Association of Intronic Single Nucleotide Polymorphism (SNP) of CALM 1 gene with Osteoarthritis of the Knee in Indian Population: A Case-control Study

Sachin Avasthi\*, Abhishek Mishra†, Divya Sanghi‡, Ajai Singh‡, Rashmi Parihar‡, Pankaj Kumar‡, S. Ganesh‡ and R. N. Srivastava§

\*Department of Orthopaedic Surgery, GSVM Medical College, Kanpur (UP), India
†Department of Orthopaedic Surgery, CSMMU, Lucknow (UP), India
‡Department of Biological Sciences and Bioengineering, Indian Institute of Technology, Kanpur (UP), India

(Received 16 March 2011 and accepted 07 June 2011)

ABSTRACT: Osteoarthritis knee is one of the most prevalent disorders in the Indian subcontinent. The wide prevalence and varying features makes it a disease of disguise. Multiple etiological factors have been described. The most recent is genetic contribution in the causation of the disease. This case control study was conducted in the Department of Orthopaedic Surgery, CSM Medical University, Lucknow in collaboration with IIT, Kanpur. 120 cases and 120 controls were enrolled. Clinico-radiological features were noted and symptomatic clinical scoring was done. Genetic polymorphism in relation to intronic region of CALM 1 gene was studied by DNA extraction, Polymerase chain reaction (PCR) and Restriction fragment length polymorphism (RFLP) method. Statistical analysis was done using Stata software. There was no significant difference between age, sex and BMI among cases and controls (p value > .05). ESR (p value =0.0000), fasting blood sugar (p value= 0.0004) and serum uric acid (p value=0.0001) were significantly different among cases and controls. SNP was found in significantly higher number in cases than controls (p value = .0022). Heterozygosity was found only in 5 cases. Logistic regression has also proved significant association of occurrence of Single Nucleotide Polymorphism (SNP) with disease. CALM-1 gene intronic SNP (rs3213718) is present in Indian population. Occurrence of this SNP is significantly affecting the disease.

KEY WORDS: Osteoarthritis knee; Single Nucleotide Polymorphism; CALM 1 gene

INTRODUCTION*

Osteoarthritis diseases are a result of both mechanical and biological events that destabilize the normal coupling of degradation and synthesis of articular cartilage chondrocytes and extracellular matrix. It is estimated that about 80% of the population has radiographic evidence of osteoarthritis knee (OA) by the age of 65 years, although only about 60% of these were symptomatic. Epidemiological profile of this disease in India is not clear but it is estimated that osteoarthritis (OA) is the second most common rheumatological problem and is most frequent joint disease with prevalence of 22% to 39% in India. The primary and secondary subtypes of osteoarthritis have been described according to the causative factors. In primary osteoarthritis, the water content of the cartilage decreases due to a reduced proteoglycan content; thus making the cartilage less resilient. In absence of proteoglycans, the collagen fibers become more susceptible to degradation. This exacerbates the degeneration. In secondary osteoarthritis, there are definite predisposing factors leading to causation of osteoarthritis such as accidental injury to joints, inflammatory diseases (such as Perthes' disease), and other forms of arthritis (e.g. gout, pseudogout, rheumatoid arthritis, etc.). The list includes healed infection of the joints and some forms of sports...
osteoarthritis being around 30%51 and odds ratio being 2.4, we had calculated the sample size. At the significance level of 0.05 and power being taken as 80%, our sample size came out to be 113. We have enrolled 120 cases and control each. **Inclusion criteria:** As per ACR (American College of Rheumatology) 1. Knee pain with osteophytes on X-ray and 2. One of the following: a. Crepitus on knee range of motion b. Age 50 years or older c. Morning stiffness of short duration (<30min) Patients of both sexes and corresponding age group who did not fulfil the diagnostic criteria of osteoarthritis of knee were included as controls. These were either patients coming with some other problem or the otherwise healthy attendants of patients. **Exclusion criteria:** All patients who did not fulfil the criteria of selection or had any feature suggestive of secondary causes of osteoarthritis were excluded from the study. After the informed consent, cases and controls were evaluated for osteoarthritis. A detailed history was taken and thorough examination of the affected joint and related regions was done. X-ray of affected joint was done and graded as per Kellgren Lawrence grading system. Patients were subjected to biochemical parameters such as Serum uric acid, ESR and Rheumatoid factor. Severity of the disease was analysed according to Visual Analogue Score, WOMAC Index and Lequesne Index. Whole Blood samples of both the groups were taken. Blood sample was stored in the deep freezer (-20°C).

