Estrogen receptor 1 gene polymorphisms in premenopausal women: interaction between genotype and smoking on lipid levels

S. Almeida¹,² and M.H. Hutz¹

¹Departamento de Genética, Instituto de Biociências, Universidade Federal do Rio Grande do Sul, Porto Alegre, RS, Brasil
²Departamento de Ciências Fisiológicas, Fundação Universidade Federal de Ciências da Saúde de Porto Alegre, Porto Alegre, RS, Brasil

Correspondence to: M.H. Hutz, Departamento de Genética, UFRGS, Caixa Postal 15053, 91501-970 Porto Alegre, RS, Brasil
Fax: +55-51-3316-7311. E-mail: mara.hutz@ufrgs.br

Estrogen has multiple effects on lipid and lipoprotein metabolism. We investigated the association between the four common single nucleotide polymorphisms in the estrogen receptor 1 (ESR1) gene locus, -1989T>G, +261G>C, IVS1-397T>C and IVS1-351A>G, and lipid and lipoprotein levels in southern Brazilians. The sample consisted in 150 men and 187 premenopausal women. The women were considered premenopausal if they had regular menstrual bleeding within the previous 3 months and were 18-50 years of age. Exclusion criteria were pregnancy, secondary hyperlipidemia due to renal, hepatic or thyroid disease, and diabetes. Smoking status was self-reported; subjects were classified as never smoked and current smokers. DNA was amplified by PCR and was subsequently digested with the appropriate restriction enzymes. Statistical analysis was carried out for men and women separately. In the study population, major allele frequencies were -1989*T (0.83), +261*G (0.96), IVS1-397*T (0.58), and IVS1-351*A (0.65). Multiple linear regression analyses indicated that an interaction between +261G>C polymorphism and smoking was a significant factor affecting high-density lipoprotein cholesterol (HDL-C) levels (P = 0.028) in women. Nonsmoking women with genotype G/C of +261G>C polymorphism had mean HDL-C levels higher than those with G/G genotype (1.40 ± 0.33 vs 1.22 ± 0.26 mmol/L; P = 0.033). No significant associations with lipid and lipoprotein levels in women and men were detected for other polymorphisms. In conclusion, the +261G>C polymorphism might influence lipoprotein and lipid levels in premenopausal women, but these effects seem to be modulated by smoking, whereas in men ESR1 polymorphisms were not associated with high lipoprotein levels.

Key words: Estrogen; Estrogen receptor 1 variants; Lipid levels; Smoking

Research supported by Programa de Apoio a Núcleos de Excelência (PRONEX #4/0895-9, Brasil) and CNPq (#471256/2006-4).

Received April 22, 2008. Accepted October 21, 2008

Introduction

It is estimated that beneficial alterations on lipoprotein levels account for one third of the decline in cardiovascular disease (CVD) incidence. Physiological studies have indicated beneficial effects of estrogen on lipoprotein metabolism, with a decrease in total cholesterol (T-chol) and low-density lipoprotein cholesterol (LDL-C), and an increase of high-density lipoprotein cholesterol (HDL-C) in plasma (1). It has been demonstrated that the cholesterol-lowering effect of estrogen in mice depends on the presence of estrogen receptor alpha (ESRα) (2). The human ESR1 gene is located at 6q25.1 and has eight exons and seven introns. About 2234 single nucleotide polymorphisms (SNPs) in this gene were reported in the dbSNP database of the National Center for Biotechnology. In the present study, we investigated the association between four common SNPs in the ESR1 gene locus, -1989T>G, +261G>C
Material and Methods

The sample consisted of 337 subjects of European descent, 150 men and 187 premenopausal women, who were selected randomly at the Clinical Analysis Laboratory of the Pharmacy School of the Federal University of Rio Grande do Sul where they went for routine blood tests. Only premenopausal women were selected for the present study because they have abundant endogenous estrogen levels, since estrogen receptors are ligand-active transcription factors dependent on the presence of the hormone. The women were considered premenopausal if they had regular menstrual bleeding within the previous 3 months and were 18-50 years of age. Informed consent for a blood sample drawn for DNA extraction to be used in studies approved by the University Ethics Committee was obtained from each subject included in the sample. Information about health and lifestyle factors (physical activity, smoking status, alcohol consumption and hormone or drug intake) was obtained from each individual by a questionnaire. After answering the questionnaire, subjects wearing light clothes had their body weight and height recorded. Body mass index (BMI) was calculated as the ratio of weight (in kg) to height squared (in meters). Smoking status was self-reported; subjects were classified as never smokers, former smokers or current smokers. Former smokers were excluded for these analyses. Only those individuals without significant disease were included in the sample. Exclusion criteria were pregnancy, secondary hyperlipidemia due to renal, hepatic or thyroid disease, and diabetes or fasting blood glucose levels above 6.9 mmol/L (3). Individuals on lipid-lowering medication were also excluded. The characteristics of men and premenopausal women included in the sample are shown in Table 1.

