Epistasis Between Two Gene Variants of Leptin and Vascular Endothelial Growth Factor Genes in the Development of Primary Knee Osteoarthritis

Matvey Sosa-Arellano¹, Norma C. González-Huerta¹, Eugenio Morales-Hernández², Carolina Duarte-Salazar³, and Antonio Miranda-Duarte¹

Departments of ¹Genomic Medicine and ²Radiology, and ³Rheumatology Service, Instituto Nacional de Rehabilitación, Luis Guillermo Ibarra Ibarra, Mexico City, Mexico

ABSTRACT

Background: The association of leptin (LEP) and vascular endothelial growth factor A (VEGFA) genes with the susceptibility to knee osteoarthritis (OA) has been analyzed; however, the epistasis between them has not been investigated. Objective: The objective of the study was to analyze the association of LEP and VEGFA variants and their interaction with primary knee OA in a Mexican Mestizo population. Methods: A case-control study was developed. Cases were ≥40 years, BMI ≤27 kg/m², with primary knee OA and radiologic Grade ≥2. Controls were participants with no knee OA and a radiologic Grade <2. The rs2167270 of LEP and rs2010963 of VEGFA were genotyped. Genotypic association was tested under codominant, dominant, and recessive models. Uni- and multi-variate analyses were developed through non-conditional logistic regression. The multifactor dimensionality reduction algorithm was developed to detect epistasis. Results: Participants comprised 103 cases and 179 controls. Allelic and genotypic distributions did not show differences between the groups. Notwithstanding, a statistically significant interaction was observed between the LEP and VEGFA genes (p = 0.02) with a testing accuracy of 0.5199 and cross-validation consistency of 10/10. This interaction model confers an increased risk to knee OA (OR [95% CI] = 1.8 [1.1-2.9]). Conclusion: Interaction between LEP and VEGFA is related with genetic susceptibility to developing primary knee OA. (REV INVEST CLIN. 2022;74(2):81-9)

Keywords: Osteoarthritis. Leptin. Vascular endothelial growth factor. Gene variant. Epistasis.
INTRODUCTION

Osteoarthritis (OA) is the most common joint disorder worldwide. The knee is the most commonly involved site, with an estimated prevalence between 10% and 13% in elderly people\(^1\). OA is classified as primary when there is no apparent underlying cause, and secondary when it is due to a known cause. OA is a whole joint disease primarily characterized by articular cartilage degeneration; however, synovial inflammation, subchondral bone sclerosis, and osteophyte formation are also observed\(^2\). Inflammation plays a role in the pathogenesis of OA. Inflammation stimulates chondrocytes and synovium cells to produce cytokines such as interleukin-1 beta (IL-1\(\beta\)) and tumor necrosis factor-alpha (TNF-\(\alpha\)), as well as collagenases and aggreganases such as matrix metalloproteinases (MMPs) and a disintegrin and metalloproteinase with thrombospondin type I motifs (ADAMTS), favoring the catabolism of articular cartilage\(^2,3\). Angiogenesis is also involved in OA development. During the osteoarthritic process, blood vessels from the subchondral bone invade the articular cartilage, which is normally avascular, facilitating the progression of OA and the formation of osteophytes. Likewise, angiogenesis occurs in the synovium, where it is strongly linked to macrophage infiltration and inflammation\(^4\).