**Genotyping**

DNA isolation was done by Flexigene DNA extraction kit (Cat no-51206). 300 μl of whole blood was used for DNA isolation. Isolated DNA was checked in agarose gel (1%) which was prepared by adding 0.25 gm of agarose in 25.0 ml of 1 X Tris-acetate-EDTA (TAE) buffer. After confirming the DNA isolation, rest of the samples were stored at –20°C for PCR analysis. The desired details of CALM-1 gene intronic SNP (IVS3-293C>T = rs3213718) was taken from the NCBI tool. The desired SNP was amplified using forward primer sequence as 5’-CTC AGG GAT GGC AGT CGG-3’ and reverse primer sequence as 5’-TGA AAG CGG AAG AAG CCA TAC-3’. The SNP created a restriction site for Hpy 188 III restriction enzyme. Polymerase chain reaction was set using the above information for each sample. 10.0μl reaction was set from each sample. For preparation of 10μl reaction mix, 0.5μl DNA, 0.5μl each of

**Sample size**

injuries7 (from exercise, athletic activity), avascular necrosis along with obesity8,9 and congenital disorders10,11

Though the pathology of osteoarthritis is well defined, the etiological factors are not fully characterized. Risk factors have been defined for incidental osteoarthritis. Body weight as risk factor has been proved by multiple studies12-18. Age is a well known risk factor associated with osteoarthritis knee9-28. Gender may influence knee OA via multiple routes including hormonal influences on cartilage metabolism, gender variation in injury risk and gender differences in mechanical environment of knee e.g. varus valgus laxity21, strength relative to body weight29. Occupation is also a definitive risk factor causing osteoarthritis. Different occupations have been tested for predisposition towards OA and significant effect of type of occupational activity has been proved21-43. The familial aggregation of clinical and radiographic features of osteoarthritis was first convincingly demonstrated by Stecher in 1940’s in studies of Heberden’s nodes44. Studies that had examined the frequency of disease among siblings of probands who had undergone total hip or knee replacement have suggested a familial effect45. Twin studies have provided a classic approach to separate the genetic influences from the shared family environment as a cause of familial aggregation46-47. Genetic exploration of genome has resulted in several susceptibility loci isolation confirming the genetic association of disease.

With the above background, a Japanese group found that their population has shown higher incidence of osteoarthritis in patients having intronic and core promoter SNP in CALM-1 gene48. At the same time, Caucasian21 and Greek25 population showed absence of any such predisposition in their population with the CALM-1 gene SNP.

We planned a case-control study in patients of primary osteoarthritis knee with aims being to study the presence of CALM-1 gene SNP in Indian population, correlation of presence of CALM-1 gene SNP with osteoarthritis in Indian population and correlation of presence of SNP with clinico-radiological stage of the disease.

**METHODOLOGY**

The study was conducted in the Department of Orthopaedic Surgery and Department of Pathology, CSMMU, UP (Formerly known as King George’s Medical University, Lucknow) in collaboration with Indian Institute of Technology, Kanpur between August 2006 to August 2010. We planned a case-control study in which all the patients and controls were selected as per inclusion and exclusion criteria.

---

**Inclusion criteria:**

1. Knee pain with osteophytes on X-ray and
2. One of the following:
   a. Crepitus on knee range of motion
   b. Age 50 years or older
   c. Morning stiffness of short duration (<30min)

Patients of both sexes and corresponding age group who did not fulfil the diagnostic criteria of osteoarthritis of knee were included as controls. These were either patients coming with some other problem or the otherwise healthy attendants of patients.

**Exclusion criteria:**

All patients who did not fulfil the criteria of selection or had any feature suggestive of secondary causes of osteoarthritis were excluded from the study. After the informed consent, cases and controls were evaluated for osteoarthritis. A detailed history was taken and thorough examination of the affected joint and related regions was done. X-ray of affected joint was done and graded as per Kellgren Lawrence grading system. Patients were subjected to biochemical parameters such as Serum uric acid, ESR and Rheumatoid factor. Severity of the disease was analysed according to Visual Analogue Score, WOMAC Index and Lequesne Index. Whole Blood samples of both the groups were taken. Blood sample was stored in the deep freezer (-20°C).

