Generalised Osteoarthritis: from Mendelian Disorder to Complex Disease

Min, J.L.

CHAPTER

4

A genome-wide scan in Dutch sibling pairs reveals the *DIO2* gene as new candidate gene for osteoarthritis at multiple joint sites

Josine L. Min, Ingrid Meulenbelt, Naghmeh Riyazi, Jeanine J. Houwing-Duistermaat, Henk-Jan van der Wijk, Albert B. Seymour, Nico Lakenberg, Cornelia M. van Duijn, Margreet Kloppenburg and P. Eline Slagboom

American Journal of Human Genetics 2006; Submitted for publication

Abstract

Osteoarthritis (OA) is a common late-onset joint disease with complex inheritance for which multiple susceptibility loci have been revealed. Here, we report a genome-wide nonparametric linkage analysis of 183 sibling pairs with a generalised OA phenotype from the Genetics OsteoARthritis and Progression (GARP) study. The GARP study consists of probands (aged 40-70 years) and their siblings of Dutch ancestry with predominantly symptomatic familial OA at multiple joint sites of the hand or in two or more of the following joints sites: hand, spine, knee or hip. Suggestive evidence for linkage was observed on chromosome 14q32.11 (LOD = 2.37, P = 0.0005). The location of the linkage peak revealed three candidate genes: calmodulin 1 (*CALM1*), fibronectin leucine rich transmembrane protein 2 (*FLRT2*) and iodothyronine deiodinase enzyme type 2 (*DIO2*). The *CALM1* gene was previously associated with hip OA in the Japanese population. Subsequent genotyping and joint modelling of linkage and association of tagging SNPs in these genes and the functional *CALM1* variant (rs12885713) did not explain the observed linkage signal. However, the C allele of *DIO2* rs225014 (P = 0.02) and the corresponding haplotype CCGC (P = 0.05) were shown to explain part of the linkage. DIO2 is a regulator of thyroid hormone metabolism in the growth plate and may confer susceptibility for OA at multiple joint sites. Together our results implicate for the first time that the thyroid hormone metabolism is involved in OA.

Introduction

Osteoarthritis (OA) is a common late-onset joint disease and is a major cause of pain and disability among the elderly. OA is diagnosed clinically and radiographically. Clinically, OA is predominantly based on symptoms of pain, stiffness and disability. Radiographic OA is characterised by osteophytes, sclerosis and joint space narrowing. OA occurs most frequently in the lumbar and cervical spine, interphalangeal joints and carpometacarpal joints of the hands, hips and knees. As early as in 1952, Kellgren and Moore[1] recognised primary generalised OA (GOA) as a different clinical entity identified in a specific subset of OA patients. Two types of GOA, occurring in either the presence or absence of Heberden's nodes (nodal or non-nodal form) were described[2]. Evidence for GOA has been confirmed further in population-based studies in which clustering between affected joint sites was observed more frequently than would be expected from an age and sex dependent effect[3-6]. However, a clear definition of GOA does not exist and varies depending on joint site (hand GOA) and the threshold of number of affected joints[7].

Heritabilities for OA range from 30-80% depending on gender or joint site investigated[8;9]. Two studies of siblings with OA, the Rotterdam sibling pairs and the Genetics osteoarthritis and Progression study, showed high familial aggregation estimates for involvement of OA at multiple joint sites[10;11]. A sumscore of affected joint groups showed a heritable estimate of 78% in sibling pairs from the Rotterdam study, a population-based cohort[10]. In addition, probands and siblings from the GARP study, selected on OA in any combination of multiple joint sites, tended to be affected at the same combination of joint sites[11]. So far, OA linkage and association studies mainly focused on OA at a single joint location with definitions based on either radiographic or symptomatic criteria. These efforts have yielded many susceptibility loci and some responsible genes which highlights the complexity of OA[8;12]. Recently, a genome-wide association study of Japanese patients with hip OA showed a compelling association at 14q32 with a functional SNP (rs12885713) in the calmodulin 1 (*CALM1*) gene[13]. In this study, it was shown that the associated T allele decreased *CALM1* transcription *in vitro* and *in vivo*[13]. The association of the functional SNP was not observed in Caucasian sibling pairs with hip replacement and showed a completely different allele frequency, indicating ethnic differences between Japanese and Caucasian individuals[14]. Other positive findings require validation in other OA studies of similar joint specific phenotypes and account only for a small proportion of the genetic component in OA susceptibility.

In the present study of middle aged sibling pairs affected with OA at multiple joint sites, we focus on identification of susceptibility loci and genes for a generalised or systemic OA phenotype which received little attention so far. We performed a genome-wide scan in 183 well-documented Dutch sibling pairs from the GARP study predominantly affected with symptomatic OA at multiple joint sites simultaneously.

Subjects and methods

The GARP study

The ongoing GARP study, which consists of Caucasian sibling pairs of Dutch origin affected predominantly by symptomatic OA at multiple sites, is aimed at identifying determinants of OA susceptibility and progression[11]. Probands (ages 40-70 years) and their siblings have OA at multiple joint sites of the hand or in two or more of the following joint sites (hand, spine (cervical or lumbar), knee or hip[11].

Subjects with symptomatic OA (as defined below) in just one joint site were required to have structural abnormalities in at least one other joint site defined by the presence of ROA in any of the four joints or the presence of two or more Heberden's nodes, Bouchard's nodes, or squaring of at least one first carpometacarpal (CMC1) joint on physical examination. Symptomatic OA in the knee and hip was defined according to the American College of Rheumatology (ACR) recommendations for knee and hip OA[15;16]. Knee OA was defined as pain or stiffness for most days of the preceding month and osteophytes at the joint margins of the tibiofemoral joint (*x* ray spurs).

Hip OA was defined as pain or stiffness in the groin and hip region on most days of the preceding month in addition to femoral or acetabular osteophytes or axial joint space narrowing on radiography. Prosthetic joints in the hips or knees as a result of end stage OA were defined as OA in that particular joint. Spine OA (cervical and lumbar) was defined as pain or stiffness in the spine on most days of the preceding month, in addition to a Kellgren/Lawrence score of two in at least one disc or one apophyseal joint.

