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

S. Almeida[1,2] and M.H. Hutz[1]

[1]Departamento de Genética, Instituto de Biociências, Universidade Federal do Rio Grande do Sul, Porto Alegre, RS, Brasil
[2]Departamento 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

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*

(rs746432), *IVS1-397T>C* (rs2234693), and *IVS1-351A>G* (rs9340799) with T-chol, LDL-C, HDL-C and triglyceride (TG) levels in order to determine whether these lipid levels are influenced by *ESR1* variation or by the interaction between genetic and environmental variables.

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 smoked and current smokers. Former smokers were ex-

cluded 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 under conditions recommended by the manufacturer: *Pvu*II (*IVS1-397T>C*), *Xba*I (*IVS1-351A>G*), *Hha*I (*-1989T>G*) and *Bst*UI (*+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 and 27 bp; *+261G>C: GG* 141 bp, *GC* 141, 87 and 54 bp, *CC* 87 and 54 bp.

Allele frequencies were estimated by gene counting. A χ^2 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 χ^2 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

Table 1. Characteristics of the study group.

	Premenopausal women (N = 187)	Men (N = 150)	P
Age (years)	34 ± 10.2	42 ± 14.0	<0.001[a]
T-chol (mmol/L)	4.93 ± 0.92	5.17 ± 1.12	0.009[a]
HDL-C (mmol/L)	1.21 ± 0.29	1.06 ± 0.27	<0.001[a]
LDL-C (mmol/L)	3.20 ± 0.84	3.31 ± 0.99	0.018[a]
TG (mmol/L)	1.18 ± 0.66	1.70 ± 1.23	<0.001[a]
BMI (kg/m^2)	25.8 ± 5.2	26.3 ± 4.3	0.170[a]
Sedentary	67.4%	52.1%	0.002[b]
Current smokers	28.3%	41.3%	0.240[b]
Hypertension	16.7%	16.8%	0.770[b]
Oral contraceptive	32%	–	–

Unadjusted data. Data are reported as means ± SD or percent. T-chol = total cholesterol; HDL-C = high-density lipoprotein cholesterol; LDL-C = low-density lipoprotein cholesterol; TG = triglycerides; BMI = body mass index. [a]Student *t*-test and [b]chi-square test for comparison of men to premenopausal women.

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 ($\chi^2_{GL=1}$ = 57.72; P < 0.001), *-1989T>G* and *IVS1-351A>G* was 0.66 ($\chi^2_{GL=1}$ = 18.3; P < 0.001), *+261G>C* and *IVS1-397T>C* was 0.75 ($\chi^2_{GL=1}$ = 12.5; P < 0.001), and *IVS1-397T>C* and *IVS1-351A>G* was 0.94 ($\chi^2_{GL=1}$ = 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.

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			
	β^a	P	Partial R^2 (%)	R^2 (%)
Total				24.1%
Age	0.198	0.014	3.4%	
Oral contraceptive[b]	0.049	0.519	0.2%	
Current smoking[c]	-0.088	0.219	0.9%	
BMI[d]	-0.134	0.062	2.0%	
TG[d]	-0.382	<0.001	13.5%	
+261G>C[e]	1.758	0.273	0.7%	
*+261G>C** age	-0.275	0.275	0.7%	
*+261G>C** oral contraceptive	-0.161	0.112	1.4%	
*+261G>C** smoking	-0.304	0.024	2.9%	
*+261G>C** BMI[d]	-2.744	0.063	0.9%	
*+261G>C** TG[d]	1.559	0.211	2.0%	

HDL-C = high-density lipoprotein cholesterol; BMI = body mass index; TG = triglycerides. *Interaction variables; [a]standardized coefficients; [b]coded as one when hormonal contraceptive users; [c]coded as one when current smokers; [d]variable In transformed; [e]subjects carrying at least one *+261*C* allele are coded as one, while subjects without this allele are zero.

(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

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

References

1. Campos H, Walsh BW, Judge H, Sacks FM. Effect of estrogen on very low density lipoprotein and low density lipoprotein subclass metabolism in postmenopausal women. *J Clin Endocrinol Metab* 1997; 82: 3955-3963.

