Short Report

Association of the estrogen receptor α gene polymorphisms with osteoporosis in the Mexican population


The estrogen receptor gene (ERα) has been implicated in the development of osteoporosis. In this study, the association of two ERα gene polymorphic markers (a TA dinucleotide repeat and a single nucleotide polymorphism, G2014A) with osteoporosis was tested in 70 osteoporotic women, 70 non-osteoporotic women and 500 subjects from the Mexican population. According to the genetic analysis of the Mexican population using eight unlinked polymorphic markers, we found that our population is structured into three subpopulations; therefore, the allele–phenotype relationship was analyzed with a statistical method that considered population stratification. We found that the G2014A polymorphism is associated with the presence of osteoporosis while the TA dinucleotide repeat is not. The G allele and the GG genotype frequencies of the G2014A marker were significantly higher in osteoporotic than in non-osteoporotic women. Likewise, subjects bearing the G allele in heterozygous or homozygous displayed lower values for lumbar bone mineral density and T score than those who did not present any G allele. The effect of confounders for osteoporosis on the association of G allele–osteoporosis was ruled out. In summary, we conclude that the G2014 polymorphism may become a useful marker for genetic studies of osteoporosis in the Mexican population.

Osteoporosis is characterized by low bone mineral density (BMD) and microarchitectural deterioration of bone tissue with increased susceptibility to fracture (1). Osteoporosis is a multifactorial disorder in which aging, hormonal, environmental, and genetic factors are major contributors to its onset and progression (2). Evidence based on twin-family studies and segregation analysis in affected families suggests that up to 75% of the variance in BMD is genetically determined (3–5). Given the complex structure of skeleton, it is likely that many skeletal genes are responsible for this genetic component. In fact, some studies have related bone metabolism with polymorphisms localized in or adjacent to different gene sequences, such as the vitamin D receptor gene (6), the
Estrogen polymorphisms and osteoporosis

Several lines of evidence have shown the important role of the estrogen endocrine system in the regulation of BMD and the occurrence of osteoporosis: (i) estrogen deficiency plays a major role in development of menopause (16), which is characterized by increased bone resorption and a potential occurrence of osteoporosis (17, 18); (ii) the supplement of estrogen significantly prevents osteoporosis (19); (iii) laboratory animals present a significant decrease in bone mass after ovariectomy (20); and (iv) a failure in the conversion of testosterone to estrogen, because of an aromatase deficiency, produces osteoporosis in men (21). The skeletal effects of estrogen are mediated by its binding to specific estrogen receptors (ERs), which are localized both at cytosolic and nuclear level. Two functional ERs (ERα and ERβ), which are encoded by different genes, have been described so far. The ERα gene, which is located on chromosome 6q25-27 (22), seems to be the major receptor mediating estrogen action in bone; in fact, an ERα null provoked severe osteoporosis in a male patient (23).

Genetic screening of the ERα gene locus has revealed the existence of several polymorphic sites (24–27), and a large number of studies analyzing the association of these polymorphisms with osteoporosis in different populations have been performed. However, some of these studies have yielded conflicting results, and the individual contribution of these genetic markers to the osteoporosis pathogenesis remains to be universally confirmed (28). Multiple factors can influence the polymorphism association studies including differences in age, menopausal status, inadequate sample size, ethnicity and the presence of admixture in the population.

In this study, we analyzed the association of two polymorphic markers located in the ERα gene with BMD in a Mexican population, a short tandem repeat polymorphism (TA) situated approximately 1 kbp upstream of exon 1 (29) and a single nucleotide polymorphism (SNP) located at exon 8 (G2014A) (30). As Mexican population consists of mestizos, individuals with an ancestry mixture of European, native American and African genes (31, 32), we analyzed the structure of the Mexican population before proceeding with the association study. By screening 500 unrelated individuals with eight unlinked polymorphic markers, the Mexican population was found to be structured into three subpopulations. Therefore, the study of allele–phenotype association was performed with the statistical method structured population association test (STRAT), which corrects for population stratification (33). We found that the G allele of the G2014A polymorphism is associated with the presence of osteoporosis in the Mexican population, while the TA dinucleotide polymorphism is not.