OA in hand joints was defined according to the ACR criteria[17] as pain or stiffness on most days of the preceding month in addition to three of the following four criteria: bony swelling of two or more of the ten selected joints (bilateral distal interphalangeal (DIP) joints 2+3, bilateral proximal interphalangeal (PIP) joints 2+3, and CMC1 joints), bony swelling of two or more DIP joints, fewer than three swollen metacarpalphalangeal (MCP) joints, and deformity of at least one of the ten selected joints.

Intrareader variability for the different joint sites, scored by the Kellgren/Lawrence method, was assessed as follows: the intraclass correlation coefficient (ICC, with 95% confidence interval) was for the hands, 0.95 (0.92 to 0.96); for the knees (tibiofemoral), 0.92 (0.86 to 0.96); for the hips, 0.95 (0.92 to 0.98); for the cervical spine (apophyseal and disc), 0.71 (0.52 to 0.84); and for the lumbar spine (apophyseal and disc), 0.67 (0.46 to 0.81). Intrareader variability was based on an examination of 40 radiographs that were selected randomly throughout the duration of the study period and were blinded for any patient characteristics.

Table 1 Characteristics of GARP sibling pairs with symptomatic OA and subjects with ROA from the Rotterdam sample

Characteristics	The GARP study, symptomatic OA[1]	The Rotterdam study, ROA[2]
No. (frequency)	372 (1)	809 (1)
Age, mean ± SD	60.4 ± 7.6	60.3 ± 2.7
BMI, mean ± SD	27.0 ± 4.6*	26.3 ± 3.6
No. of women (frequency)	301 (0.81)*	478 (0.59)
Hand (frequency)	266 (0.72)	172 (0.21)
Hip (frequency)	91 (0.24)	71 (0.09)
Knee (frequency)	128 (0.34)	143 (0.18)
Spinal DD (frequency)	294 (0.79)	176 (0.23)

R(OA) = (radiographic) osteoarthritis; GARP = Genetics, osteoARthritis and Progression; BMI = body mass index, SD = standard deviation; DD = disc degeneration

[1] The GARP sibling pairs comprise 183 sibships: 179 siblings and four trios.

[2] In the association and haplotypes analysis, the GARP sibling pairs were compared with controls, the Rotterdam sample.

* $P \leq 0.05$ versus controls (the complete Rotterdam sample).

The Rotterdam study

The Rotterdam study, which comprises 7,983 Caucasian participants is a prospective, population-based cohort study of the determinants and prognosis of chronic diseases in the elderly[18]. The medical ethical committee of the Erasmus University Medical Centre approved the study, and informed consent was obtained from all subjects. In a random sample of unrelated subjects ages 55-65 years (n = 809), radiographs were scored for the presence of radiographic OA (ROA) of the knees, hips[19] , the hands and thoracocolumbar spine[10;19]. All radiographs were scored according to the Kellgren/Lawrence grading system (grades 0-4)[20] by two independent readers, blinded to all other data of the participant. After each set of about 150 radiographs the scores of the two readers were evaluated. Whenever the scores differed two or more points, or was two for one reader but one for the other, a consensus score was agreed upon. ROA of the knee was only assessed in the tibiofemoral joint. In the hands, a total of 36 separate joints were scored comprising eight joint groups: DIP joints, PIP joints, MCP joints, CMC1 joints, interphalangeal joint of the thumb, the trapezoscaphoideal joints, the radionavicular joints and the distal radioulnar joints.

By definition, ROA of the spine is confined to the apophyseal joints, but these joints could not be assessed on the lateral radiographs of the spine that were available. Instead, we assessed spinal disc degeneration (DD), according to a Kellgren/Lawrence scale[20], at three levels i.e. thoracic (Th4 to Th12), lumbar (L1 to L4 or L5) and lumbosacral (L5-S1 or L5-L6). Definite ROA at a particular joint site was defined as a Kellgren/Lawrence score of two or more[20]. As previously described[21], subjects with two or more of the following four criteria were considered as affected with generalised ROA: hand ROA in three or more hand joint groups (the right and left hands were considered separately), spinal DD in two or more disc levels, ROA in one or two knees and ROA in one or two hips.

In the association and haplotype analysis, we compared affected sibling pairs from the GARP study to the complete Rotterdam sample (n = 809) as reference group representing the general population. Both studies comprise Caucasian subjects from the western areas of the Netherlands with a mean age of 60.3 years and may represent the same genetic background.

Genotype measurements

Short tandem repeat polymorphisms

A complete genome-wide scan containing 403 microsatellite markers with an average spacing of 10 cM was performed in 187 pairs and four trios with OA at multiple joint sites from the GARP study. Markers and 14 additional microsatellite markers for fine mapping on chromosome 6, 10, 13 and 14 were taken from Human Linkage Set v2.5 MD10 or HD5 (Applied Biosystems), respectively and measured using an ABI Prism DNA Analyzer 3700 (Applied Biosystems). Genotyping was performed using standard conditions and reagentia with some exceptions. The amount of polymerase chain reaction (PCR) primer pairs for the markers was reduced up to 5-fold and duplex PCR reactions were designed if possible to reduce costs, time expense and amount of genomic DNA used. Genotypes were analysed by using Genemapper version 2.0 and 3.0 (Applied Biosystems).