2. Ohlsson C, Hellberg N, Parini P, Vidal O, Bohlooly Y, Rudling M, et al. Obesity and disturbed lipoprotein profile in estrogen receptor-alpha-deficient male mice. *Biochem Biophys Res Commun* 2000; 278: 640-645.

3. Report of the Expert Committee on the Diagnosis and Classification of Diabetes Mellitus. *Diabetes Care* 1997; 20: 1183-1197.

4. Friedwald WT, Levy RI, Fredrickson DS. Estimation of the concentration of low-density lipoprotein cholesterol in plasma, without use of the preparative ultracentrifuge. *Clin Chem* 1972; 18: 499-502.

5. Long JC. *Multiple locus haplotype analysis*. [Computer program]. Version 2.0. Software and documentation distributed by the author. Section on population genetics and linkage. Bethesda: Laboratory of Neurogenetics, National Institutes of Health; 1999.

6. Long JC, Williams RC, Urbanek M. An E-M algorithm and testing strategy for multiple-locus haplotypes. *Am J Hum Genet* 1995; 56: 799-810.

7. Schneider S, Kueffer J-M, Roessli D, Excoffier L. *Arlequin ver 2.000: a software for population data analysis*. [Computer program]. Geneva: Genetics and Biometry Laboratory, University of Geneva; 2000.

8. Lewontin RC. On measures of gametic disequilibrium. *Genetics* 1988; 120: 849-852.

9. Shearman AM, Cupples LA, Demissie S, Peter I, Schmid CH, Karas RH, et al. Association between estrogen receptor alpha gene variation and cardiovascular disease. *JAMA* 2003; 290: 2263-2270.

10. Schuit SC, Oei HH, Witteman JC, Geurts van Kessel CH, van Meurs JB, Nijhuis RL, et al. Estrogen receptor alpha gene polymorphisms and risk of myocardial infarction. *JAMA* 2004; 291: 2969-2977.

11. Kjaergaard AD, Ellervik C, Tybjaerg-Hansen A, Axelsson CK, Gronholdt ML, Grande P, et al. Estrogen receptor alpha polymorphism and risk of cardiovascular disease, cancer, and hip fracture: cross-sectional, cohort, and case-control studies and a meta-analysis. *Circulation* 2007; 115: 861-871.

12. Demissie S, Cupples LA, Shearman AM, Gruenthal KM, Peter I, Schmid CH, et al. Estrogen receptor-alpha variants are associated with lipoprotein size distribution and particle

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

13. Molvarec A, Nagy B, Kovacs M, Walentin S, Imreh E, Rigo J Jr, et al. Lipid, haemostatic and inflammatory variables in relation to the estrogen receptor alpha (ESR1) *Pvu*II and *Xba*I gene polymorphisms. *Clin Chim Acta* 2007; 380: 157-164.

14. Almeida S, Hutz MH. Genetic variation of estrogen metabolism and the risks of cardiovascular disease. *Curr Opin Investig Drugs* 2007; 8: 814-820.

15. Lin PI, Vance JM, Pericak-Vance MA, Martin ER. No gene is an island: the flip-flop phenomenon. *Am J Hum Genet* 2007; 80: 531-538.

16. Nakamura Y, Wada K, Wada Y, Doi H, Kanaya S, Gojobori T, et al. Codon usage tabulated from the international DNA sequence databases. *Nucleic Acids Res* 1996; 24: 214-215.

17. Almeida S, Hutz MH. Estrogen receptor 1 gene polymorphisms and coronary artery disease in the Brazilian population. *Braz J Med Biol Res* 2006; 39: 447-454.

18. Shearman AM, Demissie S, Cupples LA, Peter I, Schmid CH, Ordovas JM, et al. Tobacco smoking, estrogen receptor alpha gene variation and small low density lipoprotein level. *Hum Mol Genet* 2005; 14: 2405-2413.

19. Freeman DJ, Griffin BA, Murray E, Lindsay GM, Gaffney D, Packard CJ, et al. Smoking and plasma lipoproteins in man: effects on low density lipoprotein cholesterol levels and high density lipoprotein subfraction distribution. *Eur J Clin Invest* 1993; 23: 630-640.

20. Meek MD, Finch GL. Diluted mainstream cigarette smoke condensates activate estrogen receptor and aryl hydrocarbon receptor-mediated gene transcription. *Environ Res* 1999; 80: 9-17.