well recognised attributes of the successful GWA study. In LDD the disc space narrows and osteophytes grow at the circumference of the disc. We have developed a continuous trait based on these 2 features which is measurable on all forms of imaging (plain radiograph, CT scan and MR imaging) and performed a meta-analysis of 5 cohorts of Northern European extraction having GWAS imputed to HapMap vs2. This study of >4,500 individuals identified 4 SNPs with $p < 5 \times 10\text{-}8$, the threshold set for genome-wide significance. We identified a novel variant in the PARK2 gene ($p = 2.8 \times 10^{-8}$) associated with LDD which has not hitherto been implicated in the condition. This work sheds new light on the pathogenesis of LDD.

Introduction

Lumbar disc degeneration (LDD) is a common, age-related trait [1] which has been shown to contribute to low back pain [2,3]. As low back pain is common in the general population and costly to society LDD is, therefore, of considerable public health importance [4]. The intervertebral disc comprises three distinct components; the nucleus pulposus, the annulus fibrosis and the cartilaginous endplate. Discrete biochemical, histological, metabolic and functional changes occur with age, such that the discs become dehydrated, lose disc height and there is accompanying in-growth of blood vessels and outgrowth of osteophytes from the vertebral body margin [5]. There are similarities with osteoarthritis where similar changes are seen in the cartilage. LDD has been shown to be heritable, with estimates of 65-80% according to precise phenotype studied [6,7]. Thus a considerable proportion of the variance in LDD is explained by genetic factors. To date, only a handful of genetic variants have been reliably determined using the candidate gene approach (reviewed by [8]) including VDR encoding the vitamin D receptor and some of the collagen genes such as COL9A2 and A3. A number of studies show conflicting results: these are likely due to small sample size or may reflect ethnic differences between Northern European and Asian populations, as seen in peripheral joint osteoarthritis [9]. The fact that increasing numbers of published genome-wide associations (GWAs) in common complex traits fail to replicate candidate gene findings suggest that there are limitations to the candidate gene method. An example from the bone mineral density literature is the GWAS meta-analysis of many thousands of individuals that reproduced only 9 of 150 reported candidate gene associations [10]. Even if a large proportion of replicated candidate gene studies in LDD are truly associated, a considerable proportion of the genetic variance in LDD remains unexplained [11]. In order to optimise sample size in the present study we performed meta-analysis of GWAS using a number of cohorts having the LDD phenotype. A variable was derived from measures of disc height and osteophytes obtained from lateral images on MR, CT scan or plain radiograph**.** Summing this variable over the 5 lumbar discs provided a continuous measure of disc degeneration.

Materials and methods

Using standardised coding methods and uniform transformation a continuously distributed LDD trait was derived for GWA in each cohort. Meta-analysis was performed of imputed GWA data from five population cohorts (Framingham, GARP, Rotterdam study 1 and 3 and TwinsUK) having imaging of the spine (see below). All cohorts had obtained fully informed consent from their participants and appropriate ethics committee approval. In all studies, a cumulative degeneration score was constructed from the sum of scores of degenerative changes at each level (disc space narrowing coded 0-3 and osteophytes, either anterior or posterior or both, coded 0-3). In those cohorts where only 4 disc levels were read (FHS) a fifth level was imputed by taking the mean reading for 4 discs as a surrogate for the fifth disc, and summing over 5 discs. The data underwent inverse normal transformation to generate a normally distributed variable.

Phenotyping the cohorts

1. **The Framingham Heart study (FHS)** is a longitudinal cohort of a defined population in Massachusetts, initiated in 1948 (www.framinghamheartstudy.org). It began as a study sample of 5,209 Framingham men and women between the ages of thirty and sixty. Subsequently, offspring and third generation subjects were incorporated. Every other year, after an extensive baseline examination, subjects undergo testing that includes a medical history, blood profile, echocardiogram, and bone, eye, and other tests. The subset of the Framingham subjects covered by the current analysis comprised 366 subjects from the Offspring and Generation 3 arms of the study who had undergone CT scanning of the spine, and the recruitment, conduct, and specifications of CT scanning have been reported elsewhere [12]. Measurement of the lumbar spine CTs for disc height and scoring (0-3) for anterior and posterior osteophytes was performed by a spine specialist using the mid-sagittal plane at spinal levels L2-L3, L3-L4, L4-L5, and L5-S1 by author PS using the atlas of Jarosz et al [7]. The measured values for disc height (mm) were converted to 0-3 categorical scale for disc height loss. Using the imputed value for the 5th lumbar vertebra, values for disc height loss and anterior and posterior osteophytes were summed over the 5 lumbar disc levels.

2. The Genetics, osteoARthrosis and Progression study (GARP) study comprises white sibling pairs of Dutch origin affected by osteoarthritis at multiple sites, and is aimed at identifying determinants of osteoarthritis susceptibility and progression. Probands (ages 40–70 years) and their siblings had osteoarthritis at multiple joint sites of the hand or in >2 of the following joint sites hand, spine (cervical or lumbar), knee or hip as described previously [13]. Subjects included in this study had undergone lateral radiographs of the spine (T4-S1). Each intervertebral disc level from L1/2 to L5/S1 was reviewed for the presence and severity of osteophytes (anterior) and disc narrowing, using the Lane atlas [14] where 0 = none; grade 1 = mild; grade 2 = moderate; and grade 3 = severe. The score at each level for anterior osteophytes and disc height loss were summed over the 5 lumbar levels.

3**. The Rotterdam study** is a prospective population-based follow-up study of the determinants and prognosis of chronic diseases in the elderly ([15,16]). All persons living in Ommoord, a suburb of Rotterdam, who were aged 55 years and over were invited to participate. A total of 7,983 participants were examined. For the current analysis, two subsets of the data were considered. Rotterdam cohort 1 (**RS1**) consists of 2,440 subjects; Rotterdam cohort 3 (**RS3**) consists of 974 subjects. Subjects originating from the Rotterdam study underwent plain radiography and scoring of LDD as previously described [3]. In brief, lateral lumbar radiographs were scored by a single observer for the presence of the individual radiographic features of disc degeneration. Each intervertebral disc from L1/2 to L5/S1 was reviewed for the presence and severity of osteophytes (anterior) and disc narrowing, using the Lane atlas as described above [14]. The scores for the 2 traits over the 5 lumbar discs were summed.

4. The TwinsUK registry (TUK) was described previously [17]. The register was started in 1993 and now comprises of approximately 10,000 monozygotic (MZ) and dizygotic (DZ) adult Caucasian twins aged 16 to 85 years from all over the United Kingdom, plus some parents and siblings. It now incorporates previous twin registries from the Institute of Psychiatry and Aberdeen University. This is a volunteer sample recruited by successive media campaigns without selecting for particular diseases or traits. All

twins receive a series of detailed disease and environment questionnaires. The majority of twins have been assessed in detail clinically at several time points for several hundred phenotypes related to common diseases or intermediate traits. The subset of TwinsUK covered by the current analysis consisted of 744 subjects who had participated in the spine MR study (scanned 1996-2000) using a Siemens MR machine with (Munich, Germany) 1.0-tesla superconducting magnet. Serial sagittal images of the cervical, thoraco-lumbar junction and lumbar spine (T9-L5) were obtained [7]. Images were coded for disc height loss and anterior osteophytes using a 0-3 scale in each case, where 0 is normal and 3 maximal degeneration as per the atlas of Jarosz et al [7]. All 5 lumbar discs were scored and the scores summed to give a combined LDD variable.

Genotyping and imputation

1. Framingham Heart Study subjects were genotyped using Affymetrix GeneCHip Human Mapping 500K array set and/or the 100K array set and/or the 50K array. Methods and quality controls have been described previously [18].

2. GARP subjects were genotyped using Illumina Human660W Quad BeadChips. Genotyping was performed at the genotyping Rotterdam Genotyping Centre. Positive strand, genotypes were called by clustering in Genome studio and imputation was performed using IMPUTE software and hapmap phase II v21[19,20]. Strict selection criteria were applied to the measured genotypes using a high information content (RT2 of >95%) and a minor allele frequency > 0.0025. Association analyses were performed using an in house developed software package that allows the analyses of family data using all information available in the cases and controls by extending the Cochran-Armitage trend test [21].