As quality control, approximately 8% of the samples were genotyped in duplicate and compared. In addition, 48 additional family members from 36 different sibling pairs were genotyped to improve our ability to detect genotyping errors and estimate allele sharing. Mendelian errors were checked for Mendelian inconsistencies and unlikely recombinants using Merlin[22]. These quality checks indicated that marker D6S434 and D9S158 from the Human Linkage Set v2.5 MD10 could not be genotyped reliably in our hands due to unclear one base pair differences. Subjects and markers showed an average success rate of 96% (range 77-100%) and 96% (range 83-100%), respectively. Family relationships were verified using the GRR program[23]. Eight sibling pairs showed pedigree errors and were removed for further analysis. In seven of these sibling pairs, individuals reported to be full siblings were almost certainly half siblings. The remaining siblings were monozygotic twins. A locally developed SQL database was used to store genotypic data, compare repeated genotypes and generate output files for linkage analysis. The location of the markers was taken from an integrated genetic map with interpolated genetic map positions (http://www2.qimr.edu.au/DavidD/). The position is in Decode cM, estimated via locally weighted linear regression (lo(w)ess) from the Build 35.1 (and 34.3) physical map positions and published Decode and Marshfield genetic map positions.

Single nucleotide polymorphisms

Genomic DNA was isolated from blood samples. In total, 809 subjects (331 men, 478 women) from the Rotterdam study and 179 sibling pairs and four trios (71 men, 301 women) were genotyped for highly informative tagging SNPs capturing a large fraction of all genetic variation in the *CALM1*, *DIO2* and *FLRT2*. Tagging SNPs were selected from HapMap Public Release #19 applying the efficient multimarker method with $r^2 > 0.8$ and minor allele frequency (MAF) > 0.05 implemented at in two or more http://www.hapmap.org[24]. Tagging SNPs or their proxies were chosen to fit efficiently in a Sequenom multiplex assay. We genotyped the following tagging SNPs in *CALM1*: rs3814847, rs3814845, rs2300496, rs2300502 and rs5871. For *DIO2*, we measured rs225014, rs2267872, rs225011, rs12885300 and rs10136454. For *FLRT2*, rs2239576, rs2057311, rs17121375, rs17646457 and rs1129671 were genotyped. Multiplex genotyping assays were designed using Assay Designer software (Sequenom, San Diego, CA). Tagging SNPs were genotyped by mass spectrometry (the homogeneous MassARRAY system; Sequenom, San Diego, CA) using standard conditions. PCR reactions were carried out in a final volume of 5 µl and contained standard reagents and 2.5 ng of genomic DNA. Genotypes were assigned by using Genotyper version 3.0 software (Sequenom, San Diego, CA). The functional SNP rs12885713, located in *CALM1*, was genotyped using a Taqman by design assay and an ABI Prism DNA Analyzer 7900 (Applied Biosystems) with standard conditions. In addition, genotype distributions of all SNPs were in agreement with Hardy-Weinberg equilibrium and approximately 8% of the subjects were genotyped twice and checked.

Statistical analysis

Linkage analysis

Nonparametric linkage analysis was carried out by use of the S_{all} statistics[25] implemented in the software Merlin[22]. LOD scores[26] were plotted on a common 1 cM grid. For X-chromosome analyses, we used MINX (Merlin-In-X), a modified version of Merlin. Using Merlin, we calculated the genome-wide significance level by use of 1000 gene dropping simulations. In each simulation, we generated a random dataset with the same marker allele frequencies, marker spacings and missing data patterns and preserved the original phenotypes. To evaluate the false positive rate, we analysed these simulations and calculated the probability of observing a linkage peak higher than our maximum LOD score. We calculated the contribution of each family to the maximum LOD score with Merlin. To analyse whether families affected with similar joint sites contributed to the maximum LOD score, we calculated Pearson's correlation coefficient between the LOD score and hip, knee or hand OA families.

Association approaches

Two approaches were used to investigate whether tagging SNPs explained the linkage signal. In the first approach, joint linkage and association analysis was performed to identify SNPs that fully or partly explain the observed linkage signal using the program LAMP (http://csg.sph.umich.edu/LAMP)[27]. The maximum likelihood was estimated using 50 starting points and the disease prevalence was set at 0.01. A drawback of this method is that it models one associated locus assuming that it is the only locus under the peak.

In the second more robust approach, we stratified all sibling pairs for identical by descent (IBD) status and compared each stratum with subjects from the total Rotterdam sample as reference group. To take into account uncertainty in IBD status, allele, haplotype and genotype frequencies in sibling pairs were weighted for the probability to share two, one or zero alleles IBD. The IBD status was estimated for the location of each gene separately with the genotypes of microsatellite markers only using the IBD and grid (1 cM) option as implemented in Merlin[22]. For differences in allele or haplotype frequencies between sibling pairs from the GARP study and the Rotterdam sample, standard errors were estimated from the variance between sibling pairs (robust standard errors)[28]. For calculating haplotype frequencies and their effects in the sibling pairs, posterior haplotype probabilities were estimated in the siblings and the Rotterdam sample simultaneously. The expectation maximisation algorithm implemented in SNPHAP (SNPHAP version 1.3, http://www-gene.cimr.cam.ac.uk/clayton/software/) was used to assign the most likely haplotype pairs to individuals. We used the missing data option that allows us to carry out haplotype analysis in subjects with some missing genotype measurements.

A logistic regression model was fitted to measure the strength of association, which is expressed as odds ratio (OR) with 95% confidence intervals (CI) adjusted for age (years), body mass index (BMI, kg/m^2), and sex. The strength of these effects in all sibling pairs as compared with the Rotterdam sample as reference group was determined with logistic regression with robust standard errors to adjust for family relationship[28]. Instead of adjusting *P* values a priori for multiple testing, exact *P* values are provided in order to allow the reader to interpret the level of significance. We performed robust standard error analyses using Stata SE8 software (Stata Corporation, USA). All other analyses were carried out with SPSS version 11 software (SPSS, Chicago, Illinois, USA).