**Genotyping**

DNA isolation was done by Flexigene DNA extraction kit (Cat no-51206). 300 μl of whole blood was used for DNA isolation. Isolated DNA was checked in agarose gel (1%) which was prepared by adding 0.25 gm of agarose in 25.0 ml of 1 X Tris-acetate-EDTA (TAE) buffer. After confirming the DNA isolation, rest of the samples were stored at –20°C for PCR analysis. The desired details of CALM-1 gene intronic SNP (IVS3-293C>T = rs3213718) was taken from the NCBI tool. The desired SNP was amplified using forward primer sequence as 5’-CTC AGG GAT GGC AGT CGG-3’ and reverse primer sequence as 5’-TGA AAG CGG AAG AAG CCA TAC-3’. The SNP created a restriction site for Hpy 188 III restriction enzyme. Polymerase chain reaction was set using the above information for each sample. 10.0μl reaction was set from each sample. For preparation of 10μl reaction mix, 0.5μl DNA, 0.5μl each of...
forward and reverse primers were added along with 0.6μl dNTPs, 1μl Taq polymerase buffer, 0.1μl Taq polymerase and 6.8μl autoclaved distilled water. The standardization of PCR cycles was denaturation phase (95°C x 5 min), denaturation phase (95°C x 30 sec), annealing phase (64°C x 45 sec) and extension phase (72°C x 45 sec). The size of amplicon (Figure 1) was checked using agarose gel electrophoresis against control DNA ladder of known molecular weight. SNP typing was done using Restriction Fragment Length Polymorphism method. Hpy 188 III was the desired restriction enzyme which was found to cut the 535 base pair long PCR amplicon into 81 and 454 base pair fragments. Digested products were examined using Poly acrylamide gel electrophoresis (PAGE). (Figure 2)

Figure 1: PCR amplification of CALM-1 gene (The amplicon is 535 base pair size)

Figure 2: PAGE gel results of CALM-1 gene

Statistical analysis:

Data was analysed using Stata S.E. /11.1 software. We did univariate analysis between cases and controls to know the distribution of variables in two groups. In significantly different variables, we applied multivariate analysis and logistic regression to assess the significance of variables.

RESULTS

120 cases and 120 controls were enrolled. The basic demographic profile is mentioned in Table 1. Mean height of cases was 158.7 +/- 15.4 cm and mean height of controls was 161.18 +/- 7.3 cm. The difference in cases and controls was insignificant (p value = 0.1109). Mean weight of cases was 64.48 +/- 10.2 kg and mean weight of controls was 65.85 +/- 7.1 kg. The difference in weight between cases and controls was insignificant (p value = 0.4695). Mean body mass index (BMI) in cases was 25.54 +/- 2.6 and mean BMI in controls was 25.45 +/- 2.2. The difference in cases and controls was insignificant (p value = 0.5013).

Lab investigation differences are mentioned in Table 2. Erythrocyte sedimentation rate (ESR) was significantly higher in cases (26.83 +/- 9.0 mm in first hour) as compared to controls (19.98 +/- 6.2 mm in first hour) with p value being 0.0000. Mean fasting blood sugar was 90.57 +/- 13.7 mg/dl in cases and 86.44 +/- 12.6 mg/dl in controls. The difference came out to be significant (p value = 0.0004). However postprandial blood sugar was 122.58 +/- 14.0 mg/dl in cases and 124.50 +/- 16.8 mg/dl in controls. This was insignificantly different in the groups (p value = 0.7284). Mean Serum uric acid in cases came out to be 4.79 +/- 1.0 mg/dl and it was 4.04 +/- 0.9 mg/dl in controls. The difference in cases and controls came out to be highly significant (p value = 0.0001).

We detected the intronic Single nucleotide polymorphism (SNP) in both cases and controls. 39 (32.50%) out of 120 cases showed presence on SNP in our population. In comparison, only 18 (15%) of controls showed the presence of SNP. This shows that the SNP exists in Indian population and the difference in occurrence of SNP in cases and controls is significant (P value = 0.0022) suggesting clinically important role of such SNP in Indian population. Among SNP positive cases, only 5 (8.7%) of them showed heterozygosity for the occurrence of SNP. None of the controls were heterozygous for SNP. (Table 3)

We did regression of affecting variables against SNP taking the presence of osteoarthritis as dependent variable. Taking one by one, we calculated the adjusted odds ratio of all the significant variables. Finally we did logistic regression to see the effect of these variables on the occurrence of disease. We found that age, sex, SNP presence, ESR, postprandial blood sugar and serum uric acid are the variables which affect the occurrence of disease significantly (p value < .05). (Table 4)
### Table 1: Demographic profile