3. RS1 and RS3 subjects in the Rotterdam Study sets were genotyped on the HumanHap550v3 Genotyping BeadCHip (Illumina).

4. **TUK** subjects were genotyped using a combination of Illumina arrays (Human Hap300 and the Human Hap610Q). Genotyping was performed by the Wellcome Trust Sanger Institute using the Infinium assay (Illumina, San Diego, USA) across three genome-wide SNP sets, as described previously [22]. Genotyping results had been sent to KCL for collation and analysis using statistical package, STATA (StataCorp) [23]. Strict quality control was applied: 314,075 SMPs were retained for analysis (98.7%) – 733 were excluded because their call rates were <=90% and 725 SNPs had minor allele frequency < 0.01. In TwinsUK, significant population substructure was excluded using the STRUCTURE program.

Meta-analysis of the 5 study groups

Genotypes for 2.5-3 million autosomal SNPs were imputed separately to increase coverage using HapMap version 2 (www.hapmap.org) as reference panel. In GARP and TUK imputation was performed with Impute version 2[24]and in the other studies with MACH[25]. The common reference panel led to the reporting of results for the positive strand for all cohorts. In addition, allele pairs were compared between cohorts and no detectable strand-flips were found; the minor allele frequency was also compared between data-sets. The distributions of beta values of the cohorts were found to be similar and therefore suitable for meta-analysis. All directly genotyped or imputed autosomal SNPs having information from more than one study group (n=2,552,511) were included in the meta-analysis. Association results were combined using inverse variance weighted fixed effects meta-analysis. Two meta-analyses were run: the first was unadjusted; the second was adjusted for age and sex as both known risk factors for LDD and each risk factor was correlated with LDD in each study group. Heterogeneity of estimated effect was expressed using Q (weighted sum of squares) and I^2 (ratio of true heterogeneity to total observed variation). SNPs were excluded from the meta-analysis if the cohort-specific imputation quality as assessed by r2.hat (MACH) or .info (IMPUTE) metric was <0.40.

Results

The study samples for the meta-analysis included 4683 individuals of European ancestries. Table 2 shows sample size, demographic characteristics, LDD and lumbar spine imaging method for each independent cohort. Participants were mainly females (67.0%) and had mean age 57.7 years. Across the cohorts, mean level of LDD varied from 0.011 to 3.46, reflecting differences in imaging methods. However, the variance of the LDD variable were broadly similar (range 0.958 – 1.14), as were the distributions of the estimated genetic effect sizes (beta). Results from the adjusted analysis were broadly similar to the unadjusted: the Manhattan plots for the unadjusted and adjusted analyses are shown in Figure 1 and data from the analyses are shown in Table 3 (unadjusted) and Supplementary Table 1 (adjusted) for SNPs having $p<10^{-5}$. All signals with suggestive levels of significance are listed in Table 3a and Table 3b for unadjusted and adjusted analyses respectively. Quantile-quantile plots for LDD, both unadjusted and adjusted, are presented in Figure 2. Test statistic inflation post meta-analysis, as measured by the genomic control statistic [26] was low (λ_{GC} unadjusted =1.02; λ_{GC} adjusted=1.03) suggesting that relatedness had been adequately dealt with by the individual cohorts.

Four markers achieved genome-wide significance in the unadjusted GWAS, 3 of which were on chromosome 6 (rs926849; rs2187689; rs7767277), plus an intergenic marker on chromosome 3 (rs17034687). The results of the meta-analysis adjusted for age and gender were broadly similar, with p values slightly attenuated: the top signal for was also for SNP rs926849. This SNP lies on an intronic region of the Parkinson protein 2, E3 ubiquitin protein ligase (PARK2) gene on chromosome 6. Data were available from four studies and the range of estimated minor allele frequencies was 0.23-0.32. Imputation quality was high for all four studies contributing this SNP (>0.90). The minor or C allele of rs926849 was associated with a lower level of LDD implying that the minor allele is protective. Figure 4 shows association results (for the adjusted analysis) of both genotyped and imputed SNPs within 200 Kb of the PARK2 gene, along with recombination rates.

Two of the other strongly associated SNPs are in perfect LD: rs2187689 and rs7767277 on chromosome 6. Data were available for four studies and the range of estimated allele frequency was 0.05-0.10. Imputation quality was high for all four studies (>0.90). Both SNPs are in strong LD (r^2=0.76) with an intronic marker on the proteasome subunit, beta type 9, large multifunctional peptidase 2 gene (PSMB9) that is located in the class II region of the major histocompatibility complex (MHC). Figure 5 shows association results (for the adjusted analysis) of both genotyped and imputed SNPs within 400 Kb of rs2187689, along with recombination rates.

DISCUSSION

The availability of HapMap data and development of GWA technology has provided the ability for researchers to search the genome for associated variants without *a priori* assumptions of their involvement. While disadvantaged by the multiple testing involved, the main strength of the method is that it provides an agnostic search and so novel genes and pathways may be identified. This sort of approach may play an important role in conditions such as LDD where it is difficult for cell biologists to obtain fresh normal disc specimens for examination and the pathology that underlies LDD remains incompletely understood. GWA offers an unbiased scan of common genetic variants (minor allele frequency >5%) and thus may deliver novel variants in genes not hitherto suspected of playing a role in disc degeneration. At the time of writing there are no published genome-wide studies of lumbar disc degeneration (LDD), lumbar disc disease (disc bulge/prolapsed in cases vs controls), sciatica or back pain. This work is, therefore, the first to report on a genome-wide meta-analysis being conducted for lumbar disc degeneration. LDD is an age-related process which occurs in all people to some extent and may be detected as early as the teenage years [1]. LDD is known to have genetic determinants [7,16] and its expression is also influenced by gender (women develop LDD later), body mass index [27–31] and smoking [32]. Occupational factors also play a role in LDD [33]. LDD as determined by MR imaging has been implicated in the development of episodes of severe and disabling low back pain [2,7]. LDD has, therefore, considerable social and health-related costs which it is important to address. We undertook this large meta-analysis in order to identify novel genetic variants associated with LDD and to shed light on the underlying pathology of disc degeneration.

GWA data obtained using differing chip technology may be readily compared using imputation to HapMap. In total 2,543,887 overlapping markers were available in each cohort. We identified 4 markers having significant association with the LDD phenotype ($p<5x10^{-8}$). There was a marked similarity in the results obtained with and without adjustment for the covariates age and sex. A total of twenty-six markers had $p<10^{-5}$ in both meta-analyses. In both analyses there were multiple associations to the HLA region and to markers in PARK2 (Parkinson protein 2, E3 ubiquitin protein ligase). Among the most significant findings (Table 3) SNP rs926849 lies at 6q25.2-27 within an intron in the PARK2 gene, a large gene of 1.3 Mb comprising 12 exons. The SNP encodes a change of base from T to C and is reported to have minor allele frequency of 0.23 – 0.34 in dbSNP, which is keeping with the findings in our study groups (Table 3, Figures 3 and 4). Although this SNP has not been directly genotyped by any study group, estimates suggest imputation to be accurate for rs926849 (range 95-99%, Table 3). PARK2 encodes a protein called parkin which is a component of a multiprotein E3

ubiquitin ligase complex that mediates the targeting of unwanted proteins for proteasomal degradation. This complex also controls the level of proteins involved in several critical cell activities such as the timing of cell division and growth and for this reason it is postulated to be a tumour suppressor protein. Alternative splicing of the gene produces multiple transcript variants encoding distinct isoforms. Parkin is widely expressed in solid organs as well as skeletal muscle (http://www.proteinatlas.org/). Mutations within PARK2 are associated with autosomal recessive juvenile Parkinson's disease, Alzheimer's disease, diabetes mellitus and a number of solid tumours (reviewed in [34]). It has been postulated that Parkin accounts for the observed inverse relation between Parkinson's disease and cancer epidemiology [35]. In addition, variation within the PARK2 gene have been shown to alter the risk of leprosy [36] and other infectious diseases [37].