Materials and methods

Subjects

The studied population was composed of unrelated individuals native from Mexico City that provided informed consent to participate in this study and whose families have lived in Mexico for at least three generations. Mexican population was subdivided in two groups: the first group composed of 70 osteoporotic women and 70 non-osteoporotic women, recruited by the National Rehabilitation Institute of Mexico City, was used for the allele–phenotype association studies; and the second group composed of 500 blood donor subjects from the general Mexican population (mean ± SD age, 34.3 ± 10.2 years) was employed to determine the allelic and genotypic frequencies of the ERα gene markers. The general Mexican population sample was obtained from the blood reference Center of the Genomic Medicine Department of Biomedical Investigation Institute – UNAM, which recruited individuals from the different geographic zones of Mexico; therefore, our sample is a representative group of the Mexican population. The identification of Mexican osteoporotic and non-osteoporotic women was previously reported by Magaña et al. (7) following the criteria of the World Health Organization.

Women were interviewed by trained personnel to record information about tobacco and alcohol consumption, physical activity, steroid and hormone consumption, and parity. The study was approved by the Ethics Committee of the National Rehabilitation Institute, Mexico City.

Determination of the ERα gene dinucleotide polymorphism

Genomic DNA was extracted from peripheral blood leukocytes by the phenol-chloroform method (34). Polymerase chain reaction (PCR) was performed with oligonucleotide primers designed to amplify a polymorphic (TA)n repeat in the human ERα gene [Sano et al. (29)]. PCR was performed in a total volume of 6 μl containing 0.1 μM of each primer (Forward, 5′-GAGAGATGATACATTGACCACC-3′; reverse, 5′-GCAGGATCACATTACCAACC-3′), 29 ng of human genomic DNA; 200 μM of each of the four
dNTPs; 0.6 µl of 10× reaction buffer, 2 mM of MgCl₂, and 0.05 U of Taq DNA polymerase (Roche Diagnostics GmbH, Mannheim, BW, Germany). The forward primer was an 5'-hexachloro-fluorescein-labeled oligonucleotide primer (Accessolab, Mexico City, Mexico). The thermocycling procedure consisted of 30 cycles of denaturation at 94°C for 2 min, annealing at 60°C for 1 min and extension at 74°C for 1 min. The PCR product was electrophoresed on 3% agarose gel to verify the reaction. An aliquot of the reaction was mixed with deionized formamide and the internal size standard (ABI GeneScan-500 TAMRA), heated at 95°C for 7 min, and cooled on ice for 5 min. Samples were then electrophoresed through capillary on the ABI PRISM 310 Genetic Analyzer (Applied Biosystems, Foster City, CA). Electrophoresis was carried out at 15 kV for 24 min at a constant temperature of 60°C. Analysis and genotyping were carried out using the GENSCAN software.

DNA sequencing

The number of repeats present in each allele was verified by DNA sequencing. Homozygous alleles were sequenced with the BigDye Terminator Cycle Sequencing Kit (Applied Biosystems) and analyzed on the ABI PRISM 310 Genetic Analyzer (Applied Biosystems). Heterozygous alleles were analyzed by PCR amplification, as described above, and then PCR products were purified with the QIAquick PCR purification Kit (Qiagen, Hilden, NRW, Germany) and sequenced on the Genetic Analyzer.