Blood samples were collected from subjects after at least 12-h fasting. T-chol, HDL-C, TG, and glucose were determined by standard methods using commercial kits (Labtest®, Brazil) on a Selectra autoanalyser. LDL-C was calculated according to the Friedwald formula (4) if plasma TG were not more than 4.48 mM. When plasma TG were above this level, LDL-C cannot be estimated. Genomic DNA was extracted from peripheral blood leukocytes by a salting-out procedure. DNA was amplified by polymerase chain reaction, the amplification products were subsequently digested with restriction enzymes by a method recommended by the manufacturer: PvuII (IVS1-397T>C), XbaI (IVS1-351A>G), HhaI (-1989T>G) and BstUI (+261G>C). Genotypes were determined after electrophoresis on agarose gels containing ethidium bromide, using a 50-bp ladder to score the band sizes. The fragment sizes for each genotype were: IVS1-397T>C: CC 1350 bp, CT 1350, 900 and 450 bp, TT 900 and 450 bp; IVS1-351A>G: GG 1350 bp, GA 1350, 950 and 400 bp, AA 950 and 400 bp; -1989T>G: TT 230 bp, TG 230, 203 and 27 bp, GG 203 bp; +261G>C: GG 1347 bp, GC 1350, 900 and 450 bp, AA 950 and 400 bp, CC 87 and 54 bp.

Allele frequencies were estimated by gene counting. A χ² test for goodness of fit was used to verify whether the observed genotype frequencies agreed with those expected under Hardy-Weinberg equilibrium. The maximum likelihood estimate of haplotype frequencies was calculated from multisite marker data using the Multiple Locus Haplotype Analysis version 2.0 (5,6). Linkage disequilibrium was tested by a χ² goodness of fit test using Arlequin program version 2.000 (7). D’ values, the relative magnitude of linkage disequilibrium, were calculated as described previously (8). The means were compared by the Student t-test using the Graph Pad InStat software version 2.04a (Graph Pad Software, USA).

All other tests and transformations were performed with SPSS 8.0.0 statistical package. For association with lipid and lipoprotein levels, women and men were analyzed separately. Natural logarithm transformation was used for TG levels and BMI in order to avoid skewness for statistical testing, but non-adjusted levels are shown in Tables. To identify possible interactions between genetic and environmental variables, multiple linear regressions.
were carried out for men and women separately using a forward variable selection. The variables were chosen according to their influence on lipid levels in previously published data. The variables tested were the four ESR1 polymorphisms, coded as carrier and non-carrier of each allele, age, BMI, current smoking, TG levels and interaction between these variables with ESR1 genotypes.

Results

In the study population, major allele frequencies were -1989*T (0.83), +261*G (0.96), IVS1-397*T (0.58) and IVS1-351*A (0.65). The genotype frequencies were not statistically different compared with those expected under Hardy-Weinberg equilibrium. A consistent linkage disequilibrium was detected between four pairwise combinations of ESR1 gene polymorphisms, -1989T>G and +261G>C was 0.85 (χ²GL=1 = 57.72; P < 0.001), -1989T>G and IVS1-351A>G was 0.66 (χ²GL=1 = 18.3; P < 0.001), +261G>C and IVS1-397T>C was 0.75 (χ²GL=1 = 12.5; P < 0.001), and IVS1-397T>C and IVS1-351A>G was 0.94 (χ²GL=1 = 323.70; P < 0.001). Eleven haplotypes were detected, the two most common haplotypes being -1989*T, +261*G, IVS1-397*T, IVS1-354*A and -1989*T, +261*G, IVS1-397*C, IVS1-351*G, which accounted, respectively, for 47 and 31% of the chromosomes investigated. The multiple linear regression for the HDL-C model demonstrated significant regression coefficients by age (P = 0.014), TG levels (P < 0.001) and ESR1+261G>C* smoking interaction (P = 0.024; Table 2). In non-smoker women with a heterozygous genotype at the +261G>C site, the mean HDL-C level was 1.40 ± 0.33 mmol/L and in those with a G/G genotype it was 1.22 ± 0.26 mmol/L (P = 0.033). The four SNPs studied did not influence lipoprotein levels in men, neither as single variables, nor as interaction variables. No significant associations with lipid and lipoprotein levels in women and men were shown at the haplotype level (data not shown).