Three further markers in the unadjusted meta-analysis had $p<5x10^{-8}$. Marker rs17034687 is an intergenic marker on chromosome 3. Based on 1KG/CEU data, it is not in LD ($r^2>0.3$) with any known gene-based markers. Markers rs2187689 (Fig 5) and rs7767277 are HLA-region markers, neither of which is included in the 1KG pilot data. Using data from HapMap version 3 (release 2), rs2187689 and rs776277 are in perfect LD with each other and in LD ($r^2=0.76$) with an intronic marker in PSMB9 (proteasome (prosome, macropain) subunit, beta type 9; large multifunctional peptidase 2). Proteasomes are distributed throughout eukaryotic cells at high concentration and cleave peptides in an ATP/ubiquitin-dependent process in a non-lysosomal pathway. The gene is located in the class II region of the MHC (major histocompatibility complex). Expression of the gene is induced by gamma interferon and this gene product replaces catalytic subunit 1 (proteasome beta 6 subunit) in the immunoproteasome.

While lumbar degeneration is not considered an inflammatory process and has not been reported to be auto-immune in aetiology, there is evidence of proinflammatory cytokine activation in degenerate, particularly herniated, lumbar discs [38] and further, anti-TNF has been used successfully to treat disc herniation [39]. Of note, the COL11A2 gene lies 164 KB upstream from rs2187689. A SNP (rs2076311) within this very reasonable candidate gene has been shown to be associated with MR-determined disc signal intensity in a candidate gene study of Finnish male twins [40]. SNP rs2076311 is not, however, in LD with our top hit, rs2187689 ($R^2 = 0.017$) so it seems unlikely that this collagen-encoding gene accounts for our results. So many published GWA studies have identified SNPs in intergenic regions and gene deserts, it is clear that there is an as-yet undefined role for these regions. Long range enhancers, for example, could operate in this region so an influence on Col XI alpha2 expression cannot be excluded.

Of possible significance is SNP rs4802666($p = 3.76x10^{-06}$, adjusted meta-analysis) which lies within the MYH14 gene which encodes myosin, heavy chain 14, non-muscle. It is expressed in cell lines derived from bone (www.proteinatlas.org) and is implicated in autosomal dominant hearing impairment. It is of interest in LDD because it lies on chromosome 19 under the linkage peak we have reported in twins for LDD [41] and a peak reported by the Framingham group for hand osteoarthritis [42]. As there is a known phenotypic relationship between these two sites, this region on chromosome 19 forms a plausible candidate region for OA. It is not impossible that a muscle-expressed protein plays a role LDD through mechanisms similar to those proposed for OA [43].

The main limitation of the study is one of obtaining accurate phenotype on individuals which is known to be an important factor in the success of GWA [44]. There is no agreed gold standard imaging method in the determination of LDD, although it is recognised that MR imaging offers the most sensitive, widely available tool. Even so, MR is still expensive and many of the largest cohorts of spine imaging have plain radiographs, with more limited phenotypic information. The coding method applied to the imaging is also yet to be formally standardised. In order to obtain sufficient sample size, a number of cohorts were recruited having different imaging methods, but traits were selected such that they were comparable across the cohorts. Thus cohorts recoded their imaging where necessary to meet uniform requirements for inclusion. We included measures of disc height (coded 0-3) and anterior osteophytes (in RS1, RS3, GARP and TUK, also coded 0-3) and posterior osteophytes in FHS (coded 0-3). These sub-phenotypes were summed over the 5 discs and underwent inverse normal transformation to give a normal distribution. A further limitation is that 4 cohorts are population samples while GARP is derived from OA-affected sib pairs. We included GARP because it has made a contribution to similar analyses performed for OA [45]and, with adjustment for relatedness, appears to provide data comparable to other studies. While the differing methods of imaging provide different amounts of information, so the LDD variable has lower mean in those cohorts with radiographs, the variance of each groups LDD variable is comparable. Where GARP samples made a contribution to the meta-analysis (a number of the significant SNPs did not include a contribution from GARP, Table 3) the MAF was similar to those of other groups. The TUK group has a disproportionate number of women, for historical reasons. The men were retained, however, as they did not differ significantly from women in the LDD variable or BMI (data not shown). This study lacks a replication group. A second sample of similar size to the first is considered important to show that the findings of the first sample are true positives. Unfortunately there are, to our knowledge, no other collections of Northern Europeans having spine imaging which together would approach 4,500 individuals. There is considerable evidence in the literature that the genetic predisposition between Northern Europeans and Asians to OA is different [46] and given the phenotypic and genetic similarities between OA and LDD, replication should be made in Northern Europeans. We elected to include all the subjects in a single, powerful study rather than split the sample and reduce the chances of finding significant novel loci.

This is the first large-scale GWA study of LDD and we have identified several novel variants in the PARK2 gene and in PSMB9 within MHC class 2.

Reference List

1. Powell MC, Wilson M, Szypryt P, Symonds EM, Worthington BS (1986) Prevalence of lumbar disc degeneration observed by magnetic resonance in symptomless women. Lancet 2: 1366-1367. S0140-6736(86)92008-8 [pii].

2. MacGregor AJ, Andrew T, Sambrook PN, Spector TD (2004) Structural, psychological, and genetic influences on low back and neck pain: a study of adult female twins. Arthritis Rheum 51: 160-167.

3. de Schepper EI, Damen J, van Meurs JB, Ginai AZ, Popham M, Hofman A, Koes BW, Bierma-Zeinstra SM (2010) The association between lumbar disc degeneration and low back pain: the influence of age, gender, and individual radiographic features. Spine (Phila Pa 1976) 35: 531-536. 10.1097/BRS.0b013e3181aa5b33 [doi].

4. Phillips CJ (2006) Economic burden of chronic pain. Expert Rev Pharmacoecon Outcomes Res 6: 591-601. 10.1586/14737167.6.5.591 [doi].

5. Roberts S, Evans H, Trivedi J, Menage J (2006) Histology and pathology of the human intervertebral disc. J Bone Joint Surg Am 88 Suppl 2: 10-14. 88/1_suppl_2/10 [pii];10.2106/JBJS.F.00019 [doi].

6. Battie MC, Videman T, Levalahti E, Gill K, Kaprio J (2008) Genetic and environmental effects on disc degeneration by phenotype and spinal level: a multivariate twin study. Spine 33: 2801-2808.

7. Sambrook PN, MacGregor AJ, Spector TD (1999) Genetic influences on cervical and lumbar disc degeneration: a magnetic resonance imaging study in twins. Arthritis Rheum 42: 366-372.

8. Ryder JJ, Garrison K, Song F, Hooper L, Skinner J, Loke Y, Loughlin J, Higgins JP, MacGregor AJ (2007) Genetic associations in peripheral joint osteoarthritis and spinal degenerative disease: a systematic review. Ann Rheum Dis 67: 584-591.

9. Valdes AM, Loughlin J, Oene MV, Chapman K, Surdulescu GL, Doherty M, Spector TD (2006) Sex and ethnic differences in the association of ASPN, CALM1, COL2A1, COMP, and FRZB with genetic susceptibility to osteoarthritis of the knee. Arthritis Rheum 56: 137-146.

10. Richards JB, Kavvoura FK, Rivadeneira F, Styrkarsdottir U, Estrada K, Halldorsson BV, Hsu YH, Zillikens MC, Wilson SG, Mullin BH, Amin N, Aulchenko YS, Cupples LA, Deloukas P, Demissie S, Hofman A, Kong A, Karasik D, van Meurs JB, Oostra BA, Pols HA, Sigurdsson G, Thorsteinsdottir U, Soranzo N, Williams FM, Zhou Y, Ralston SH, Thorleifsson G, van Duijn CM, Kiel DP, Stefansson K, Uitterlinden AG, Ioannidis JP, Spector TD (2009) Collaborative meta-analysis: associations of 150 candidate genes with osteoporosis and osteoporotic fracture. Ann Intern Med 151: 528-537. 151/8/528 [pii].

11. Manolio TA, Collins FS, Cox NJ, Goldstein DB, Hindorff LA, Hunter DJ, McCarthy MI, Ramos EM, Cardon LR, Chakravarti A, Cho JH, Guttmacher AE, Kong A, Kruglyak L, Mardis E, Rotimi CN, Slatkin M, Valle D, Whittemore AS, Boehnke M, Clark AG, Eichler EE, Gibson G, Haines JL, Mackay TF, McCarroll SA, Visscher PM (2009) Finding the missing heritability of complex diseases. Nature 461: 747-753. nature08494 [pii];10.1038/nature08494 [doi].