Results

Linkage analysis

The characteristics of the GARP sibling pairs are shown in Table 1. Initial nonparametric linkage analysis provided several suggestive linkage signals on chromosome 6, 10, 13, 14 (Table 2, Figure 1). The highest linkage signal was observed on chromosome 13q12 with a LOD score of 2.23 (P = 0.0007). Typing 14 additional markers in the four areas with suggestive linkage, evidence on chromosome 6, 10 and 13 weakened. In contrast, the LOD score on chromosome 14 increased from 1.32 (P = 0.007) to 2.37 (P = 0.0005) with exact location on 14q32.11 (Figure 2). The one LOD-drop interval of this signal encompasses 26 cM (73-97 cM) containing the markers D14S74, D14S1037, D14S1044 and D14S280. However, the maximum obtained LOD score of 2.37 did not reach genome-wide significance as we observed in 133/1000 simulations a LOD score of 2.37 or higher (P = 0.13). A positive LOD score was observed for 90 sibling pairs out of 183 at this locus (86-87 cM). Among these pairs, we observed no significant correlation between the LOD score per family and any combination of affected joint sites.

Notably, the location of the linkage peak coincided with the calmodulin 1 (*CALM1)* gene, previously associated with symptomatic hip OA in the Japanese population[13]. Upon a search in the public genome resources, two other attractive candidate genes were located in the linkage area. Fibronectin leucine rich transmembrane protein 2 (*FLRT2*) encodes a small leucine-rich proteoglycan found in the extracellular matrix[29]. *DIO2* encodes iodothyronine deiodinase Type II (D2), a selenoprotein that catalyzes the 5-prime deiodination of thyroxine (T4) to generate an active thyroid hormone, 3,3-prime,5-triiodothyronine (T3). D2 is an important regulator of endochondral bone formation in epiphyseal growth plates[30-33]. To examine whether the possible contribution of genetic variation in these genes is underlying the observed linkage signal, tagging SNPs covering the haplotype blocks were genotyped.

Table 2 lists all chromosomal regions yielding a LOD score > 1.0 in non-parametric analyses

Chrom	Nearest Marker	Position Marker (cM)	Info[1]	Interval (cM)[2]	Position (cM)	Max LOD	*P*
6	D6S1574	14	0.56	0-43	0	1.13	0.011
10	D10S196	69	0.33	48-84	69	1.36	0.006
13	D13S175	1	0.38	0-12	0	2.23	0.0007
14	D14S74;D14S280	76;92	0.39;0.49	32-117	83	1.32	0.007

Chrom = chromosome; cM = centiMorgan; LOD = log of odds; info= informativity
[1] Informativity at the position of the marker.
[2] One LOD-drop interval.

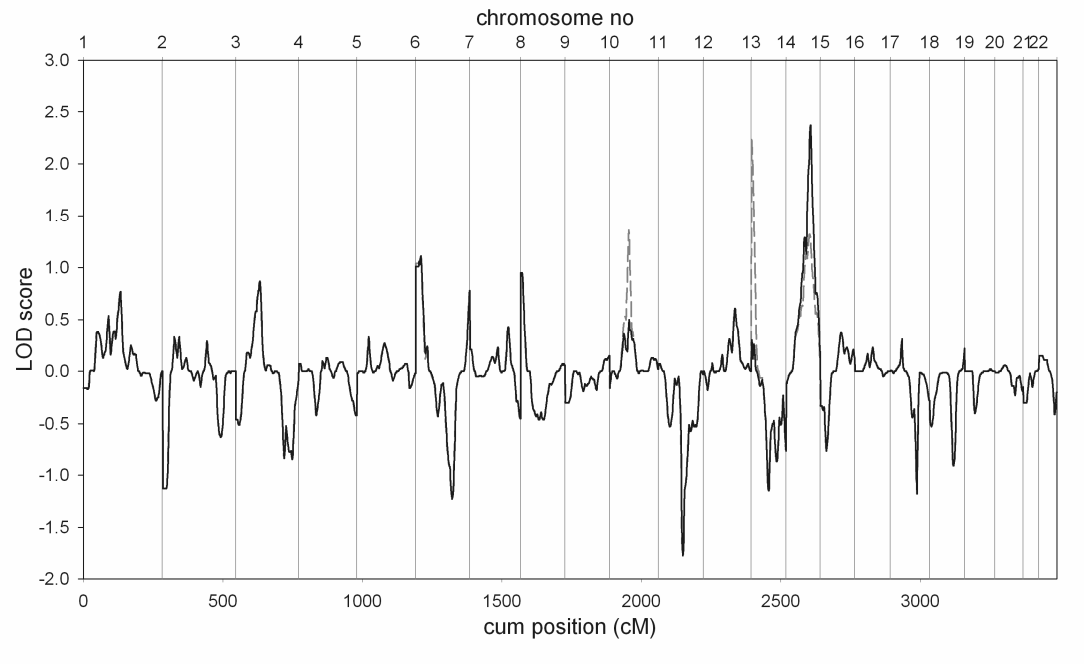

Figure 1 Results from the genome-wide scan in 183 sibling pairs from the GARP study
Dashed line represents initial results (mean informativity = 0.41; range, 0.08-0.63); solid line represents results upon finemapping (mean informativity = 0.42; range, 0.08-0.69).

