The additive effect of individual genes in predicting risk of knee osteoarthritis

A M Valdes, M Doherty and T D Spector

Ann Rheum Dis 2008;67:124-127; originally published online 17 Aug 2007; doi:10.1136/ard.2007.075838

Updated information and services can be found at:
http://ard.bmj.com/cgi/content/full/67/1/124

These include:

References
This article cites 13 articles, 3 of which can be accessed free at:
http://ard.bmj.com/cgi/content/full/67/1/124#BIBL

Rapid responses
You can respond to this article at:
http://ard.bmj.com/cgi/eletter-submit/67/1/124

Email alerting service
Receive free email alerts when new articles cite this article - sign up in the box at the top right corner of the article

Notes

To order reprints of this article go to:
http://journals.bmj.com/cgi/reprintform

To subscribe to Annals of the Rheumatic Diseases go to:
http://journals.bmj.com/subscriptions/
The additive effect of individual genes in predicting risk of knee osteoarthritis

A M Valdes, M Doherty, T D Spector

ABSTRACT

Objective: Genetic factors are important determinants of osteoarthritis (OA) but most individual genetic associations appear relatively modest. We aimed to answer whether carrying several genetic variants associated with knee OA could result in a greater risk of OA.

Methods: Genotypes for 36 single nucleotide polymorphisms (SNPs) in 17 candidate genes previously associated with OA were analysed in 298 men and 305 women diagnosed with knee OA who met American College of Rheumatology (ACR) criteria, and in 297 male and 299 female age- and ethnicity-matched controls. The S-sum statistic method was used to select SNPs that contributed to knee OA, separately for men and women, and the coefficients from a logistic regression were used to add the genotypes in a new genetic risk variable.

Results: The odds ratio for individuals in the top quartile of the “genetic risk” variable compared to those in the bottom quartile was found to be 8.68 (95% CI 5.20–14.49, p < 2 × 10^{-14}) for women and 5.06 (95% CI 3.10–8.27, p < 1 × 10^{-10}) for men.

Conclusions: Our data suggest that the additive information from a number of genetic variants can predict a substantial proportion of risk of knee OA.

Osteoarthritis (OA) of the knee is a common complex disorder resulting in joint disability with known constitutional and environmental risk factors for development and progression, such as age, obesity, hormonal status, bone density, physical activity and past history of trauma. Knee OA, also has an important genetic component, and several studies have investigated the role of candidate genes in the risk of hip and knee OA. Several genes with common polymorphisms consistently affecting risk of OA have been reported to date. However, the genetic variants involved are not mutants with large attributable risks. Rather, the increased risks for carrying a predisposing genetic variant appear to be fairly modest, with most of them having odds ratios between 1.3 and 2.0. If an individual carries risk variants at several genes, does his/her risk of OA increase in proportion, or do genetic risks remain modest when compared with environmental/constitutional risk factors such as obesity and past history of trauma?

To answer this question we have combined the previously published genotypes of 36 variants on 17 genes in a multi-centre case–control study of knee OA. Variation at all the genes included in this study have been reported to be associated with risk of knee or hip OA in at least one independent study. We assessed the risk of OA for individuals carrying several gene variants associated with disease susceptibility compared to individuals who carried few or none. Because several of the genes included (eg, FRZB, BMP2) have been reported to be associated only or predominantly in women, and the prevalence of OA is higher in females, all analyses have been performed separately in each gender.

SUBJECTS AND METHODS

Subjects for this study have been described elsewhere. Briefly, 603 knee OA cases (298 men and 305 women) were recruited in Nottingham, UK. Osteoarthritis was assessed both clinically and radiographically, patients had standardised extended weight-bearing anteroposterior radiographs and met American College of Rheumatology (ACR) criteria. In addition, 596 Caucasian age-matched controls aged 50–80 without clinical signs or symptoms typical of osteoarthritis were recruited from two centres: Nottingham (111 women and 50 men) and Oxford (185 women and 250 men). No radiographs were performed on controls. Controls were only characterised by clinical criteria. The mean age was: female cases, 73.5 years (SD = 7.16); female controls, 72.1 (SD = 8.5); male cases, 72.1 (SD = 6.9); and male controls, 71.0 years (SD = 7.8). All study subjects gave informed consent to participate and the Oxford and Nottingham Research Ethics Committees approved the protocol.

Genotypes

The genotyping data included in this study have been previously reported. For each SNP we identified the allele that was more common among cases than among controls and assigned a value of 1 to the homozygote carrying two copies of such allele, 0.5 to the heterozygote and 0 for the lower risk homozygote. The genes and SNPs included are shown in table 1.