12. Hoffmann U, Siebert U, Bull-Stewart A, Achenbach S, Ferencik M, Moselewski F, Brady TJ, Massaro JM, O'Donnell CJ (2006) Evidence for lower variability of coronary artery calcium mineral mass measurements by multi-detector computed tomography in a community-based cohort--consequences for progression studies. Eur J Radiol 57: 396-402. S0720-048X(05)00450-X [pii];10.1016/j.ejrad.2005.12.027 [doi].

13. Riyazi N, Meulenbelt I, Kroon HM, Ronday KH, Hellio le Graverand MP, Rosendaal FR, Breedveld FC, Slagboom PE, Kloppenburg M (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. 10.1136/ard.2004.024661 [doi];ard.2004.024661 [pii].

14. Lane NE, Nevitt MC, Genant HK, Hochberg MC (1993) Reliability of new indices of radiographic osteoarthritis of the hand and hip and lumbar disc degeneration. J Rheumatol 20: 1911-1918.

15. Hofman A, Breteler MM, van Duijn CM, Janssen HL, Krestin GP, Kuipers EJ, Stricker BH, Tiemeier H, Uitterlinden AG, Vingerling JR, Witteman JC (2009) The Rotterdam Study: 2010 objectives and design update. Eur J Epidemiol 24: 553-572. 10.1007/s10654-009-9386-z [doi].

16. Bijkerk C, Houwing-Duistermaat JJ, Valkenburg HA, Meulenbelt I, Hofman A, Breedveld FC, Pols HA, van Duijn CM, Slagboom PE (1999) Heritabilities of radiologic osteoarthritis in peripheral joints and of disc degeneration of the spine. Arthritis Rheum 42: 1729-1735.

17. Spector TD, Williams FM (2006) The UK Adult Twin Registry (TwinsUK). Twin Res Hum Genet 9: 899-906.

18. Suchindran S, Rivedal D, Guyton JR, Milledge T, Gao X, Benjamin A, Rowell J, Ginsburg GS, McCarthy JJ (2010) Genome-wide association study of Lp-PLA(2) activity and mass in the Framingham Heart Study. PLoS Genet 6: e1000928. 10.1371/journal.pgen.1000928 [doi].

19. 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. ng2088 [pii];10.1038/ng2088 [doi].

20. Howie BN, Donnelly P, Marchini J (2009) A flexible and accurate genotype imputation method for the next generation of genome-wide association studies. PLoS Genet 5: e1000529. 10.1371/journal.pgen.1000529 [doi].

21. Uh HW, Wijk HJ, Houwing-Duistermaat JJ (2009) Testing for genetic association taking into account phenotypic information of relatives. BMC Proc 3 Suppl 7: S123.

22. Richards JB, Rivadeneira F, Inouye M, Pastinen TM, Soranzo N, Wilson SG, Andrew T, Falchi M, Gwilliam R, Ahmadi KR, Valdes AM, Arp P, Whittaker P, Verlaan DJ, Jhamai M, Kumanduri V, Moorhouse M, van Meurs JB, Hofman A, Pols HA, Hart D, Zhai G, Kato BS, Mullin BH, Zhang F, Deloukas P, Uitterlinden AG, Spector TD (2008) Bone mineral density, osteoporosis, and osteoporotic fractures: a genome-wide association study. Lancet 371: 1505-1512.

23. (2007) Stata for Windows, version version 10 [computer program]. College Station, Texas, USA: Statacorp.

24. Howie BN, Donnelly P, Marchini J (2009) A flexible and accurate genotype imputation method for the next generation of genome-wide association studies. PLoS Genet 5: e1000529. 10.1371/journal.pgen.1000529 [doi].

25. Li Y, Willer CJ, Ding J, Scheet P, Abecasis GR (2010) MaCH: using sequence and genotype data to estimate haplotypes and unobserved genotypes. Genet Epidemiol 34: 816-834. 10.1002/gepi.20533 [doi].

26. Devlin B, Roeder K (1999) Genomic control for association studies. Biometrics 55: 997-1004.

27. Liuke M, Solovieva S, Lamminen A, Luoma K, Leino-Arjas P, Luukkonen R, Riihimaki H (2005) Disc degeneration of the lumbar spine in relation to overweight. Int J Obes (Lond) 29: 903-908.

28. Pye SR, Reid DM, Adams JE, Silman AJ, O'Neill TW (2007) Influence of weight, body mass index and lifestyle factors on radiographic features of lumbar disc degeneration. Ann Rheum Dis 66: 426-427. 66/3/426 [pii];10.1136/ard.2006.057166 [doi].

29. Symmons DP, van Hemert AM, Vandenbroucke JP, Valkenburg HA (1991) A longitudinal study of back pain and radiological changes in the lumbar spines of middle aged women. II. Radiographic findings. Ann Rheum Dis 50: 162-166.

30. Yoshimura N, Dennison E, Wilman C, Hashimoto T, Cooper C (2000) Epidemiology of chronic disc degeneration and osteoarthritis of the lumbar spine in Britain and Japan: a comparative study. J Rheumatol 27: 429-433.

31. Williams FM, Popham M, Livshits G, Sambrook PN, Spector TD, MacGregor AJ (2010) A response to Videman et al., "challenging the cumulative injury model: positive effects of greater body mass on disc degeneration". Spine J 10: 571-572. S1529-9430(10)00194-4 [pii];10.1016/j.spinee.2010.03.002 [doi].

32. Battie MC, Videman T, Gill K, Moneta GB, Nyman R, Kaprio J, Koskenvuo M (1991) 1991 Volvo Award in clinical sciences. Smoking and lumbar intervertebral disc degeneration: an MRI study of identical twins. Spine 16: 1015-1021.

33. Videman T, Battie MC (1999) The influence of occupation on lumbar degeneration. Spine 24: 1164-1168.

34. Kay DM, Stevens CF, Hamza TH, Montimurro JS, Zabetian CP, Factor SA, Samii A, Griffith A, Roberts JW, Molho ES, Higgins DS, Gancher S, Moses L, Zareparsi S, Poorkaj P, Bird T, Nutt J, Schellenberg GD, Payami H (2010) A comprehensive analysis of deletions, multiplications, and copy number variations in PARK2. Neurology 75: 1189-1194. 75/13/1189 [pii];10.1212/WNL.0b013e3181f4d832 [doi].

35. Gao X, Ning Y (2011) Cancer and Parkinson's disease: the odd couple. Drugs Today (Barc) 47: 215-222. 1519657 [pii];10.1358/dot.2011.47.3.1519657 [doi].

36. Mira MT, Alcais A, Nguyen VT, Moraes MO, Di FC, Vu HT, Mai CP, Nguyen TH, Nguyen NB, Pham XK, Sarno EN, Alter A, Montpetit A, Moraes ME, Moraes JR, Dore C, Gallant CJ, Lepage P, Verner A, van d, V, Hudson TJ, Abel L, Schurr E (2004) Susceptibility to leprosy is associated with PARK2 and PACRG. Nature 427: 636-640. 10.1038/nature02326 [doi];nature02326 [pii].

37. Ali S, Vollaard AM, Widjaja S, Surjadi C, van d, V, van Dissel JT (2006) PARK2/PACRG polymorphisms and susceptibility to typhoid and paratyphoid fever. Clin Exp Immunol 144: 425-431. CEI3087 [pii];10.1111/j.1365-2249.2006.03087.x [doi].

38. Shamji MF, Setton LA, Jarvis W, So S, Chen J, Jing L, Bullock R, Isaacs RE, Brown C, Richardson WJ (2010) Proinflammatory cytokine expression profile in degenerated and herniated human intervertebral disc tissues. Arthritis Rheum 62: 1974-1982. 10.1002/art.27444 [doi].

39. Korhonen T, Karppinen J, Paimela L, Malmivaara A, Lindgren KA, Bowman C, Hammond A, Kirkham B, Jarvinen S, Niinimaki J, Veeger N, Haapea M, Torkki M, Tervonen O, Seitsalo S, Hurri H (2006) The treatment of disc-herniation-induced sciatica with infliximab: one-year follow-up results of FIRST II, a randomized controlled trial. Spine (Phila Pa 1976) 31: 2759-2766. 10.1097/01.brs.0000245873.23876.1e [doi];00007632-200611150-00002 [pii].