Leptin (LEP) belongs to the type 1 cytokine superfamily. LEP is mainly produced by adipocytes and is primarily known for its role as a hypothalamic modulator of food intake; however, it also has a role as a cytokine upregulating the secretion of inflammatory cytokines, including IL-6, IL-12, and TNF-\(\alpha\)\(^5,6\). LEP plays a key role in the OA process. LEP levels are increased in synovial fluid and serum of OA patients; furthermore, there is a correlation between the LEP levels of synovial fluid and OA severity\(^7,8\). LEP is also produced by chondrocytes and promote cartilage damage by inducing the production of IL-1\(\beta\), IL-6, IL-8, MMP-1, MMP-3, MMP-13, ADAMTS-4, ADAMTS-5, and nitric oxide (NO)\(^9,10\). Vascular endothelial growth factors A are a family of platelet-derived growth factors being VEGF-A the most prominent\(^11\). VEGF-A is a potent angiogenic factor and also plays an active role in OA. In OA patients, VEGF-A levels are increased in the synovial fluid and serum, and there is also a correlation between synovial fluid VEGF-A levels and the grade of OA severity\(^12,13\). Hypoxic, mechanical, and inflammatory stimuli increase VEGF-A production in chondrocytes\(^14\). In chondrocytes and synovial fibroblasts, VEGF-A induces cartilage catabolic mediators such as NO, MMP-1, MMP-3, and MMP-13, as well as IL-1\(\beta\), IL-6, and TNF-\(\alpha\)\(^15,16\).

Despite the prominent role of LEP and VEGF-A in OA pathogenesis, to date, there are few studies analyzing the association of LEP and VEGFA gene variants with knee OA\(^17-19\). Genetic association studies in primary OA have typically focused on independent effects of candidate genes. Epistasis is the interaction between two or more genes. This interaction can modify the genetic associations independently of the particular effects of the genes and can be a factor in the susceptibility to developing a disease\(^20\). It is, therefore, necessary to explore epistasis in OA, as well as single-gene effects, which prompted us to analyze the association of LEP and VEGFA gene variants with primary knee OA as well as for an interaction between them in the Mexican Mestizo population.

METHODS

Study design and sample population

A case-control study was conducted, which protocol was approved by the Ethics and Investigation Committee of the Instituto Nacional de Rehabilitación Luis Guillermo Ibarra Ibarra, in Mexico City. All study participants were recruited during the same period and were of Mexican mestizo origin, this being defined as a person born in Mexico with a Spanish-derived surname, and a family of Mexican ancestors back to the third generation. Cases were individuals aged ≥ 40 years, with clinical diagnosis of primary knee OA (radiographic knee OA defined as grade of ≥ 2); body mass index (BMI, kg/m\(^2\)) of ≤ 27, no history of severe injury or knee surgery, and no other joint diseases. Controls were subjects ≥ 40 years of age with no clinical diagnosis of knee OA, no radiographic knee OA defined as grade of < 2); body mass index (BMI, kg/m\(^2\)) of ≤ 27, no history of severe injury or knee surgery, and no other joint diseases. Controls were subjects ≥ 40 years of age with no clinical diagnosis of knee OA, no radiographic knee OA defined as radiologic grade < 2, a BMI of ≤ 27, and no history of serious knee injury or joint diseases.

Controls arrived mainly due to orthopedic problems involving no serious knee damage (i.e., comprising shoulder lesions or fractures). The main classification criteria of the study groups was radiologic. The Kellgren-Lawrence classification method was assessed in
anteroposterior weight-bearing and lateral knee x-rays. The classification is as follows: grade 0: no pathology; grade 1: doubtful joint space narrowing and possible osteophytes; grade 2: definite osteophytes and possible joint space narrowing; grade 3: definite osteophytes with moderate joint space narrowing and some sclerosis; grade 4: large osteophytes, marked joint space narrowing, severe sclerosis, and bone contour deformity21,22.

The radiological evaluation was performed by a sole experienced observer blinded to patients’ data. All subjects were interviewed and responded to a structured questionnaire designed for this study to collect demographic data, comorbidities, and occupational and sport activities, among other information.