Determination of G2014A SNP

Genotyping was carried out with the 5'-nuclease assay with minor groove binder probes fluorescently labeled with 6-carboxyfluorescein or VIC™ and using the protocol recommended by the supplier (Applied Biosystems). Reactions were run in 96-well plates, and the PCR reaction mixture included 20 ng of genomic DNA in a 10 µl reaction volume and the following concentrations of other reagents: probes (100 nM each), primers (900 nM each), and 1× TaqMan PCR Master Mix (Applied Biosystems). PCR cycling conditions consisted of a pre-incubation period of 2 min at 50°C, an initial denaturation period of 10 min at 95°C, an annealing period of 30 s at 62°C, and denaturation period of 30 s at 95°C for 45 cycles. The ABI Prism 7700 sequence detector (Applied Biosystems) was employed for data acquisition.

Population stratification

Genotyping was performed with eight polymorphic markers: four SNP markers (rs1800795, rs1800796, rs2228480 and rs2238136) and four-microsatellite markers [AFM165xb10, GDB: 611352, GDB:196364 and a CA dinucleotide repeat in 11p15.1 (35)]. The SNP markers were analyzed in the ABI Prism 7700 sequence detector (Applied Biosystems), whereas the alleles of microsatellite markers were analyzed using a capillary electrophoresis on the ABI PRISM 310 Genetic Analyzer (Applied Biosystems). The program STRUCTURE (University of Chicago, Chicago, IL) (36), available from http://www.stats.ox.ac.uk/~pritch/home.html, was used to identify population subgroups and infer admixture information from genotype data. This method uses genotypic correlations among unlinked markers to learn about the structure of the population under study and about the genetic background of the sampled individuals. Genetic associations among unlinked markers are assumed to be the result of population structure, and in agreement genetic background, individuals are assigned to subpopulations, diminished the genetic structure and eliminate these associations. The program models the sampled individuals as having inherited their genes from a pool of K ‘unstructured’ populations (where K may be unknown). The subpopulations are ‘unstructured’ and in each subpopulation all loci are in Hardy–Weinberg equilibrium (HWE) with no linkage disequilibrium between loci provided they are not tightly linked. The allele frequencies at each locus within each population are assumed to be unknown. Each individual’s genetic background is represented by a vector, \( q = q_1, \ldots, q_K \), where \( q_k \) is the proportion of the individual’s genome that originated in the subpopulation k. The program uses the Markov chain Monte Carlo (MCMC) method to estimate the number of subpopulation, the allele frequencies in each subpopulation, and the value of \( q \) for each sampled individual. All runs were performed with 150,000 length of burn-in period, and 700,000 number of MCMC repeats after burn in.

Statistical analysis

All results are expressed as mean ± SD. To examine allele and genotype frequencies, differentiation between pair of groups defined by disease status was computed by Chi-squared test (\( \chi^2 \)); simple and multiple logistic regression models were adjusted by confounders using the STATA 8.0 software package (Stat Corporation,
College Station, TX). The odds ratio analysis was used to determine the association at the ERα locus with case-control phenotypes and subsequently, the STRAT software (University of Oxford, Oxford, UK) (33) with correction for three sub-populations was employed; the runs were performed assuming the following population structure (150,000 length of burn-in period and 700,000 number of MCMC repeats after burn in). Differences in individual characteristics between osteoporotic and non-osteoporotic groups were tested using Student’s t-test for unpaired data. Analysis of the deviation of frequencies from HWE by the Weir and Cockerham Fis tests and the genetic analysis (Shannon’s index and heterozygosity) for each population were performed with the GENETIX software (University of Montpellier, Montpellier, France) (37). The levels of significance were determined empirically with 10,000 permutations.

Results

To genotype the TA repeat and the G2014A polymorphisms of the ERα gene, the Mexican population under study was classified into the osteoporotic group (70 osteoporotic women), the control group (70 non-osteoporotic women) and the general population group (500 subjects). We previously reported the identification of Mexican osteoporotic and non-osteoporotic women (7).