40. Videman T, Saarela J, Kaprio J, Nakki A, Levalahti E, Gill K, Peltonen L, Battie MC (2009) Associations of 25 structural, degradative, and inflammatory candidate genes with lumbar disc desiccation, bulging, and height narrowing. Arthritis Rheum 60: 470-481. 10.1002/art.24268 [doi].

41. Williams FM, Kato BS, Livshits G, Sambrook PN, Spector TD, MacGregor AJ (2008) Lumbar disc disease shows linkage to chromosome 19 overlapping with a QTL for hand OA. Ann Rheum Dis 67: 117-119. ard.2007.073262 [pii];10.1136/ard.2007.073262 [doi].

42. Demissie S, Cupples LA, Myers R, Aliabadi P, Levy D, Felson DT (2002) Genome scan for quantity of hand osteoarthritis: the Framingham Study. Arthritis Rheum 46: 946-952.

43. Brandt KD, Radin EL, Dieppe PA, van de PL (2006) Yet more evidence that osteoarthritis is not a cartilage disease. Ann Rheum Dis 65: 1261-1264.

44. Amos CI (2007) Successful design and conduct of genome-wide association studies. Hum Mol Genet 16 Spec No. 2: R220-R225. ddm161 [pii];10.1093/hmg/ddm161 [doi].

45. Zhai G, van Meurs JB, Livshits G, Meulenbelt I, Valdes AM, Soranzo N, Hart D, Zhang F, Kato BS, Richards JB, Williams FM, Inouye M, Kloppenburg M, Deloukas P, Slagboom E, Uitterlinden A, Spector TD (2009) A genome-wide association study suggests that a locus within the ataxin 2 binding protein 1 gene is associated with hand osteoarthritis: the Treat-OA consortium. J Med Genet 46: 614-616. jmg.2009.067314 [pii];10.1136/jmg.2009.067314 [doi].

46. Valdes AM, Spector TD (2011) Genetic epidemiology of hip and knee osteoarthritis. Nat Rev Rheumatol 7: 23-32. nrrheum.2010.191 [pii];10.1038/nrrheum.2010.191 [doi].

Table 1 .Characteristics of the study samples

	FHS	GARP	RS1	RS3	TUK
N	330	192	2440	974	744
Age, years	54.3 (11.0)	60.3 (7.1)	65.7 (6.7)	54.7 (3.4)	53.6 (8.3)
Women (%)	42.2	79.7	57	59	96.8
BMI, kg/m ^{2}	28.1 (5.1)	XXX	26.3 (3.4)	27.12 (4.6)	24.9 (4.4)
Lumbar spine imaging	CT	radiograph	radiograph	radiograph	MRI
LDD variable	2.49 (0.97)	0.02 (0.958)	0.006 (0.978)	0.011 (0.965)	3.46 (1.14)

Legend to Table 1

FHS represents Framingham Heart Study; GARP, Genetics of OsteoArthrosis and Progression study; RS1, Rotterdam study cohort 1; RS3:, Rotterdam study cohort 3; TUK, TwinsUK: BMI, body mass index; MRI, magnetic resonance imaging; CT, computed tomography

Values are mean (SD) unless specified otherwise.

Table 2. Genotyping and imputation methods by study

	FHS	GARP	RS1	RS3	TUK
Genotyping					
Platform	Affymetrix	Illumina	Illumina	Illumina	Illumina
Chip	500K or 1000K or 10 K	Human660W quad	Human Hap550	Human Hap550	Human Hap300 or Human Hap610Q
Call rate	>0.97	> 0.95	> 0.975	> 0.975	>0.90
MAF	> 0.05	> 0.0025	> 0.01	> 0.01	>0.01
HWE p-value	> 10^{-6}	XXX	> 10^{-6}	> 10^{-6}	> 10^{-4}
Imputation					
Variants included for imputation	534,982	xxx	512,349	514,073	314,075
Imputation software	MACH vs1	IMPUTE	MACH	MACH	IMPUTE vs 2
Genome build	HapMap Phase II release 22	Hapmap Phase II release 21	Hapmap Phase II release 22	Hapmap Phase II release 22	
Total number of SNPs available	2,543,887	1,830,601	2,543,887	2,543,887	3,044,064

Legend to Table 2

FHS represents Framingham Heart Study; GARP, Genetics of OsteoArthrosis and Progression study; RS1, Rotterdam study cohort 1; RS3:, Rotterdam study cohort 3; TUK, TwinsUK: BMI, body mass index; MAF, Minor Allele Frequency; HWE, Hardy-Weinberg equilibrium

Table 3. Results of the GWA meta-analysis (unadjusted, showing those SNPs having $p<10^{-5}$).

SNP	Gene	Chr	position	N	MAF.R1	MAF.R3	MAF.T	MAF.F	MAF.G	R2.R1	R2.R3	Inf.T	R2.F	R2.G	Eff All	beta	SE	P
rs17034687	NA	3	3654852	4429	0.09	0.10	0.05	0.08	NA	0.92	0.93	0.88	0.82	NA	C	0.23	0.0383	1.82E-09
rs2187689	NA	6	32960625	4457	0.08	0.07	0.06	0.09	NA	0.98	0.99	0.95	0.93	NA	C	0.23	0.0411	2.72E-08
rs7767277	NA	6	32961020	4457	0.08	0.07	0.06	0.09	NA	0.98	0.99	0.95	0.93	NA	A	0.23	0.0411	2.81E-08
rs926849	PARK2	6	162081609	3939	0.31	0.32	NA	0.31	0.23	0.98	0.99	NA	0.90	95.04	C	-0.13	0.0243	3.25E-08
rs7744666	NA	6	32967690	4466	0.10	0.09	0.06	0.10	NA	0.99	1.00	0.97	0.96	NA	C	0.2	0.0367	5.58E-08
rs11969002	NA	6	32967726	4466	0.10	0.09	0.06	0.10	NA	0.99	1.00	0.97	0.96	NA	A	0.2	0.0367	5.59E-08
rs6457690	NA	6	32963695	4464	0.10	0.09	0.07	0.11	NA	0.98	1.00	0.96	0.97	NA	A	0.19	0.0364	9.36E-08
rs1029296	NA	6	32964359	4464	0.10	0.09	0.07	0.11	NA	0.98	1.00	0.96	0.97	NA	C	0.19	0.0364	9.39E-08
rs6936004	NA	6	32964912	4462	0.10	0.09	0.07	0.11	NA	0.98	1.00	0.96	0.97	NA	C	0.19	0.0364	1.04E-07
rs3749982	NA	6	32970585	4458	0.10	0.09	0.06	0.10	NA	0.99	1.00	0.96	0.96	NA	A	0.19	0.0367	1.46E-07
rs9469300	NA	6	32968730	4482	0.10	0.09	0.07	0.10	NA	0.99	1.00	0.92	0.96	NA	A	0.19	0.0365	1.47E-07
rs10214886	NA	6	32965397	4479	0.10	0.09	0.07	0.11	NA	0.98	1.00	0.92	0.97	NA	A	0.19	0.0363	2.32E-07
rs10046257	NA	6	32962675	4461	0.10	0.09	0.08	0.11	NA	0.98	1.00	0.96	0.97	NA	A	0.19	0.0366	3.22E-07
rs4875102	NA	8	4272100	4608	0.26	0.26	0.27	0.25	0.29	0.99	0.99	0.95	0.91	97.65	A	-0.12	0.0238	3.61E-07
rs3019449	PARK2	6	162085083	4636	0.32	0.32	0.31	0.31	0.28	0.98	0.98	0.98	0.99	97.18	A	-0.12	0.0228	3.68E-07
rs1029295	NA	6	32964460	3747	0.10	0.09	NA	0.11	NA	0.98	1.00	NA	0.97	NA	C	0.19	0.0382	4.58E-07
rs9301951	GPC6	13	93750833	4397	0.04	0.04	0.03	0.05	NA	0.96	0.97	0.74	0.78	NA	C	-0.26	0.0539	9.47E-07
rs7896691	PFKP	10	3145173	4652	0.10	0.10	0.09	0.13	0.07	0.99	1.00	0.96	0.67	97.61	C	0.17	0.0354	2.02E-06
rs6602024	PFKP	10	3145237	4673	0.10	0.10	0.09	0.13	0.07	0.99	1.00	0.98	0.67	97.62	A	0.17	0.0354	2.20E-06
rs1884158	PARK2	6	162081656	4491	0.31	0.32	0.31	0.31	NA	0.98	0.99	0.99	0.91	NA	C	-0.11	0.0231	3.56E-06
rs10998466	NA	10	70296435	4446	0.01	0.01	0.01	0.02	NA	0.75	0.97	0.73	0.51	NA	A	-0.53	0.1136	3.59E-06
rs17132175	PFKP	10	3140814	4659	0.10	0.10	0.09	0.13	0.07	0.98	0.99	0.97	0.63	97.13	C	0.16	0.0356	3.73E-06
rs1981483	PIGQ	16	570666	3939	0.42	0.43	NA	0.42	0.47	0.99	0.98	NA	0.44	99.64	A	0.11	0.023	3.75E-06
rs1154053	NA	8	4272798	4639	0.17	0.17	0.18	0.16	0.20	0.99	1.00	0.96	0.89	96.17	C	-0.13	0.028	3.99E-06
rs737631	PARK2	6	162082423	4303	0.28	0.29	0.23	0.28	NA	0.92	0.91	0.89	0.83	NA	A	-0.11	0.0246	4.32E-06
rs2484990	NA	10	31555138	4429	0.01	0.01	0.00	0.02	NA	0.56	0.82	0.47	0.93	NA	C	0.68	0.148	4.39E-06
rs1250307	NA	10	31535980	4415	0.01	0.01	0.00	0.02	NA	0.56	0.82	0.45	0.94	NA	A	0.68	0.1484	4.39E-06
rs7204439	RAB40C	16	601336	3939	0.42	0.43	NA	0.43	0.47	0.97	0.98	NA	0.43	97.37	C	0.11	0.023	4.43E-06
rs4802666	MYH14	19	55412886	4094	0.27	0.27	0.18	0.27	NA	0.70	0.74	0.76	0.64	NA	A	-0.13	0.0286	4.55E-06
rs2484992	NA	10	31552104	4440	0.01	0.01	0.00	0.02	NA	0.56	0.82	0.48	0.93	NA	C	0.68	0.1485	4.81E-06
rs9488238	NA	6	114123826	4464	0.04	0.03	0.02	0.03	NA	0.87	0.88	0.87	0.97	NA	A	-0.28	0.0609	5.17E-06