Set association

This method is used to find a set of SNPs that are jointly associated with disease by computing an association statistic for each SNP ie, a χ² for a 2×3 table, where the two rows correspond to cases and controls, and the three columns refer to SNP genotypes. Markers are then ordered by the size of their test statistics, and sums are formed sequentially, starting with the largest test statistic and gradually adding one after another SNP and sums, S_{i} are formed sequentially, starting with the largest test statistic and gradually adding one after another SNP, S_{i} is the sum of the i largest test statistics for the SNPs. For each S_{i} an associated significance level is computed with permutation.
testing (10,000 replicates). Only those SNPs that are significantly associated as part of a set (having taken into account any lack of independence between markers) result in a p value <0.05 after correcting for multiple testing. The S statistic software available at http://www.genemapping.cn/sumstat.html was used. This method has been shown to have consistently higher statistical power than individual marker associations under a range of simulated models. Although set association can incorporate multiplicative effects, no interactions have been reported in the literature for these genes and no such exploratory analyses were undertaken.

Logistic regression

In order to combine all the significant SNPs resulting from the set association analysis into a single variable (not part of the set association method), we then fitted a logistic regression model.

Table 1  Genes and single nucleotide polymorphisms (SNPs) used to assess the combined risk of knee osteoarthritis (OA)

<table>
<thead>
<tr>
<th>Gene symbol</th>
<th>Gene name</th>
<th>SNP alias</th>
<th>SNP rs no.</th>
<th>Chromosome</th>
<th>Position</th>
<th>Higher risk/lower risk allele</th>
<th>Freq of higher risk allele in controls (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>AACT</td>
<td>Alpha1 antiproteinase antitrypsin</td>
<td>aact</td>
<td>rs4934</td>
<td>14</td>
<td>941 505 556</td>
<td>(G/A)</td>
<td>50.9</td>
</tr>
<tr>
<td>ADAM12</td>
<td>A disintegrin and metalloproteinase domain 12 (meltrin alpha)</td>
<td>adam_48</td>
<td>rs3740199</td>
<td>10</td>
<td>128 009 015</td>
<td>(G/C)</td>
<td>47.6</td>
</tr>
<tr>
<td>ASPN</td>
<td>Asporin</td>
<td>aspn_5p</td>
<td>rs7022562</td>
<td>9</td>
<td>92 285 895</td>
<td>(C/T)</td>
<td>69.7</td>
</tr>
<tr>
<td>BMP2</td>
<td>Bone morphogenetic protein 2</td>
<td>bmp2_87</td>
<td>rs1049007</td>
<td>20</td>
<td>6 699 034</td>
<td>(T/C)</td>
<td>35.7</td>
</tr>
<tr>
<td>CALM1</td>
<td>calmodulin 1</td>
<td>calm1_int</td>
<td>rs3213718</td>
<td>14</td>
<td>89 939 666</td>
<td>(T/C)</td>
<td>60.6</td>
</tr>
<tr>
<td>CD36</td>
<td>CD36 antigen (collagen type I receptor, thrombospondin receptor)</td>
<td>cd36_5p</td>
<td>rs1049654</td>
<td>7</td>
<td>80 113 391</td>
<td>(C/A)</td>
<td>54.9</td>
</tr>
<tr>
<td>CILP</td>
<td>Cartilage intermediate protein, nucleotide pyrophosphohydrolase</td>
<td>cilp_3p</td>
<td>rs511822</td>
<td>7</td>
<td>80 116 562</td>
<td>(G/A)</td>
<td>59.4</td>
</tr>
<tr>
<td>COL2A1</td>
<td>Collagen, type II, alpha 1</td>
<td>col2_1405</td>
<td>rs207039</td>
<td>12</td>
<td>46 654 243</td>
<td>(T/C)</td>
<td>90.3</td>
</tr>
<tr>
<td>COX2</td>
<td>Prostaglandin-endoperoxide synthase 2</td>
<td>cox2_102</td>
<td>rs5277</td>
<td>1</td>
<td>184 914 820</td>
<td>(G/C)</td>
<td>18.4</td>
</tr>
<tr>
<td>ESR1</td>
<td>Destrogen receptor alpha</td>
<td>esr_int</td>
<td>rs207311</td>
<td>6</td>
<td>184 807 681</td>
<td>(A/G)</td>
<td>96.3</td>
</tr>
<tr>
<td>FRZB</td>
<td>Frizzled-related protein</td>
<td>frz_int</td>
<td>rs912428</td>
<td>13</td>
<td>46 065 904</td>
<td>(T/C)</td>
<td>20.4</td>
</tr>
<tr>
<td>LRCH1</td>
<td>Leucine-rich repeats and calponin homology (CH) domain containing</td>
<td>lrch1_int</td>
<td>rs912428</td>
<td>13</td>
<td>46 065 904</td>
<td>(T/C)</td>
<td>20.4</td>
</tr>
<tr>
<td>NCO2</td>
<td>Nuclear receptor co-repressor 2</td>
<td>ncr_int</td>
<td>rs2229840</td>
<td>12</td>
<td>123 392 415</td>
<td>(G/A)</td>
<td>82.0</td>
</tr>
<tr>
<td>OPN</td>
<td>Tumour necrosis factor receptor superfamily, member 11b (osteoprotegerin)</td>
<td>opg_5p</td>
<td>rs1546458</td>
<td>8</td>
<td>120 014 347</td>
<td>(A/G)</td>
<td>11.7</td>
</tr>
<tr>
<td>TNA</td>
<td>Tetranectin (plasminogen binding protein)</td>
<td>tna_int</td>
<td>rs93090</td>
<td>3</td>
<td>45 052 127</td>
<td>(A/G)</td>
<td>31.1</td>
</tr>
<tr>
<td>TNFAIP6</td>
<td>Tumour necrosis factor, alpha-induced protein</td>
<td>tnf_int</td>
<td>rs104668</td>
<td>2</td>
<td>151 934 816</td>
<td>(G/A)</td>
<td>87.9</td>
</tr>
<tr>
<td>VDR1</td>
<td>Vitamin D receptor</td>
<td>vdr_int</td>
<td>rs10735810</td>
<td>12</td>
<td>46 559 162</td>
<td>(A/G)</td>
<td>63.7</td>
</tr>
</tbody>
</table>