<table>
<thead>
<tr>
<th>Subject</th>
<th>Number</th>
<th>Sex</th>
<th>Age (in years)</th>
<th>BMI</th>
<th>KL Grade</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Male</td>
<td>Female</td>
<td>Mean</td>
<td>SD</td>
</tr>
<tr>
<td>Cases</td>
<td>120</td>
<td>48(40%)</td>
<td>72(60%)</td>
<td>53.81</td>
<td>8.3</td>
</tr>
<tr>
<td>Controls</td>
<td>120</td>
<td>66(55%)</td>
<td>54(45%)</td>
<td>50.73</td>
<td>7.2</td>
</tr>
</tbody>
</table>

### Table 2: Laboratory investigations

<table>
<thead>
<tr>
<th>Subjects</th>
<th>Number</th>
<th>ESR (mm in 1st hr)</th>
<th>Blood Sugar fasting (mg/dl)</th>
<th>Blood Sugar Postprandial (mg/dl)</th>
<th>Serum uric Acid (mg/dl)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Mean +/- SD</td>
<td>Mean +/- SD</td>
<td>Mean +/- SD</td>
<td>Mean +/- SD</td>
</tr>
<tr>
<td>Cases</td>
<td>120</td>
<td>26.83 +/- 9.0</td>
<td>90.57 +/- 13.7</td>
<td>122.58 +/- 14.0</td>
<td>4.79 +/-1</td>
</tr>
<tr>
<td>Controls</td>
<td>120</td>
<td>19.98 +/- 6.2</td>
<td>86.44 +/- 12.6</td>
<td>124.50 +/- 16.8</td>
<td>4.04 +/-0.9</td>
</tr>
<tr>
<td>Total</td>
<td>240</td>
<td>23.40 +/- 8.5</td>
<td>88.50 +/- 13.3</td>
<td>123.54 +/- 15.4</td>
<td>4.41 +/-0.1</td>
</tr>
<tr>
<td>P value</td>
<td></td>
<td>0.0000</td>
<td>0.0004</td>
<td>0.7284</td>
<td>0.0001</td>
</tr>
</tbody>
</table>

### Table 3: Single nucleotide polymorphism

<table>
<thead>
<tr>
<th>SNP</th>
<th>Cases</th>
<th>Controls</th>
</tr>
</thead>
<tbody>
<tr>
<td>Absent</td>
<td>81 (67.50%)</td>
<td>102 (85%)</td>
</tr>
<tr>
<td>Present</td>
<td>39 (32.50%)</td>
<td>18 (15%)</td>
</tr>
<tr>
<td></td>
<td>Homozygous: 34</td>
<td>All homozygous</td>
</tr>
<tr>
<td></td>
<td>Heterozygous: 05</td>
<td></td>
</tr>
</tbody>
</table>

P value = 0.0022; OR = 2.73 (95% CI: 1.40, 5.45)

### Table 4: Statistical distributions

<table>
<thead>
<tr>
<th>Group</th>
<th>β coefficient</th>
<th>Std. Error</th>
<th>P value</th>
<th>95% confidence interval</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age</td>
<td>.0616</td>
<td>.021</td>
<td>.004</td>
<td>.0192, .1039</td>
</tr>
<tr>
<td>Sex</td>
<td>.8977</td>
<td>.325</td>
<td>.006</td>
<td>.2595, 1.535</td>
</tr>
<tr>
<td>SNP</td>
<td>.9866</td>
<td>.385</td>
<td>.010</td>
<td>.2312, 1.741</td>
</tr>
<tr>
<td>ESR</td>
<td>1.357</td>
<td>.350</td>
<td>.000</td>
<td>.6698, 2.045</td>
</tr>
<tr>
<td>blood sugar (PP)</td>
<td>-2.931</td>
<td>.774</td>
<td>.000</td>
<td>-4.449, -1.412</td>
</tr>
<tr>
<td>Uric Acid</td>
<td>.9648</td>
<td>.187</td>
<td>.000</td>
<td>.597, 1.332</td>
</tr>
<tr>
<td>Constant</td>
<td>-9.742</td>
<td>1.68</td>
<td>.000</td>
<td>-13.03, -6.449</td>
</tr>
</tbody>
</table>

Log likelihood = -117.752; Pseudo R² = .2922

Copyrighted © by Dr. Arun Kumar Agnihotri. All right reserved
DISCUSSION

The study of Osteoarthritis knee (OA) was planned as the disease is the most prevalent Orthopaedic degenerative disorder in our country. Almost every third individual beyond the age of fifty years is suffering from osteoarthritis knee. The wide spectrum of etiopathological factors is associated with socioeconomic factors in leading to causation of OA. We have limited number of studies to depict the role of these factors in our population. It is probably because of the poor socioeconomic status of majority of population that the patients suffering from this disorder are presenting late when the disease is more advanced and cartilage degeneration is almost complete. The subjective nature of clinical symptomatology poses more difficulty in assessing the severity of disease. The real nature of the disease and the stage can be assessed only after full clinicoradiological assessment supplemented with pathological findings.