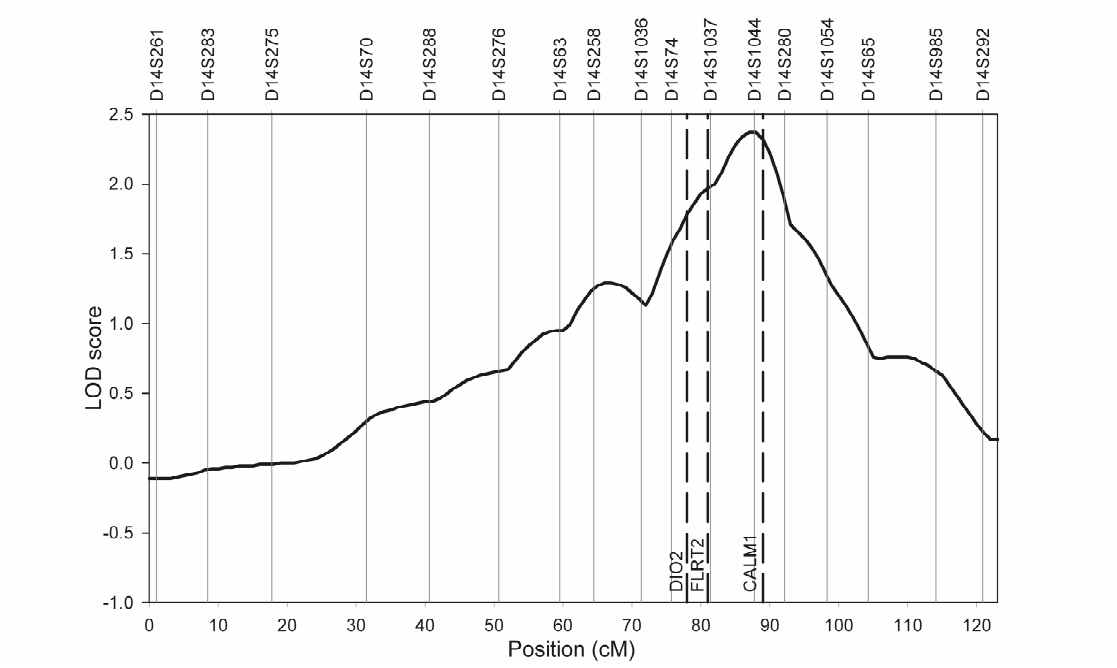

Figure 2 Suggestive evidence for linkage on chromosome 14q32.11
Dashed lines represent the genes *DIO2* (78 cM; informativity =0.51), *FLRT2* (82 cM; informativity = 0.54) and *CALM1* (89 cM; informativity = 0.47) located in the one LOD-drop interval (73-97 cM; mean informativity = 0.50 (range, 0.46-0.56)).

Joint modelling and association with LAMP

Joint modelling of linkage and association revealed no evidence for association to individual SNPs or haplotypes in *CALM1* (including the functional SNP rs12885713) or *FLRT2* with the putative disease locus (Table 3). However, we observed significant associations (P = 0.012) with the C allele of *DIO2* rs225014, the C allele of *DIO2* rs12885300 (P = 0.05) and a borderline significant association with the C allele of *DIO2* rs225011 (P = 0.08) and the putative disease locus (Table 3). There was no linkage disequilibrium ($r^2 < 0.01$) between this SNP and the putative disease locus indicating that the allele is unlikely to be causal as calculated by LAMP. These tagging SNPs (rs12885300, rs2267872, rs225011 and rs225014) in *DIO2* showed high linkage disequilibrium and comprised six common haplotypes (>1%). The common CGCC haplotype, encompassing the associated C alleles, showed a significant association (P = 0.04).

Association analysis stratified for IBD status

Alternatively, allele and haplotype frequencies in sibling pairs sharing two alleles IBD, indicating the subjects that contribute to the linkage at the *DIO2* locus were compared with subjects from the total Rotterdam sample as reference group (n = 809) to confirm the observed *DIO2* associations as observed by the LAMP program. The frequency of the C allele of the rs225014 (P = 0.02) and the corresponding haplotype CGCC (P = 0.05) were again significantly increased with this approach (Table 4). Among carriers of at least one risk allele, the risk of having OA at multiple joint sites, adjusted for age, BMI and sex and family relationship was 1.5 (95% CI 1.1-2.1, P = 0.018). The effect sizes did not alter significantly when these covariables (age, BMI and sex) were removed. Allele, genotype and haplotype frequencies were identical between unstratified sibling pairs and the random population.
In addition, we investigated whether genetic variation in *DIO2* increases radiographic OA susceptibility at multiple joint sites in the Rotterdam sample in the way we previously analysed other loci for their contribution to generalised ROA[21]. Allele frequencies of subjects with ROA at multiple joint sites (generalised ROA) did not differ from the complete population-based Rotterdam sample excluding those with ROA at multiple joint sites (data not shown).

Table 4 Allele frequencies of *DIO2* SNPs in GARP sibling pairs stratified for IBD status and unrelated subjects from the Rotterdam sample

		GARP sibling pairs			Rotterdam sample	Carrier risk[3,4]	
SNP/ haplotype	Allele	IBD = 0 MAF[1]	IBD = 1 MAF[1]	IBD = 2 MAF[1]	MAF[2]	OR (95% CI)	*P*
rs12885300	C>T	0.34 (94/48)	0.38 (220/134)	0.33 (157/77)	0.35 (968/530)	0.9 (0.6-1.3)	0.46
rs2267872	G>A	0.11 (127/15)	0.11 (308/37)	0.07 (216/16)	0.09 (1,349/137)	0.7 (0.4-1.4)	0.32
rs225011	T>C	0.44 (80/62)	0.40 (205/137)	0.48 (120/110)	0.42 (848/624)	1.3 (0.9-1.8)	0.20
rs225014	T>C	0.35 (93/49)	0.32 (237/111)	0.44 (130/102)*	0.35 (965/519)	1.5 (1.1-2.1)	0.018
rs10136454	C>T	0.007 (142/1)	0.014 (353/5)	0.029 (233/7)	0.016 (1,448/24)	2.1 (0.7-6.2)	0.18
haplotype	TGTT[5]	0.33 (95/47)	0.35 (233/124)	0.30 (167/73)	0.34 (1,069/549)	0.8 (0.6-1.2)	0.34
haplotype	CGCC[5]	0.33 (95/47)	0.28 (259/98)	0.40 (145/95)	0.32 (1,097/521)	1.5 (1.0-2.1)	0.05
haplotype	CGTT[5]	0.22 (111/31)	0.23 (275/82)	0.18 (196/44)	0.21 (1,275/343)	0.8 (0.6-1.3)	0.40
haplotype	CACT[5]	0.11 (127/15)	0.10 (320/37)	0.067 (224/16)	0.09 (1,480/138)	0.8 (0.4-1.4)	0.36
haplotype	TGCC[5]	0.007 (141/1)	0.025 (348/9)	0.021 (235/5)	0.014 (1,595/23)	1.2 (0.4-3.6)	0.80
haplotype	CGTC[5]	0.007 (141/1)	0.011 (236/4)	0.021 (235/5)	0.017 (1,591/27)	1.5 (0.4-5.1)	0.54

SNP = single nucleotide polymorphism; MAF = minor allele frequency; IBD = identical by descent; OR = odds ratio

[1] Minor allele frequency stratified for IBD status (sum of IBD probability of common alleles/ sum of IBD probability of minor alleles).