Genotyping

The rs2167270 (G>A) of LEP and rs201093 (C>T) of VEGFA are non-coding gene variants located in the 5’ untranslated region (UTR) region (https://www.ncbi.nlm.nih.gov/snp/). Those gene variants have been associated with entities whose pathologic processes are related to inflammation and angiogenesis (e.g., rheumatoid arthritis and lupus erythematosus); therefore, we chose to analyze their association with knee OA. After obtaining signed, written, informed consent from each participant, a 5 mL sample of peripheral blood was collected. Genomic DNA was extracted, and TaqMan SNP Genotyping Assays were used to identify the rs2167270 and rs2010963 according to the manufacturer’s instructions (Assay IDs C_15966471_20 and C_8311614_10; Applied Biosystems, Foster City, CA, USA). Genotyping was performed in a Rotor-Gene Q Real-Time PCR fluorescence instrument (QUIAGEN, Hilden, Germany), and software TeeChart (Steema Software SL) was utilized to analyze the PCR genotyping results.

Statistical analysis

Data distribution was assessed through a Shapiro–Wilk test in both study groups. The comparison of continuous variables was performed using Student’s t-test for normally distributed data or Mann–Whitney U-test for non-normally distributed data. For categorical variables, Chi-squared statistics (χ²) or Fisher’s exact tests were applied. For each gene variant, allelic and genotypic frequencies as well as Hardy-Weinberg equilibrium (HWE) were calculated. The genotypic association analysis was tested under codominant, dominant, and recessive inheritance models. Univariate and multivariate analyses were developed through non-conditional logistic regression considering each genotype as the main effect. To estimate the association magnitude of LEP and VEGFA variants with primary knee OA, unadjusted and adjusted odds ratios (ORs) with their 95% confidence intervals (95% CI) were reported. STATA ver. 15.0 statistical software package was utilized for calculations. To detect gene–gene interaction, the multifactor dimensionality reduction (MDR) algorithm was developed. This approach was designed specifically to detect, characterize, and interpret gene–gene interactions in the absence of statistically detectable independent effects23. The software MDR 3.0.2 is available at epistasis.org, and the full procedure can be reviewed elsewhere23,24. Briefly, the data are randomly divided into 10 parts for cross-validation. From the pool of genetic factors, a set of n factors are selected and their multifactor cells are represented in n-dimensional space (e.g., for two polymorphisms, there are nine two-locus genotype combinations).

The ratio of the number of cases to the number of controls is calculated within each multifactor cell and is labeled as high and low risk depending on the threshold ratio (i.e., the ratio of the number of cases to that of controls), reducing the n-dimensional space to one dimension with two levels (low and high risk). All possible combinations of factors are evaluated for their ability to classify cases and controls in 9/10 of the data (training set) to select the best n-factor model. The remaining 1/10 data (testing set) are used for independent testing for cross-validation consistency (CVC), which is a measure of the number of times a particular set of factors is identified in each possible 9/10 of the subjects. The process is repeated 10 times with the data split into 10 different training and testing sets. The best models are shown with their training and testing accuracy (i.e., the proportion of subjects that are grouped correctly according to their status), the CVC, and the statistical significance of the model (evaluated using a sign test to compare the observed testing accuracies to those expected under the null hypothesis of no association). The best model is that with the highest CVC and testing accuracy values since this model shows more consistent results and is the most likely to generalize to independent data. In addition, the
ORs (95% CI) are provided. A graphical model is obtained showing the distribution of each genotype combination in cases and controls, as well as an interaction graph showing the independent effects of the genes and their interactive relationships.

RESULTS

A total of 282 participants were included: 103 comprised cases with primary knee OA and 179 were controls. Females were more frequent in cases and controls (71.8% and 67.6%, respectively; p = 0.4). Mean age was 65.5 ± 14.6 years in cases and 60.8 ± 15.6 years in controls (p = 0.014). BMI showed a mean of 25.1 ± 3.1 and 25.0 ± 3.8 in cases and controls, respectively (p = 0.8). Arterial hypertension and Type II diabetes mellitus were the most frequent comorbidities (p = 0.03 and 0.05, respectively) (Table 1).