rs763014	RAB40C	16	615681	3939	0.42	0.43	NA	0.43	0.47	0.98	0.99	NA	0.42	99.40	C	0.1	0.023	5.43E-06
rs1205863	NA	6	12051512	3939	0.06	0.06	NA	0.06	0.07	0.99	1.00	NA	0.99	98.23	G	0.21	0.0465	5.64E-06
rs11918654	ARL8B	3	5163246	4310	0.27	0.27	0.29	0.27	NA	0.97	0.98	0.90	0.99	NA	C	-0.11	0.0243	7.10E-06
rs4881085	PFKP	10	3145540	4490	0.10	0.10	0.10	0.13	NA	1.00	1.00	0.99	0.67	NA	A	0.16	0.0356	7.44E-06
rs2657195	NA	8	92629091	4316	0.22	0.21	0.18	0.21	NA	0.85	0.85	0.87	0.74	NA	A	-0.13	0.0281	7.54E-06
rs11754641	NA	6	65692644	4478	0.03	0.03	0.02	0.04	NA	0.93	0.95	0.91	1.00	NA	C	0.29	0.0644	7.84E-06
rs12805875	NA	11	103034842	4674	0.42	0.41	0.42	0.43	0.41	1.00	1.00	1.00	1.00	98.12	A	0.09	0.0209	8.51E-06
rs980238	NA	8	4270026	4491	0.30	0.29	0.31	0.29	NA	1.00	1.00	1.00	1.00	NA	A	-0.1	0.0234	9.11E-06
rs7103004	NA	11	103031234	4667	0.42	0.41	0.42	0.43	0.41	1.00	1.00	0.99	1.00	97.09	C	0.09	0.0209	9.20E-06
rs4554859	NA	11	103035576	4674	0.42	0.41	0.42	0.43	0.41	1.00	1.00	1.00	1.00	98.12	G	0.09	0.0209	9.24E-06
rs7118412	NA	11	103030964	4669	0.42	0.41	0.42	0.43	0.41	1.00	1.00	0.99	1.00	97.09	A	0.09	0.0209	9.37E-06
rs2017567	NA	16	577213	4683	0.42	0.43	0.42	0.42	0.47	0.99	1.00	1.00	0.45	100.00	C	0.1	0.0214	9.38E-06
rs710924	PIGQ	16	573354	4683	0.42	0.43	0.42	0.42	0.47	1.00	1.00	1.00	0.45	99.64	C	0.09	0.0214	9.93E-06

Legend to Table 3

Studies contributing data are denoted RS1: Rotterdam study cohort 1; RS3: Rotterdam study cohort 3; TUK: TwinsUK: BMI: body mass index; FHS: Framingham Heart Study; GARP: Genetics of OsteoArthrosis and Progression study;

SNP single nucleotide polymorphism; Chr chromosome; position, SNP location in base pairs; MAF minor allele frequency;

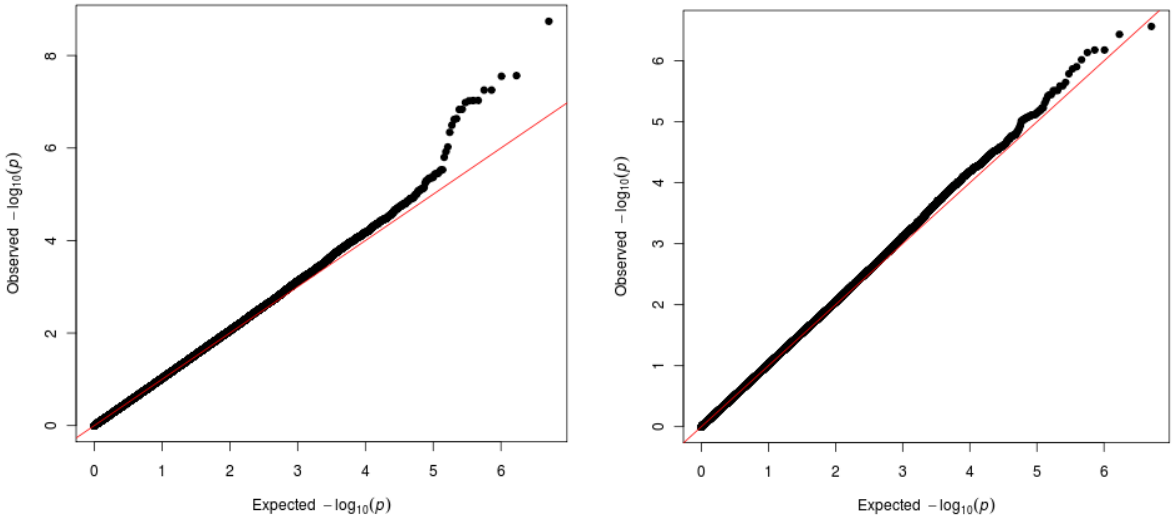
R2

Inf

Eff All, effector allele; beta, effect size; SE, standard error of beta; p, p value

Figure 1. Quantile-quantile plot of observed versus expected distribution of p values for the GWA meta-analysis