[2] Minor allele frequency (number of common alleles/number of minor alleles).

[3] Siblings conferring at least one minor allele and stratified for IBD 2 status, versus the complete Rotterdam sample. OR and *P* values were adjusted for age, BMI, sex and family relationship.

[4] OR for carriers of particular haplotype versus the complete Rotterdam sample, excluding those with that particular haplotype.

[5] Alleles in the following order: rs12885300, rs2267872, rs225011, rs225014.

* $P \leq 0.05$, allele frequency among siblings stratified for IBD 2 status and adjusted for family relationship versus the complete Rotterdam sample.

Table 3 Combined linkage and association analysis in the GARP sibling pairs

Gene	SNP reference	allele	MAF[1]	*P*[2]
DIO2	rs12885300	C>T	0.36 (472/260) *	0.05
	rs2267872	G>A	0.09 (653/67)	0.20
	rs225011	T>C	0.43 (406/310)	0.08
	rs225014	T>C	0.36 (461/263) *	0.012
	rs10136454	C>T	0.02 (730/12)	0.40
	haplotype[3]	CGCC	0.33 (500/242) *	0.04
FLRT2	rs2057311	G>C	0.42 (423/303)	1.0
	rs17121375	A>G	0.02 (718/14)	0.6
	rs17646457	G>A	0.14 (627/103)	0.6
	rs2239576	T>C	0.14 (628/98)	1.0
	rs1129671	C>T	0.28 (532/208)	0.6
CALM1	rs3814847	C>G	0.34 (471/243)	0.4
	rs12885713	T>C	0.43 (398/298)	0.4
	rs2300496	A>C	0.43 (418/312)	0.3
	rs2300502	G>A	0.11 (653/81)	0.4
	rs5871	T>C	0.08 (674/60)	0.3
	rs3814845	C>G	0.07 (678/52)	0.7

GARP = Genetics, osteoARthritis and Progression, MAF = minor allele frequency
[1] Minor allele frequency (number of common alleles/ number of minor alleles) in GARP sibling pairs.
[2] *P* values of the combined linkage and association analyses using the program LAMP.
[3] Alleles in the following order: rs12885300, rs2267872, rs225011, rs225014.
* $P \leq 0.05$

Discussion

Here, we report the results of a nonparametric linkage analysis of a genome-wide scan of sibling pairs with OA at multiple joint sites (GARP study) from whom detailed clinical and radiographic data had been collected. Suggestive evidence for linkage was observed on 14q32.11 with a maximal LOD score of 2.37 ($P = 0.0005$; genome-wide $P = 0.13$). This signal was mainly attributable to OA at different combinations of multiple joint sites indicating that systemic genetic factors may account for the generalised disease phenotype selected for in the GARP study.

This locus contains the positional candidate genes, *DIO2*, *FLRT2* and *CALM1*. Here, by two combined linkage and association approaches, we did not find evidence that the linkage signal could be attributed to genetic variation in *CALM1* or *FLRT2* using tagging SNPs or the functional variant (rs12885713) in *CALM1* that was previously associated with hip OA in Japanese patients[13]. However, the C alleles of *DIO2* SNPs rs225014, rs225011 and rs12885300 were emerged from the combined linkage and association analysis as the corresponding common CGCC haplotype in patients contributing to the linkage. In addition, carriers of the C allele of rs225014 conferred a

relative risk of 1.5 (1.1-2.1, *P* =0.02) to have OA at multiple joint sites as compared to the Rotterdam sample representing the random population. Given the low prevalence of OA at multiple joint sites and the high frequency of the associated risk alleles (0.36, 0.43 and 0.36), this variant may be not causal itself. The true causal variant may have a lower frequency and be in LD with this variant. Since other SNPs of the HapMap are not available to further delineate these haplotypes, resequencing of *DIO2* may be required to detect the causal variant itself.

DIO2 encodes a selenoenzyme D2 that catalyzes the conversion of thyroxin (T4) to triiodothyronine (T3) via 5-prime-deiodination and regulates the local thyroid hormone bioactivity in the growth plate[33;34]. T3 inhibits chondrocyte proliferation but stimulates chondrocyte differentiation and matrix synthesis[35]. Genetic variation in *DIO2* may therefore have functional consequences for deiodinase activity but also for circulating iodothyronine levels. Previously, the T allele of rs12885300 but not the C allele of rs225014 was associated with lower levels of plasma T4, free T4 and rT3 and a higher T3/T4 and T3/rT3 ratio indicating higher D2 activity[36;37]. In contrast, the C allele of rs12885300 encompassing our associated haplotype, may therefore have lower D2 activity. Indeed, lower expression of D2 in the growth plate contributes to the pathogenesis of tibial chondrodysplasia in chicken[38].

The linkage region found in our study overlapped with the location of the *CALM1* gene previously associated with hip OA in two independent studies of Japanese patients[13]. A functional assay in which the associated allele modulated chondrogenic activity confirmed the relevance of the promoter SNP (rs12885713). In our studies, genetic variation including the functional promoter SNP in the *CALM1* gene was not associated with OA at multiple joint sites. Lack of association of the functional promoter SNP was also observed for hip replacement in the Caucasian population. Possible ethnic differences are illustrated by the difference of frequency of the promoter SNP in Japanese versus Caucasian subjects[39]. A combined linkage association approach might have limited power for complex diseases as OA when multiple genetic variants with small effect sizes are involved[27]. Therefore, we can not exclude that other (rare) variants in the *CALM1* gene contribute to OA in GARP.