Both gene variants, rs2167270 of LEP and rs2010963 of VEGFA, were in agreement with HWE (p = 0.3 and 0.4, respectively). The allelic distribution of both SNPs did not show differences among the groups (p = 0.02 and 0.4; respectively) (Table 2). In the genotypic association analysis, a tendency toward an increased risk under a codominant model and a recessive model was observed for AA homozygous of rs2167270. In the case of rs2010963, a tendency toward a lower risk for CC homozygous was observed under a codominant and a recessive model. However, these tendencies were not statistically significant (Table 2).

Through MDR analysis, an interaction was observed between the LEP and VEGFA. This model had a testing accuracy of 0.5199 and a maximum CVC of 10/10. The model was significant at the level of 0.02, indicating that all testing accuracies were greater than 0.5 during cross validation (Table 3). In addition, this model was associated with an increased risk to knee OA [OR (95% CI) = 1.8 (1.1-2.9)]. Figure 1A summarizes the distribution of the two-locus genotype combination of LEP and VEGFA in cases and controls.
Table 2. Allelic and genotypic association of LEP and VEGFA with primary knee osteoarthritis

<table>
<thead>
<tr>
<th>SNP</th>
<th>Cases (n = 103)</th>
<th>Controls (n = 179)</th>
<th>OR (CI 95%)</th>
<th>p</th>
<th>OR (CI 95%)</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>n (%)</td>
<td>n (%)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>LEP</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>rs2167270</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>A</td>
<td>105 (51.0)</td>
<td>162 (45.0)</td>
<td>1.3 (0.9-1.8)</td>
<td>0.2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>G</td>
<td>101 (49.0)</td>
<td>196 (55.0)</td>
<td>0.8 (0.6-1.1)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Codominant</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AA</td>
<td>31 (30.1)</td>
<td>40 (22.3)</td>
<td>1.5 (0.8-2.9)</td>
<td>0.2</td>
<td>1.6 (0.8-3.1)</td>
<td>0.2</td>
</tr>
<tr>
<td>AG</td>
<td>43 (41.7)</td>
<td>82 (45.8)</td>
<td>1.03 (0.6-1.8)</td>
<td></td>
<td>1.04 (0.6-1.9)</td>
<td></td>
</tr>
<tr>
<td>GG</td>
<td>29 (28.2)</td>
<td>57 (31.8)</td>
<td>1</td>
<td></td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>Dominant</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AA/AG versus GG</td>
<td>74 (71.8)</td>
<td>122 (68.2)</td>
<td>1.2 (0.7-2.0)</td>
<td>0.5</td>
<td>1.2 (0.7-2.1)</td>
<td>0.4</td>
</tr>
<tr>
<td>Recessive</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AA versus AG/GG</td>
<td>31 (30.1)</td>
<td>40 (22.3)</td>
<td>1.5 (0.9-2.6)</td>
<td>0.1</td>
<td>1.5 (0.9-2.7)</td>
<td>0.1</td>
</tr>
<tr>
<td>VEGFA</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>rs2010963</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>C</td>
<td>71 (34.0)</td>
<td>135 (38.0)</td>
<td>0.8 (0.6-1.3)</td>
<td>0.4</td>
<td></td>
<td></td>
</tr>
<tr>
<td>T</td>
<td>135 (66.0)</td>
<td>223 (62.0)</td>
<td>1.1 (0.8-1.7)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Codominant</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CC</td>
<td>11 (10.7)</td>
<td>28 (15.7)</td>
<td>0.6 (0.3-1.4)</td>
<td>0.3</td>
<td>0. (0.3-1.5)</td>
<td>0.3</td>
</tr>
<tr>
<td>CT</td>
<td>49 (47.6)</td>
<td>79 (44.1)</td>
<td>1.0 (0.6-1.7)</td>
<td></td>
<td>1.0 (0.6-1.7)</td>
<td></td>
</tr>
<tr>
<td>TT</td>
<td>43 (41.7)</td>
<td>72 (40.2)</td>
<td>1</td>
<td></td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>Dominant</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CC/CT versus TT</td>
<td>60 (58.2)</td>
<td>107 (59.8)</td>
<td>0.9 (0.6-1.5)</td>
<td>0.8</td>
<td>0.9 (0.6-1.5)</td>
<td></td>
</tr>
<tr>
<td>(0.5-1.5)</td>
<td>0.7</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Recessive</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CC versus CT/TT</td>
<td>11 (10.7)</td>
<td>28 (15.7)</td>
<td>0.6 (0.3-1.3)</td>
<td>0.2</td>
<td>0.6 (0.3-1.4)</td>
<td>0.2</td>
</tr>
</tbody>
</table>

Unadjusted odds ratios and 95% confidence intervals.