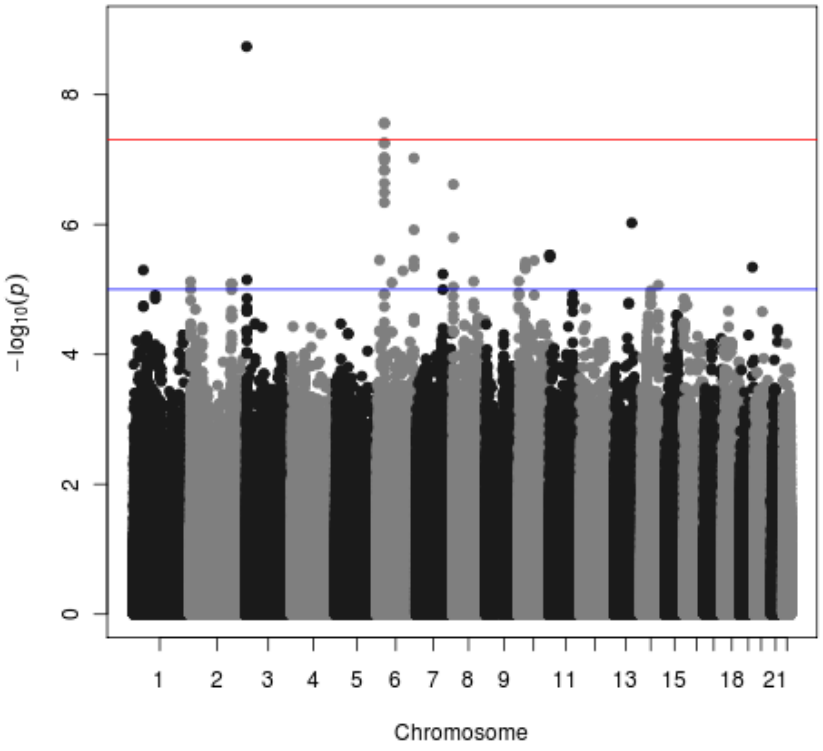
a) b)



Legend to Figure 1

The plots show GWA meta-analysis quantile-quantile plot a) unadjusted and b) adjusted for age and sex

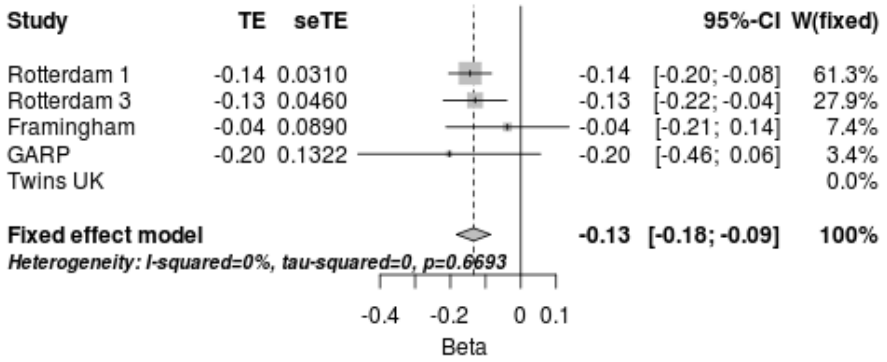
Figure 2. Manhattan plot for GWA meta-analysis unadjusted results



Legend to Figure 2

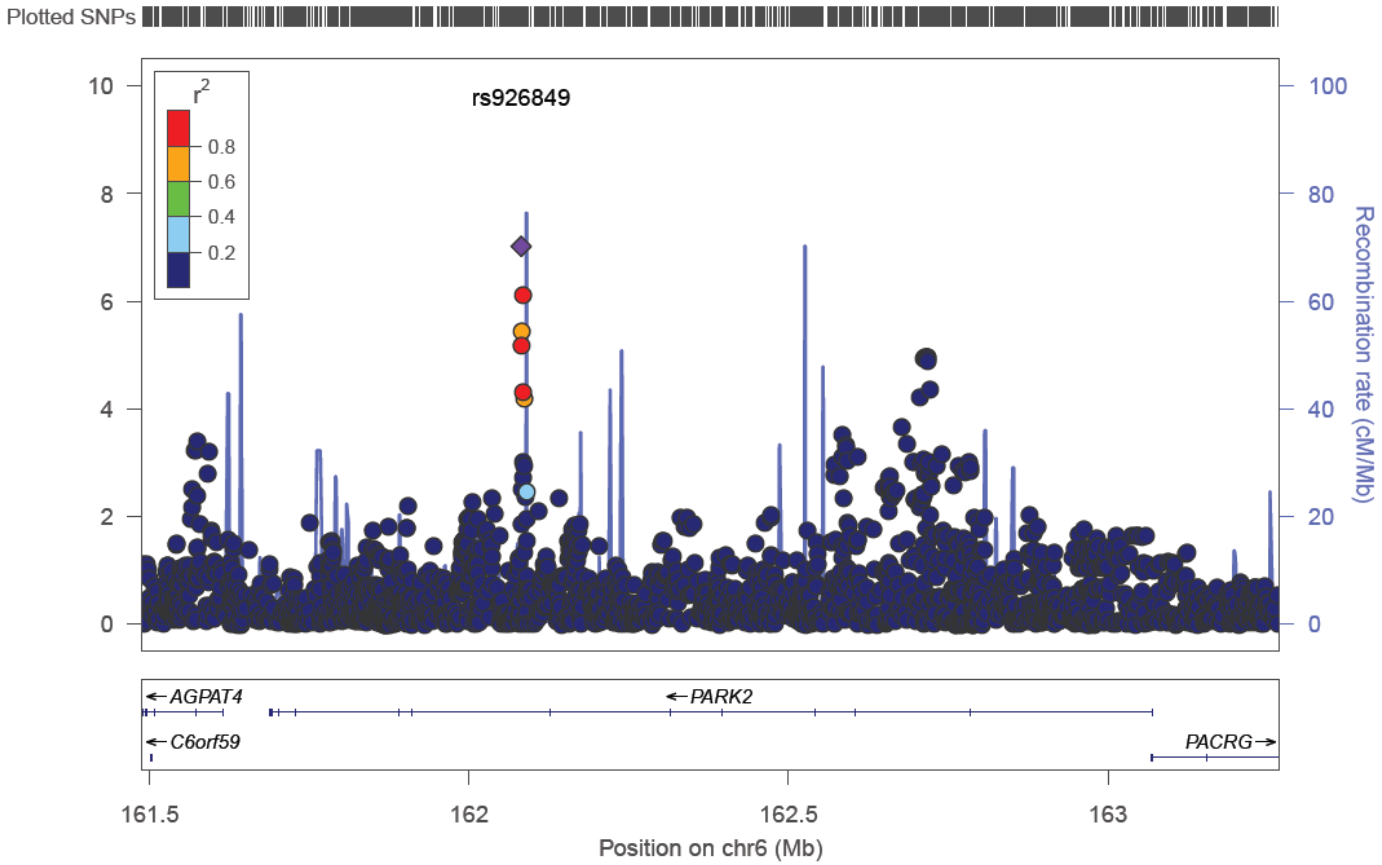
Plot shows combined results for the 5 studies included in the meta-analysis

Figure 3 Forest plot of rs926849 in PARK2 unadjusted for covariates



Legend to Figure 3

Figure 4. Regional plot of association results and recombination rates for the PARK2 gene adjusted for age and sex.