Previous associations of a functional *FRZB* variant with hip replacement[40] were confirmed for OA at multiple joint sites[21], indicating that similar pathways may be involved in advanced hip OA and OA at multiple joint sites. However, the extent of overlapping aetiology of these advanced OA phenotypes remains to be solved. From the simulations, our linkage signal of 2.37 was not genome-wide significant. However, thresholds for genome-wide significance have been debated extensively and range between 3.0 and 3.6[41-42]. Obviously, there is limited statistical power for mapping a complex trait with 183 informative sibling pairs. Additional data including LD mapping and concordance with previous published studies may support discriminating true from false positive results.

In the association analysis, we stratified for IBD status and therefore the size of the case group (i.e. sibling pairs contributing to the linkage result) declines. In addition, we mainly included affected sibling pairs and only several additional sibling or parents. Consequently, the genetic information for estimating the IBD probabilities is restricted[43]. The lack of association of tagging SNPs in the *CALM1* gene in sibling pairs sharing two alleles IBD may therefore be explained by low information content at this locus. This limitation will probably be addressed in the future through an increase of genetic information via a dense panel of SNP markers.

In our OA sibling pair study, primarily selected for OA at multiple joint sites, a major locus on chromosome 14q32.11 was identified which coincided with associations of common SNPs in *DIO2*. Our data suggest that common variants within the *DIO2* gene predispose to OA at multiple joint sites.

Acknowledgements

The Leiden University Medical Centre, the Dutch Arthritis Association and Pfizer Inc., Groton, CT, USA support the GARP study. We further thank for the GARP study the support of the cooperating hospitals and referring rheumatologists, orthopaedic surgeons and general practitioners and for the Rotterdam study the support of general practitioners and pharmacists of the Ommoord district in Rotterdam. The Rotterdam Study is supported by the Erasmus Medical Centre and Erasmus University Rotterdam, the Netherlands Organization for Scientific Research, the Netherlands Organization for Health Research and Development, the Research Institute for Diseases in the Elderly, the Ministry of Education, Culture and Science, the Ministry of Health, Welfare and Sports, the European Commission (DG XII), and the Municipality of Rotterdam. This work was supported by the Netherlands Organization of Scientific Research (MW 904-61-095), Leiden University Medical Centre and the Centre of Medical System Biology. Finally, we thank Ruud van der Breggen, Dennis Kremer and Eka Suchiman for genotyping and Hein Putter for performing simulations.

References

1. Kellgren JH, Moore R. Generalized osteoarthritis and Heberden's nodes. Br Med J 1952; 1(4751):181-187.
2. Kellgren JH, Lawrence JS, Bier f. Genetic factors in generalized osteo-arthrosis. Ann Rheum Dis 1963; 22:237-255.

3. Chaisson CE, Zhang Y, McAlindon TE, Hannan MT, Aliabadi P, Naimark A et al. Radiographic hand osteoarthritis: incidence, patterns, and influence of pre-existing disease in a population based sample. J Rheumatol 1997; 24(7):1337-1343.
4. Cooper C, Egger P, Coggon D, Hart DJ, Masud T, Cicuttini F et al. Generalized osteoarthritis in women: pattern of joint involvement and approaches to definition for epidemiological studies. J Rheumatol 1996; 23(11):1938-1942.
5. Egger P, Cooper C, Hart DJ, Doyle DV, Coggon D, Spector TD. Patterns of joint involvement in osteoarthritis of the hand: the Chingford Study. J Rheumatol 1995; 22(8):1509-1513.
6. Gunther KP, Sturmer T, Sauerland S, Zeissig I, Sun Y, Kessler S et al. Prevalence of generalised osteoarthritis in patients with advanced hip and knee osteoarthritis: the Ulm Osteoarthritis Study. Ann Rheum Dis 1998; 57(12):717-723.
7. Vignon E. Hand osteoarthritis and generalized osteoarthritis: a need for clarification. Osteoarthritis Cartilage 2000; 8 Suppl A:S22-S24.
8. Loughlin J. Genetic epidemiology of primary osteoarthritis. Curr Opin Rheumatol 2001; 13(2):111-116.
9. Spector TD, MacGregor AJ. Risk factors for osteoarthritis: genetics. Osteoarthritis Cartilage 2004; 12 Suppl A:S39-S44.
10. Bijkerk C, Houwing-Duistermaat JJ, Valkenburg HA, Meulenbelt I, Hofman A, Breedveld FC et al. Heritabilities of radiologic osteoarthritis in peripheral joints and of disc degeneration of the spine. Arthritis Rheum 1999; 42(8):1729-1735.
11. Riyazi N, Meulenbelt I, Kroon HM, Ronday KH, Hellio le Graverand MP, Rosendaal FR et al. Evidence for familial aggregation of hand, hip, and spine osteoarthritis (OA), but not knee OA in siblings with OA at multiple sites. The GARP study. Ann Rheum Dis 2005; 64(3):438-443.
12. Peach CA, Carr AJ, Loughlin J. Recent advances in the genetic investigation of osteoarthritis. Trends Mol Med 2005; 11(4):186-191.
13. Mototani H, Mabuchi A, Saito S, Fujioka M, Iida A, Takatori Y et al. A functional single nucleotide polymorphism in the core promoter region of CALM1 is associated with hip osteoarthritis in Japanese. Hum Mol Genet 2005; 14(8):1009-1017.
14. Loughlin J, Sinsheimer JS, Carr A, Chapman K. The CALM1 core promoter polymorphism is not associated with hip osteoarthritis in a United Kingdom Caucasian population. Osteoarthritis Cartilage 2005.
15. 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 1986; 29(8):1039-1049.
16. Altman R, Alarcon G, Appelrouth D, Bloch D, Borenstein D, Brandt K et al. The American College of Rheumatology criteria for the classification and reporting of osteoarthritis of the hip. Arthritis Rheum 1991; 34(5):505-514.
17. Altman R, Alarcon G, Appelrouth D, Bloch D, Borenstein D, Brandt K et al. The American College of Rheumatology criteria for the classification and reporting of osteoarthritis of the hand. Arthritis Rheum 1990; 33(11):1601-1610.
18. Hofman A, Grobbee DE, de Jong PT, van den Ouweland FA. Determinants of disease and disability in the elderly: the Rotterdam Elderly Study. Eur J Epidemiol 1991; 7(4):403-422.
19. Odding E, Valkenburg HA, Stam HJ, Hofman A. Determinants of locomotor disability in people aged 55 years and over: the Rotterdam Study. Eur J Epidemiol 2001; 17(11):1033-1041.
20. Kellgren JH, Jeffrey MR, Ball J. The epidemiology of chronic rheumatism. Volume II: Atlas of standard radiographs of arthritis. Oxford: Blackwell Scientific Publications, 1963.
21. Min JL, Meulenbelt I, Riyazi N, Kloppenburg M, Houwing-Duistermaat JJ, Seymour AB et al. Association of the Frizzled-related protein gene with symptomatic osteoarthritis at multiple sites. Arthritis Rheum 2005; 52(4):1077-1080.
22. Abecasis GR, Cherny SS, Cookson WO, Cardon LR. Merlin--rapid analysis of dense genetic maps using sparse gene flow trees. Nat Genet 2002; 30(1):97-101.