The allelic distribution of the TA repeat marker in the Mexican population is shown in Fig. 1. We identified 22 different alleles in the general population containing 8–29 TA repeats and ranking from 156 to 198 bp. Alleles designed from letter A to letter T were previously reported (29), whereas three newly identified alleles (156, 158 and 196 bp) were named herein as A–, A– and S, respectively. The number of repeats of each identified allele was confirmed by sequencing. The E allele (14 TA repeats) was the most frequent, followed by D, N and F alleles (13, 23 and 15 TA repeats, respectively). As expected from such allelic distribution, the predominant genotype was EE (9%) followed by genotypes ED (6%), NE (6%) and FD (5%) (data not shown). In addition, this polymorphic marker displayed high genetic diversity (Shannon’s index of 2.4), and high heterozygosity (0.9348). From the set of 22 alleles, only 15 were present in both the osteoporotic and the non-osteoporotic groups, however, no statistically significant differences in the TA repeat marker allelic frequency were observed between them (Fig. 1).

With respect to the G2014A SNP polymorphism, the G allele (69.7%) and the GG genotype (47.4%) were predominant in the general population (Table 1); interestingly, frequencies for the G allele (77.9% vs 62.1%, \( p = 0.006 \)) and GG genotype (60% vs 38.6%, \( p = 0.017 \)) were clearly higher in osteoporotic than in non-osteoporotic women (Table 1).

A pre-requisite for a reliable allele–phenotype association study is that the allelic distribution of the polymorphic marker conforms to the HWE. Noticeably, we found that both the TA repeat (\( F_{\text{is}} = 0.00762 \)) and the G2014A (\( F_{\text{is}} = -0.0067 \)) markers displayed deviation from the HWE in the general population (\( p = 0.0003 \)). As deviation from the HWE could be caused by the presence of admixture in the population, among other factors (migration, mutation, etc.), we decided to test the general Mexican population (500 subjects) for stratification using eight different unlinked polymorphic markers and the program STRUCTURE. This method estimates the number of subpopulations based on the allele frequencies of each marker and the genetic background of each sampled individual. The program assumes that linkage among unlinked markers indicates population stratification; therefore, to eliminate these genetic associations the program assigns individuals to different subpopulations (see Materials and Methods). Under this analysis, the Mexican population was classified into three sub-populations: osteoporotic, non-osteoporotic and the general population (Table 1).

![Fig. 1. Allelic distribution of the (TA) repeat polymorphism of the ERα gene in 70 osteoporotic women (140 alleles), 70 non-osteoporotic women (140 alleles) and 500 subjects from the Mexican general population (1000 alleles).](image-url)
individuals were clustered into three different subpopulations ($p = 1$): subpopulation 1 ($F_{st} = 0.1558$), subpopulation 2 ($F_{st} = 0.1181$) and subpopulation 3 ($F_{st} = 0.0590$) (Fig. 2; the vertices of the triangle represents the three subgroups). Next, we analyzed whether the ER<sub>a</sub> markers conform to the HWE within each of the three Mexican subpopulations. Both the G2014A SNP polymorphism (subpopulation 1: $F_{is} = 0.0316$, $p = 0.561$; subpopulation 2: $F_{is} = 0.0501$, $p = 0.204$; subpopulation 3: $F_{is} = 0.253$, $p = 0.998$) and the TA repeat marker (subpopulation 1: $F_{is} = 0.0016$, $p = 0.402$; subpopulation 2: $F_{is} = 0.0501$, $p = 0.204$; subpopulation 3: $F_{is} = -0.04362$, $p = 0.9124$) were in HWE. Therefore, a model with three subpopulations was employed to evaluate residual allelic associations with osteoporosis using the STRAT software. For the TA repeat polymorphism, we observed no significant differences in the allelic frequencies between osteoporotic and non-osteoporotic women ($p = 0.15$), whereas for the G2014A marker we did found association of the G allele with osteoporosis ($p < 0.0001$). As shown in Table 1, the risk to present osteoporosis, which was associated with the G allele, increased when this allele is homozygous (the odds ratio values were adjusted with the following confounders for osteoporosis: age, BMI, smoking, alcohol consumption, physical activity, hormone and steroid consumption and parity). Likewise, subjects who possessed two G alleles had lower BMD and <i>t</i> score average values than those who possessed one or did not present any G allele (lumbar BMD: 79.86%, 84.59%, and 87.08%; <i>T</i> score: −2.04, −1.26, and −1.05, respectively).