Adjusted by age, diabetes mellitus II, and arterial hypertension. VEGFA: vascular endothelial growth factor A.
and for the new MDR-constructed variable. As can be observed, the pattern of high- and low-risk cells for the LEP variant differed depending on the value of VEGFA variant, which provides evidence of epistasis. The interaction graph showed in figure 1B indicates the amount of information gained about case–control status by putting the two gene variants together. A low percentage of information gain of case–control status explained by LEP or VEGFA considered independently was observed (0.53% and 0.36%; respectively). However, a large percentage of information gain explained by the interaction between these two loci was observed (1.53%), which is interpreted as a synergistic or non-additive relationship between the VEGFA and LEP.

**DISCUSSION**

Our results showed no association between LEP and VEGFA gene variants with primary knee OA when their independent effects were analyzed. Nevertheless, MDR analysis provided evidence of epistasis between LEP and VEGFA, which could be implicated in the development of primary knee OA. To the best of our knowledge, there are no studies analyzing the gene-gene interaction between those genes in the context of OA. Moreover, to date, only few studies have analyzed the association of LEP and VEGFA gene variants with primary knee OA. In the Han Chinese population, three variants of LEP were assessed for their contribution to knee OA: the rs11761556 and rs12706832, located in the intron region, and the rs2071045 of the 3´UTR region. In general, the genotypes of those variants were significantly associated with a lower risk to knee OA in normal weight and overweight groups, but not in the obese group19. Two independent studies analyzed the association of VEGFA variants with knee OA in the Mexican population. In a population sample settled in West Mexico, the –460 T/C (rs833061) and +405 C/G (rs2010963), located in intron and 5´UTR regions of VEGFA, respectively, were assessed using polymerase chain reaction-restriction fragment length polymorphisms. The variants did not show significant association with knee OA. Notably, the cases had a significantly high BMI (29.9 [17.5-46.9])17. The other study analyzed the variants rs699947 and rs3025039 (located in intron and 3´UTR regions of VEGFA, respectively) in cases with knee OA who had a BMI of 29.0 ± 4.19; however, there was no association. This study also analyzed epistasis among several genes, and a non-statistically significant interaction of VEGFA with other genes such as MMP3, COL3A1, HF1AN, and EGF was found (LEP was not analyzed)18. The rs2167270 of LEP has not been previously investigated in OA. This variant was not associated with knee OA, and we only observed a non-statistically significant tendency toward an increased risk for AA homozygous. Regarding rs2010963 of VEGFA, there was no association, and we merely observed a tendency toward a lower risk. This variant was previously analyzed in the population of West Mexico, and the authors did not find an association with OA, which seems to be in line with our findings. However, according to the genotype frequencies reported in that study, there is an HWE deviation indeed (p = 0.0006), which could affect their results17. Of notice, in our study, the LEP and VEGFA variants follow HWE, and the minor allele frequencies are not importantly different to those reported in population with Mexican ancestry in The 1,000 Genomes Project (https://www.ensembl.org/index.html/). Therefore, we consider that our results are reliable.