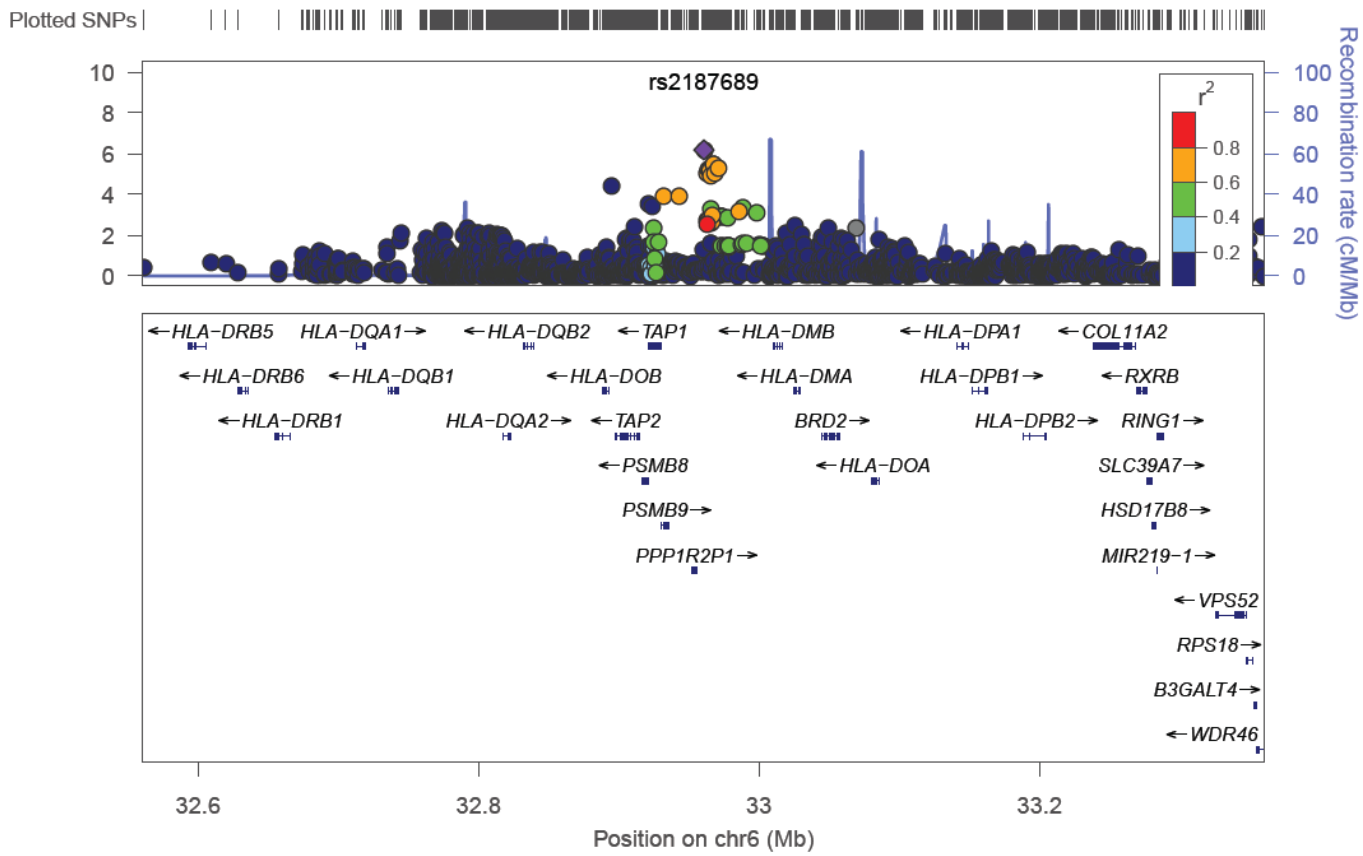


Legend to Figure 4

$-\log_{10}$ *P* values (*y* axis) of the SNPs are shown according to their chromosomal positions (*x* axis)

The colour intensity of each symbol reflects the extent of LD with the rs926849, coloured red ($r^2 > 0.8$) through to blue ($r^2 < 0.2$). Genetic recombination rates (cM/Mb), estimated using HapMap CEU samples, are shown with a light blue line. Physical positions are based on build 36 (NCBI) of the human genome. Also shown are the relative positions of genes mapping to the region of association. Genes have been redrawn to show the relative positions, and therefore, the maps are not to physical scale.

Figure 5. Regional plot of association results and recombination rates for the class II MHC region adjusted for age and sex.



Legend to Figure 5

$-\log_{10}$ *P* values (*y* axis) of the SNPs are shown according to their chromosomal positions (*x* axis)

The colour intensity of each symbol reflects the extent of LD with the rs926849, coloured red ($r^2 > 0.8$) through to blue ($r^2 < 0.2$). Genetic recombination rates (cM/Mb), estimated using HapMap CEU samples, are shown with a light blue line. Physical positions are based on build 36 (NCBI) of the human genome. Also shown are the relative positions of genes mapping to the region of association. Genes have been redrawn to show the relative positions, and therefore, the maps are not to physical scale.

Supplementary Table 1. Results of the GWA meta-analysis adjusted for age and sex, showing those SNPs having $p<10^{-5}$.

SNP	gene	Chr	position	N	MAF.R1	MAF.R3	MAF.T	MAF.F	MAF.G	R2.R1	R2.R3
rs926849	PARK2	6	162081609	3939	0.31	0.32	NA	0.31	0.23	0.98	0.99
rs17034687	NA	3	3654852	4429	0.09	0.10	0.05	0.08	NA	0.92	0.93
rs2187689	NA	6	32960625	4457	0.08	0.07	0.06	0.09	NA	0.98	0.99
rs7767277	NA	6	32961020	4457	0.08	0.07	0.06	0.09	NA	0.98	0.99
rs9488238	NA	6	114123826	4464	0.04	0.03	0.02	0.03	NA	0.87	0.88
rs3019449	PARK2	6	162085083	4636	0.32	0.32	0.31	0.31	0.28	0.98	0.98
rs9301951	GPC6	13	93750833	4397	0.04	0.04	0.03	0.05	NA	0.96	0.97
rs11136566	CSMD1	8	2922480	4610	0.39	0.39	0.36	0.39	0.46	0.99	0.98
rs7744666	NA	6	32967690	4466	0.10	0.09	0.06	0.10	NA	0.99	1.00
rs11969002	NA	6	32967726	4466	0.10	0.09	0.06	0.10	NA	0.99	1.00
rs365962	NA	16	85267450	4608	0.44	0.44	0.48	0.47	0.42	0.99	0.99
rs387953	NA	16	85271706	4683	0.46	0.45	0.48	0.46	0.43	1.00	0.99
rs737631	PARK2	6	162082423	4303	0.28	0.29	0.23	0.28	NA	0.92	0.91
rs4802666	MYH14	19	55412886	4094	0.27	0.27	0.18	0.27	NA	0.70	0.74
rs1154053	NA	8	4272798	4639	0.17	0.17	0.18	0.16	0.20	0.99	1.00
rs1205863	NA	6	12051512	3939	0.06	0.06	NA	0.06	0.07	0.99	1.00
rs3749982	NA	6	32970585	4458	0.10	0.09	0.06	0.10	NA	0.99	1.00
rs6457690	NA	6	32963695	4464	0.10	0.09	0.07	0.11	NA	0.98	1.00
rs1029296	NA	6	32964359	4464	0.10	0.09	0.07	0.11	NA	0.98	1.00
rs4875102	NA	8	4272100	4608	0.26	0.26	0.27	0.25	0.29	0.99	0.99
rs7234567	NA	18	3410806	4682	0.26	0.26	0.22	0.22	0.29	1.00	1.00
rs6936004	NA	6	32964912	4462	0.10	0.09	0.07	0.11	NA	0.98	1.00
rs1884158	PARK2	6	162081656	4491	0.31	0.32	0.31	0.31	NA	0.98	0.99
rs1981483	PIGQ	16	570666	3939	0.42	0.43	NA	0.42	0.47	0.99	0.98
rs10890236	NA	1	43348908	4683	0.09	0.09	0.09	0.09	0.09	1.00	1.00
rs1029295	NA	6	32964460	3747	0.10	0.09	NA	0.11	NA	0.98	1.00
rs763014	RAB40C	16	615681	3939	0.42	0.43	NA	0.43	0.47	0.98	0.99
rs10046257	NA	6	32962675	4461	0.10	0.09	0.08	0.11	NA	0.98	1.00
rs7204439	RAB40C	16	601336	3939	0.42	0.43	NA	0.43	0.47	0.97	0.98
rs9469300	NA	6	32968730	4482	0.10	0.09	0.07	0.10	NA	0.99	1.00
rs7205409	RAB40C	16	582611	4590	0.42	0.43	0.45	0.43	0.47	0.99	0.99
rs10998466	NA	10	70296435	4446	0.01	0.01	0.01	0.02	NA	0.75	0.97

Legend to Supplementary table 4

Studies contributing data are denoted RS1: Rotterdam study cohort 1; RS3: Rotterdam study cohort 3; TUK: TwinsUK: BMI: body mass index; FHS: Framingham Heart Study; GARP: Genetics of OsteoArthrosis and Progression study;

SNP single nucleotide polymorphism; Chr chromosome; position, SNP location in base pairs; MAF minor allele frequency;

R2 Inf

Eff All, effecter allele; beta, effect size; SE, standard error of beta; p, p value