Discussion

In this study, we analyzed the association of osteoporosis with two polymorphic markers located at the ER<sub>a</sub> gene (a TA repeat and an SNP G2014A polymorphisms) in a Mexican population. To accomplish that, we first determined the distribution of these markers in our population: for the TA dinucleotide marker, we found that E allele and EE genotype were the most frequent; for the G2014A polymorphism, we observed that G allele and GG genotype were predominant.

Table 1. Allelic and genotypic distribution of the G2014A single nucleotide polymorphism (SNP) polymorphism and its effect on osteoporosis

<table>
<thead>
<tr>
<th></th>
<th>Osteoporotic (%)</th>
<th>Non-osteoporotic (%)</th>
<th>General population (%)</th>
<th>p-Value (osteoporotic vs non-osteoporotic)</th>
<th>Odds ratio</th>
<th>95% CI</th>
</tr>
</thead>
<tbody>
<tr>
<td>G allele</td>
<td>77.9</td>
<td>62.1</td>
<td>69.7</td>
<td>0.006&lt;sup&gt;a&lt;/sup&gt;</td>
<td>4.34</td>
<td>1.04–18.14</td>
</tr>
<tr>
<td>A allele</td>
<td>22.1</td>
<td>37.9</td>
<td>30.3</td>
<td>0.006&lt;sup&gt;a&lt;/sup&gt;</td>
<td>0.46</td>
<td>0.26–1.62</td>
</tr>
<tr>
<td>GG genotype</td>
<td>60.0</td>
<td>38.5</td>
<td>47.4</td>
<td>0.017&lt;sup&gt;a&lt;/sup&gt;</td>
<td>2.21</td>
<td>1.06–4.94</td>
</tr>
<tr>
<td>GA genotype</td>
<td>37.5</td>
<td>47.1</td>
<td>44.6</td>
<td>0.039&lt;sup&gt;a&lt;/sup&gt;</td>
<td>0.62</td>
<td>0.11–19.00</td>
</tr>
<tr>
<td>AA genotype</td>
<td>4.3</td>
<td>14.3</td>
<td>8.0</td>
<td>0.083</td>
<td>0.03</td>
<td>0.12–19.87</td>
</tr>
</tbody>
</table>

Allelic and genotypic distribution of the G2014A SNP polymorphism if ER<sub>a</sub> gene in 70 osteoporotic women, 70 non-osteoporotic women and 500 subjects from the general Mexican population. Differences in the allele and genotype distribution were assessed by the chi-squared test ($\chi^2$). For the odds ratio, the models were adjusted by age, BMI, smoking, alcohol consumption, physical activity, hormone and steroid consumption and parity.

CI, confidence interval.

Significant values ($p < 0.05$) are indicated in bold.

<sup>a</sup>Significant differences among the three groups at $p < 0.05$.  