23. Abecasis GR, Cherny SS, Cookson WO, Cardon LR. GRR: graphical representation of relationship errors. Bioinformatics 2001; 17(8):742-743.
24. de Bakker PI, Yelensky R, Pe'er I, Gabriel SB, Daly MJ, Altshuler D. Efficiency and power in genetic association studies. Nat Genet 2005; 37(11):1217-1223.
25. Whittemore AS, Halpern J. A class of tests for linkage using affected pedigree members. Biometrics 1994; 50(1):118-127.
26. Kong A, Cox NJ. Allele-sharing models: LOD scores and accurate linkage tests. Am J Hum Genet 1997; 61(5):1179-1188.
27. Li M, Boehnke M, Abecasis GR. Joint modeling of linkage and association: identifying SNPs responsible for a linkage signal. Am J Hum Genet 2005; 76(6):934-949.
28. Diggle P.J., Liang K.Y., Zeger S.L. Analysis of longitudinal data. Oxford University Press, 1994.
29. Lacy SE, Bonnemann CG, Buzney EA, Kunkel LM. Identification of FLRT1, FLRT2, and FLRT3: a novel family of transmembrane leucine-rich repeat proteins. Genomics 1999; 62(3):417-426.
30. Kohrle J. Local activation and inactivation of thyroid hormones: the deiodinase family. Mol Cell Endocrinol 1999; 151(1-2):103-119.
31. Miura M, Tanaka K, Komatsu Y, Suda M, Yasoda A, Sakuma Y et al. Thyroid hormones promote chondrocyte differentiation in mouse ATDC5 cells and stimulate endochondral ossification in fetal mouse tibias through iodothyronine deiodinases in the growth plate. J Bone Miner Res 2002; 17(3):443-454.
32. Robson H, Siebler T, Stevens DA, Shalet SM, Williams GR. Thyroid hormone acts directly on growth plate chondrocytes to promote hypertrophic differentiation and inhibit clonal expansion and cell proliferation. Endocrinology 2000; 141(10):3887-3897.
33. Bianco AC, Salvatore D, Gereben B, Berry MJ, Larsen PR. Biochemistry, cellular and molecular biology, and physiological roles of the iodothyronine selenodeiodinases. Endocr Rev 2002; 23(1):38-89.
34. Salvatore D, Bartha T, Harney JW, Larsen PR. Molecular biological and biochemical characterization of the human type 2 selenodeiodinase. Endocrinology 1996; 137(8):3308-3315.
35. Bassett JH, Williams GR. The molecular actions of thyroid hormone in bone. Trends Endocrinol Metab 2003; 14(8):356-364.
36. Peeters RP, van den Beld AW, Attalki H, Toor H, de Rijke YB, Kuiper GG et al. A new polymorphism in the type II deiodinase gene is associated with circulating thyroid hormone parameters. Am J Physiol Endocrinol Metab 2005; 289(1):E75-E81.
37. Peeters RP, van Toor H, Klootwijk W, de Rijke YB, Kuiper GG, Uitterlinden AG et al. Polymorphisms in thyroid hormone pathway genes are associated with plasma TSH and iodothyronine levels in healthy subjects. J Clin Endocrinol Metab 2003; 88(6):2880-2888.
38. Shen S, Berry W, Jaques S, Pillai S, Zhu J. Differential expression of iodothyronine deiodinase type 2 in growth plates of chickens divergently selected for incidence of tibial dyschondroplasia. Anim Genet 2004; 35(2):114-118.
39. Loughlin J, Sinsheimer JS, Carr A, Chapman K. The CALM1 core promoter polymorphism is not associated with hip osteoarthritis in a United Kingdom Caucasian population. Osteoarthritis Cartilage 2005.
40. Loughlin J, Dowling B, Chapman K, Marcelline L, Mustafa Z, Southam L et al. Functional variants within the secreted frizzled-related protein 3 gene are associated with hip osteoarthritis in females. Proc Natl Acad Sci U S A 2004; 101(26):9757-9762.
41. Lander E, Kruglyak L. Genetic dissection of complex traits: guidelines for interpreting and reporting linkage results. Nat Genet 1995; 11(3):241-247.
42. Morton NE. Significance levels in complex inheritance. Am J Hum Genet 1998; 62(3):690-697.
43. Evans DM, Cardon LR. Guidelines for genotyping in genomewide linkage studies: single-nucleotide-polymorphism maps versus microsatellite maps. Am J Hum Genet 2004; 75(4):687-692.