The control frequency refers to the combined male and female population of the present study.
genes that have a larger influence on risk of OA, the coefficients from the logistic regression were used as weights for the significant SNPs to generate a new “genetic risk” variable. The distribution was standardised to have mean = 0 and SD = 1. Odds ratios and 95% CI were derived from the logistic regression results. The genetic risk distribution between cases and controls was also compared using a Wilcoxon-rank sum test with a continuity correction. Analyses were carried out in S-Plus (0.5113 + (0.9611 × adam_int) + (1.7504 × cilp_3p) + (1.0828 × aspn_5p) + (1.5776 × aspn_3p) + (0.9611 × bmp2_87) + (0.8695 × esr_594)) - 4.59064).

Genetic risk (men) = (1/0.8338) × ((1.9538 × cilp_3p) + (0.5236 × adam_int) + (0.3414 × vdr_365) + (0.3894 × esr1_325) + (2.8437 × cilp_395) + (0.7996 × col2_int) + (0.7187 × lrch4) + (0.4279 × aspn_5p) - 3.5814).

Where adam_int, adam_504 etc denote the genotype (0, 0.5, 1) of the SNP listed in table 1 coded as described under the methods section.

Figure 1 shows the frequency distribution of the genetic risk. The odds ratios between individuals in the top and bottom half of the distribution was 3.32 (95% CI 2.38–4.64, p<1×10^-17) in women and 3.12 (95% CI 2.24–4.86, p<1×10^-10) in men. A Wilcoxon-rank sum test revealed that the genetic risk distribution was significantly different between cases and controls in females with p<1×10^-16 and in males with p<1×10^-13.

When the top and bottom quartiles were used the odds ratios became 8.68 (95% CI 5.20–14.49, p=2×10^-16) for women and 5.06, 2×10^-16 (95% CI 3.10–8.27, p<1×10^-18) for men.

**DISCUSSION**

The data shown here indicate that it is possible to identify individuals at high risk of knee OA using genotype data. Although family and epidemiological studies have consistently indicated an important genetic contribution to OA, to date no single large genetic effect has been found. From a biological point of view these results confirm that the genetic risk to knee OA is likely to be due to the sum of many loci making a small contribution each. Consistent with the gender differences previously reported these data suggest that by using gender-specific genetic factors it may be possible to predict a high risk of knee OA both in men and women.

All the genes included in this model have been reported to be significantly associated with clinical or radiographic features of knee or hip OA in at least one independent study. The odds ratios obtained using the genetic risk variable are comparable than those reported for obesity or knee injury by some studies. For example the odds ratios for obesity/BMI as a risk factor range from 3.012 to 10.012 and the odds ratios for knee injury range from 3.0 for bilateral disease to 6.0 for unilateral disease.

There are some limitations to the present study. First, the genes and all the model parameters were derived from within the same population. Therefore, although all the genes have been independently associated with OA before, we cannot assess the value of these data in predicting knee OA risk without testing this model first in an independent sample. In addition, there are other candidate genes that have been consistently associated with OA that were not genotyped as part of this study (eg, interleukin genes) and that could result in improved prediction.

If these data are replicated in independent studies, it should be possible to indeed use genetic information for risk assessment. An important insight from this study is that the inclusion of more genes as genome-wide association scans and new candidates become available is likely to improve prediction of risk of knee OA to a level where it can be clinically useful.

**Acknowledgements:** The authors wish to thank the two anonymous referees for their valuable comments on an earlier version of this manuscript.

**Funding:** This work was supported by the Arthritis Research Campaign and by the Wellcome Trust.

**Competing interests:** None declared.

**REFERENCES**