We hardly know anything about the genetic database of our population. Detection of genetic changes and development of gene therapy is the need of future. We are already lagging behind in isolation of genetic changes existing in our population in spite of several research projects running in the country. Probably more extensive research is required to highlight these changes in our country otherwise we will have to depend on western literature. The serious implications of such lacking will result in prescription of future gene therapeutic modalities to our population which is actually not explored for such changes.

40% of our patients were males and 60% were females. This shows the female preponderance of the disease in our population in spite of random nature of case selection. Logistic regression between cases and controls has also shown that the sex incidence is significantly different in cases and controls (p value = .006).

In our study, mean age of presentation in cases was 53.31+/-.95 years. The mean age of controls was 50.64+/-.73 years. The difference in age incidence in cases and controls came out to be significant. Regression model for cases and controls has shown significant effect of age on occurrence of disease (p value = .004). Our observations are consistent with Mangat et al\(^1\) who have shown that the mean age of onset was 48.9 yrs in females and 54.7 yrs in males and mean duration of illness was 5.43 yrs in females and 3.39 yrs in males. Sharma et al\(^2\) have also shown that Osteoarthritis was present in only 50.2% of the elderly aged 65-74 years, whereas it was 97.7% in elderly aged 84 years or older (p <0.001). Framingham\(^12,13\) and Baltimore studies have also shown that the disease increases with age.

Mean body mass index (BMI) in cases was 25.54+/-.2.6 and mean BMI in controls was 25.45+/-.2.2. The difference in cases and controls was insignificant (p value = 0.5013). Our data is not comparable with other studies. Felson DT (1995) and Creamer P (1997) in their studies have shown the importance of weight in affecting joint degeneration. Mangat et al\(^1\) have also shown higher BMI in their study (mean BMI = 28.09+/-.4.43). Sharma et al\(^2\) have shown that BMI significantly affects the disease process.

Erythrocyte sedimentation rate was significantly different between cases and controls (p value = 0.0000). Regression analysis between cases and controls has shown very significant effect of ESR in disease (p value = 0.000). However, among cases, there was no significant difference between SNP positive and negative cases. This was further confirmed on logistic regression which showed that ESR is not significant in association with severity of disease. We found that since ESR is associated with multiple physiological factors including age and sex, this is not associated with causation and severity of disease.

Fasting blood sugar was significantly different between cases and controls (p value = .0004) whereas postprandial blood sugar was not (p value = .7284). However, in regression model between cases and controls, we found that postprandial blood sugar is significantly associated with disease. This may be because of dichotomization of values for regression model. Among cases, the difference in SNP positive and negative cases for both fasting and postprandial blood sugar came out to be insignificant (p value > .05). Logistic regression also ruled out any association of blood sugar with severity of the disease. Our findings are consistent with Mangat et al study\(^3\) which showed that diabetes was an associated illness in only 22 patients out of 300 enrolled in their study. Mean serum uric acid came out to be 4.41+/-.1 mg/dl. In case its value was 4.79+/-1 mg/dl and in controls, it was 4.04+/-.9 mg/dl. This difference in cases and controls was significant (p value = .0001). In regression model of cases and controls, uric acid has come out to be a very important variable affecting disease (p value = .000).

We detected the intronic Single nucleotide polymorphism (SNP) in both cases and controls. 39 (32.50%) out of 120 cases showed presence on SNP in our population. In comparison, only 18 (15%) of the controls showed the presence of SNP. This shows that the SNP exists in Indian population and the difference in occurrence of SNP in cases and controls is significant (P value = 0.0022) suggesting clinically important role of such SNP in Indian population. Regression model for cases and controls has also shown that SNP is significantly associated with disease (p value = .010). However, heterogeneity in relation to occurrence is not
We conclude that CALM 1 gene intronic SNP (rs3213718) is present in Indian Population. The target SNP significantly affects the disease as the difference between cases and controls is highly significant (p value = .0022). Females are more predisposed to OA in comparison to males. Mean age of presentation was 53.31 +/- 9.5 years. Age of occurrence of disease is a significant factor in causation of this disease. However it is not influenced by existence of SNP. Height, weight and BMI did not show any significant difference in disease occurrence between cases and controls.

CONCLUSION

We suggest the possibility of association between SNP and the disease process.

REFERENCES