Fig. 2. Triangle plot showing the structure of the Mexican population. Five hundred subjects from the Mexican general population were genotyped with eight different polymorphic markers and the number of Mexican subpopulations was estimated using the program Structure, as described in Material and Methods. Mexican population was divided into three subpopulations, which are depicted by the triangle vertices. According to genetics features, each individual (represented by a single point) was assigned to one of the subpopulations.
Surprisingly, the allelic distribution of both polymorphic markers displayed a deviation from the HWE, which may indicate inbreeding, mutation, natural selection, migration and/or population stratification (37, 38). The majority of the contemporary Mexican population consists of mestizos, an admixture group with genetic background derived from the original native American inhabitants of Mexico, the European settlers (primarily from Spain) who arrived after the conquest of Mexico by Cortes in the 16th century, and to a lesser extent, West Africans who were brought to Mexico mainly during the 16th and 18th centuries as a consequence of the slave trade in the American Continent (31, 32, 39). Therefore, it is likely that Mexican population is indeed an admixed population. Considering this scenario, we decided to analyze the structure of the Mexican population using eight unlinked polymorphic markers, before proceeding with the association studies. The Mexican population was found to be structured into three different subpopulations, which is in agreement with a previous study that analyzed the structure of a population from Mexico City with 69 autosomal ancestry-informative markers and determined that it is also composed of three subpopulations, to which the European and the native American groups make the major genetic contribution (40).

It is well known that population subdivision can result in association between a genotype and markers that are not linked to any causative loci. Such association can occur when the disease frequency varies across subpopulations, thereby increasing the probability that affected individuals will be sampled from a particular subpopulation. Any marker allele that is in high frequency in the overexpressed subpopulation will then be associated with the phenotype (33, 41). In response to this problem, we tested our population for phenotype–allele association using an approach that provides an adequate correction for population stratification, the STRAT program (36, 42, 43). Under this analysis, we found that the G2014A marker is associated with osteoporosis in the Mexican population while the ERα gene TA repeat polymorphism is not.

The TA repeat marker has been broadly used in association studies for osteoporosis; originally, a correlation between the (TA)$_n$ polymorphism and BMD was described in a Japanese population (8, 29). Likewise, two large-scale studies performed in Italian and Dutch populations, found a statistically significant correlation between the TA repeat length and fracture risk: women with a low number of TA repeats showed lower BMD and higher incidence of vertebral fractures, compared with women having high number of TA repeats (44–46). A similar association was described in pre- and perimenopausal women from Michigan (47, 48). Surprisingly, an opposite association between the TA repeat marker and BMD was observed in Chinese women where the presence of high number of repeats correlated with low BMD values (49). Finally, in correlation with the present study, no overall association between the TA marker and BMD was observed in post-menopausal women from Scotland and Korea (50, 51). With respect to the G2014A SNP polymorphism, we found that frequencies for the G allele and GG genotype were significantly higher in osteoporotic than in non-osteoporotic Mexican women, which suggests an association between the presence of the G allele and osteoporosis. Supporting this interpretation, we observed that women homozygous for the G allele showed lower lumbar BMD and $T$ score values than women with AG and AA genotypes. As the G2014 SNP polymorphism does not alter the amino acid sequence of the ERα gene, it could be speculated that this silent polymorphism is linked with a mutation in an encoding or regulatory sequences of the ERα gene or of the neighboring genes (52–54), which may ultimately cause osteoporosis. Further molecular studies are required to clarify this point. In contrast with this study, an association of the A allele with the presence and severity of osteoporosis was revealed in Thai post-menopausal women (30). The reason for these discrepancies is not clear; however, it may be based on ethnic and environmental factors as well as on the design of each study. We noticed that previous reports were limited to the analysis of post-menopausal women cohort (29, 45–51), while ours was a case–control study based on the diagnosis of osteoporosis rather than the hormonal condition. Although, menopause has a substantial effect on bone loss (55) our study included both post-menopausal osteoporosis and primary osteoporosis. In addition, the present work included the population analysis of the ERα markers while the others did not, which offers external validation to the allele–phenotype association. Finally, our study was smaller in scale compared with previous reports; therefore, to obtain conclusive results about the association of the ERα markers with osteoporosis, it is necessary to increase the size of the Mexican population sample.

In summary, we conclude that the G2014 polymorphism may become a useful marker for genetic studies of osteoporosis in the Mexican population. However, predictive therapy for osteoporosis may consider the analysis of several markers associated with osteoporosis in a larger
References