Single-nucleotide variants (SNVs) are the most common variant type in the human genome.
Genome-wide association studies have uncovered robust associations between SNVs and numerous complex diseases. However, only ~4% of the associated variants identified result in differences in the protein product. As a result, an SNV can be the “best available marker,” rather than the “probable causal variant”. SNVs in the UTRs, generally do not change the protein sequence, but may regulate gene expression. The 5´UTR possesses a consensus sequence containing the translation initiation codon, as well as regulatory elements (e.g., RNA-binding proteins sites). Thus, variants at 5´UTR can have an impact on the production of the protein. For example, the 90 A/G variant at the 5´UTR of the tryptophan hydroxylase 2 gene alters the mRNA structure and/or RNA-protein interaction, affecting gene expression. However, the functional consequences of variants within UTRs are not always immediately obvious and often not characterized. To date, there are no functional consequences described for the rs2167270 and rs211093, both located in the 5´UTR region; nevertheless, they can be a useful genetic marker to analyze the association of LEP and VEGFA with knee OA. LEP and VEGF-A are implicated in inflammation and angiogenesis, respectively; nevertheless, they may act

Figure 1. (A) Graphical model showing the distribution of knee OA cases and controls for each genotype combination of vascular endothelial growth factor A (VEGFA) and LEP gene variants. The cells shaded dark are high risk for disease and cells shaded light are low risk. Left bars represent the number of cases and right bars represent the number of controls. The new variable constructed by multifactor dimensionality reduction is shown on the right. The pattern of high and low risk for the VEGFA polymorphism differs depending on the value of the LEP polymorphism, which is evidence of epistasis or gene-gene interaction. (B) Interaction graph comprises a node of each gene variant and a pairwise connection between them. The percentage of entropy removed (i.e., information gain) by each variable is visualized for each node shown in a box. The pairwise Cartesian product of variables is visualized for the connection depicted as a line accompanied by a percent of entropy explained by that interaction. The bold line indicates a strong synergistic interaction.
synergistically, favoring both processes as demonstrated in different types of cancer. LEP and VEGF-A are involved in the pathophysiology of OA; it is, therefore, possible that, as in cancer, they could play a synergistic role. During osteoarthritic process, LEP and VEGF-A can modulate the production of IL-1β, IL-6, IL-8, and different MMPs. Strikingly, VEGF-A per se may cause characteristic features of OA since intra-articular injection of VEGF-A in the knee joints of mice induces OA-like changes (i.e., proteoglycan loss, cartilage calcification and degradation, bone sclerosis, osteophyte formation, and synovial hyperplasia).

We are aware that our study could have some limitations. Being a hospital-based case–control study, there is an inherent possibility of incurring in selection bias; therefore, to avoid bias, we assessed incident cases with primary knee OA with their proper controls during the same period of time through specific inclusion criteria, performing the same procedures on both study groups, performing a blinded radiologic evaluation, and rigorously controlling potential confounders through exclusion criteria or through multivariate statistical analysis. Sample size could also be a disadvantage. In this regard, it may be a challenge to achieve a large sample size in hospital-based case–control studies; to compensate for this, we attempted to increase our internal validity through the strategies mentioned above. Obesity is one of the most relevant risk factors associated with OA, mainly with secondary OA. The overload conferred by obesity is an important factor, but there is also a chronic inflammatory state observed in overweight people. Genetic associations in OA could be influenced by differences in BMI, as observed in the study in the Han Chinese population. Subsequently, we controlled for BMI, and the participants were cases and controls which were not overweight. In consequence, we believed that our results are not influenced by overweight. In fact, our patients have lower BMI in comparison to other studies. In summary, while cognizant of our limitations, we are confident that our main strength is that our study groups were adequately selected and that we are properly including primary knee OA.

In conclusion, our results suggest that the independent effects of LEP and VEGFA gene variants are not associated with primary knee OA; nevertheless, the interaction between them is associated with the genetic susceptibility to develop primary knee OA.

ACKNOWLEDGMENTS

The authors wish to thank the radiology support staff for providing the X-ray films for this study. Dr Matvey Sosa-Arellano was supported by grant No. 438437 of the Consejo Nacional de Ciencia y Tecnología (CONACyT), Mexico.

REFERENCES