Discussion

Although several SNPs were reported in ESR1 gene, most studies have focused on IVS1-351A>G and IVS1-397T>C variants in diverse ethnic groups in both genders. Three large, well-designed and well-implemented studies were published with apparent contradictory results in relation to ESR1 IVS1-397T>C polymorphism and myocardial infarction (9-11). In relation to ESR1 IVS1-397T>C polymorphism and lipid levels, Demissie et al. (12) demonstrated an association between IVS1-397T>C polymorphism T/T genotype and LDL-C particle size reduction in women from the Framingham Heart Study. Molvarec et al. (13) reported that T-chol concentrations were significantly higher in healthy Caucasian premenopausal women with the C/C genotype than in those with the T/T or T/C genotype at the IVS1-397T>C polymorphism. In addition to these large studies, other investigations have reported associations between ESR1 polymorphisms and lipid levels with conflicting results (14). These contradictory and inconsistent data might be explained if the effects of IVS1-397T>C polymorphism on cardiovascular disease or cardiovascular disease markers are small and/or influenced by other genetic or environmental factors. This situation may occur when multiple loci act in concert to cause a disease; therefore, single loci associations may be confounded by other loci or by environmental factors. These contradictory results can occur not only when a multilocus effect is not accounted for, but they may also be attributable to investigation of a non-causal variant in linkage disequilibrium with a genuine causal variant (15). In the present study, the IVS1-397T>C polymorphism was in linkage disequilibrium with +261G>C (D' = 0.75; P < 0.001). Although the functional implications of +261G>C polymorphism have not been fully elucidated, this is a synonymous mutation, the wild sequence (+261*G allele) corresponds to the one that codes least frequently for alanine in humans (7.1% of cases) and the synonymous mutated codon (+261*C allele) corresponds to the codon that codes most

| Table 2. Regression analysis predicting HDL-C levels from age, oral contraceptive usage, TG levels, BMI, current smoking, and ESR1 +261G>C polymorphism in women. |
|------------------------|--------|----------|-------|--------|--------|--------|--------|--------|--------|
| HDL-C                  | 24.1%  |          |       |        |        |
| Age                    | 0.198  | 0.014    | 3.4%  |        |        |
| Oral contraceptive     | 0.049  | 0.519    | 0.2%  |        |        |
| Current smoking        | -0.088 | 0.219    | 0.9%  |        |        |
| BMI                    | -0.134 | 0.062    | 2.0%  |        |        |
| TG                     | -0.382 | <0.001   | 13.5% |        |        |
| +261G>C* age           | 1.758  | 0.273    | 0.7%  |        |        |
| +261G>C* oral contraceptive | -0.161 | 0.112  | 1.4%  |        |        |
| +261G>C* smoking       | -0.304 | 0.224    | 2.9%  |        |        |
| +261G>C* BMI           | -2.744 | 0.063    | 0.9%  |        |        |
| +261G>C* TG            | 1.559  | 0.211    | 2.0%  |        |        |

HDL-C = high-density lipoprotein cholesterol; BMI = body mass index; TG = triglycerides. *Interaction variables; standardized coefficients; coded as one when hormonal contraceptive users; coded as one when current smokers; variable ln transformed; subjects carrying at least one +261*C allele are coded as one, while subjects without this allele are zero.
frequently (29.5% of cases) for this amino acid (16), which may cause a different gene expression pattern. The synonymous changes can also alter the predicted messenger RNA (mRNA) folding and lead to a decrease in mRNA stability and translation (16).

In a previous study, we detected an association of +261*C allele with cardiovascular disease prevalence in men (17); therefore, we hypothesize that the association observed might occur due to an association between this SNP with lipid levels. In the present study, we detected a significant smoking *+261G>C interaction on HDL-C levels. Recently, Shearman et al. (18) reported an association between IVS1-397T>C and IVS1-351A>G polymorphisms and low LDL-C levels in smoking women. Several studies have shown that smoking causes alteration in estrogen metabolism and increased lipid oxidation. Freeman et al. (19) reported that smoking cigarettes has a major effect on plasma lipoprotein metabolism through multiple mechanisms. Hepatic lipase is activated and lecithin cholesterol acyl transferase is inhibited in smokers; these plasma enzymes are known to affect HDL metabolism. Additionally, estradiol level variation, within the physiological range, influences plasma Apo AI, which is the main HDL-C apolipoprotein constituent. Meek and Finch (20) demonstrated that some of the compounds in tobacco smoke may bind to estrogen receptor 1, as an antagonist creating an anti-estrogenic effect.

The action of estrogen is mediated by estrogen receptors. Some of the compounds in tobacco smoke may activate estrogen receptor 1, and consequently smoking may modify ESR-regulated gene induction. The +261G>C polymorphism might cause a different ESR1 expression pattern and, consequently, the compounds in tobacco smoke may have different effects in female carriers of different ESR1 genotypes. Our data suggest that the +261G>C polymorphism might influence HDL-C levels in premenopausal women, but this effect seems to be modulated by smoking. Clearly, further studies are necessary before reaching definitive conclusions about the role of these polymorphisms on lipoprotein metabolism. Our results indicate that environmental and genetic interactions should be considered in association studies of complex traits.

Acknowledgments

Thanks are due to Ana Lúcia S. Antunes and Maria Perpétua de O. Pinto from the Clinical Analysis Laboratory of the Pharmacy College. We are also grateful to André Vargas, Fabiana M. de Andrade, Fabiano Roldão Silveira, Marilu Fiegenbaum, Marcel Arsand, and Vanessa S. Mattevi for help in sample collection, and to Dr. Sidia M. Callegari-Jacques for her advice on the statistical analysis.

References

levels in women: the Framingham Heart Study. *Atherosclerosis* 2006; 185: 210-218.