Genetic Polymorphism *CYP17* rs2486758 and Metabolic Risk Factors Predict Daily Salivary 17β-Estradiol Concentration in Healthy Premenopausal Norwegian Women. The EBBA-I Study

Anita Iversen, Inger Thune, Anne McTiernan, Karen W. Makar, Tom Wilsgaard, Peter T. Ellison, Grazyna Jasienska, Vidar Flote, Elizabeth M. Poole, and Anne-Sofie Furberg

Department of Community Medicine (A.I., I.T., T.W., A.-S.F.), Faculty of Health Sciences, University of Tromsø, 9037 Tromsø, Norway; Department of Oncology (I.T., V.F.), Oslo University Hospital Ullevål, 0450 Oslo, Norway; Fred Hutchinson Cancer Research Center (A.M., K.W.M., E.M.P.), Seattle, Washington 98109; Department of Human Evolutionary Biology (P.T.E.), Harvard University, Cambridge, Massachusetts 02138; Department of Epidemiology and Population Studies (G.J.), Jagiellonian University Medical College, Krakow, Poland 31-531; Channing Laboratory (E.M.P.), Harvard University Medical School and Brigham and Women's Hospital, Boston, Massachusetts 02115; and Department of Microbiology and Infection Control (A.-S.F.), University Hospital of North Norway, 9038 Tromsø, Norway

Context: The relationship between low-penetrance genes, metabolic risk factors, and levels of endogenous 17β-estradiol and progesterone, which play a role in breast cancer risk, remains unclear.

Objective: The aim of this study was to determine whether common polymorphisms in *CYP17*, in combination with metabolic risk factors (individually or clustered), alter salivary concentrations of free biologically active 17β-estradiol and progesterone among healthy premenopausal Norwegian women.

Design: Eight single nucleotide polymorphisms in *CYP17* were genotyped in 203 healthy premenopausal women aged 25–35 yr in the Norwegian EBBA-I Study, conducted in 2000–2002. Daily salivary concentrations of 17β-estradiol and progesterone were measured throughout one menstrual cycle. A clustered metabolic score was calculated, including waist circumference, mean arterial pressure, insulin resistance, fasting triglycerides, and total cholesterol/high-density lipoprotein cholesterol ratio. The study hypothesis was tested in multivariable linear regression and generalized estimating equation models.

Results: Women in the upper tertile of clustered metabolic score with the *CYP17* rs2486758 minor allele had daily salivary 17β-estradiol concentrations that were 53% higher than other study women throughout the menstrual cycle ($P < 0.001$). Similarly, women in the upper tertile of total cholesterol/high-density lipoprotein cholesterol ratio, fasting triglycerides, and insulin resistance had 44, 32, and 24% higher daily salivary 17β-estradiol concentrations, respectively (all $P < 0.05$).

Conclusion: The *CYP17* rs2486758 minor allele may predispose to higher 17β-estradiol levels, particularly in premenopausal women with a high clustered metabolic score. Thus, modification of metabolic risk factors may have significant implications for the prevention of breast cancer in women with the minor allele of *CYP17* rs2486758. **(*J Clin Endocrinol Metab* 97: E852–E857, 2012)**

ISSN Print 0021-972X ISSN Online 1945-7197
Printed in U.S.A.

doi: 10.1210/jc.2011-2577 Received September 16, 2011. Accepted February 10, 2012.
First Published Online March 14, 2012

Abbreviations: HDL-C, High-density lipoprotein cholesterol; HOMA, homeostatic model assessment; SNP, single nucleotide polymorphism.

Exposure to ovarian hormones is critical to breast cancer development; the genes that control estrogen and progesterone biosynthesis in the ovaries may contribute to inherent variability in breast cancer susceptibility. The present study focuses on *CYP17*, which is expressed in ovarian theca cells and encodes cytochrome P450 enzymes that control the early steps of endogenous estrogen biosynthesis by converting progesterone into precursors of estrogen.

So far, reports have failed to demonstrate a consistent relationship between *CYP17* genotype and levels of ovarian hormones in premenopausal women (1–4) or breast cancer risk (5, 6). However, mainly rs743572, which is a single nucleotide polymorphism (SNP) located in the promoter region of *CYP17*, has been studied, whereas recent evidence indicates that SNPs in noncoding regions of a gene may effect gene expression. Interestingly, when stratifying by body mass index, Small *et al.* (2) documented an association between *CYP17* rs743572 and serum estradiol concentration among leaner women. Importantly, metabolic risk factors (*i.e.* body fatness, hypertension, dyslipidemia, and insulin resistance) increase levels of ovarian hormones (7–10) and breast cancer risk (5, 6), but little is known about interacting effects of genetic predisposition caused by other common SNPs in *CYP17*.

Based on these previous studies, there is a need for further investigations to address possible modifying effects of metabolic risk factors on the association between common SNPs in *CYP17* and ovarian hormones. Therefore, the main aim of this study was to examine whether eight SNPs in *CYP17*, in combination with higher levels of metabolic risk factors (individually or clustered), are associated with increased daily salivary 17β-estradiol and progesterone concentrations in healthy premenopausal women with regular ovulatory cycles.

Subjects and Methods

Participants and study design

The Norwegian Energy Balance and Breast Cancer Aspects (EBBA-I) Study was conducted in 2000–2002 and included 204 women aged 25–35 yr with regular menstrual cycles (length, 20–40 d) who did not use any daily medication; were not pregnant, lactating, or using steroid contraceptives in the 6 months before recruitment; and had no gynecological or chronic disorders (7). Characteristics including reproductive history and lifestyle factors were collected using questionnaires at the time of recruitment. One woman was excluded due to low call frequency in the genotyping analysis; thus, 203 participants are included in this report. All participants gave written informed consent, and ethical approval was obtained for the study.

Clinical examination and collection of blood samples

All participants underwent clinical examination at the Clinical Research Center, University Hospital of North Norway, Tromsø, Norway, at three scheduled visits during their menstrual cycle. The present report considered measures from the first visit only. Morning blood samples were taken at the first scheduled visit (between d 1 and 5 of the menstrual cycle), after a fasting period that started at 2400 h and included abstaining from smoking and exercise. Details of the EBBA-I Study, anthropometric measurements, and serum lipids, glucose and SHBG assays have been described (7) and are included in the Supplemental Materials and Methods (published on The Endocrine Society's Journals Online web site at http://jcem.endojournals.org).

Collection and analysis of saliva samples

Participants self-collected daily morning saliva samples at home for one entire menstrual cycle, starting on the first day of bleeding (7, 11).

Salivary 17β-estradiol concentration was assayed for 20 d and progesterone for 14 d of the menstrual cycle, using I-125-based RIA kits (see Supplemental Materials and Methods for details). All cycles were aligned at midcycle based on the identification of the drop in 17β-estradiol concentration (aligned cycle day 0), which provides a reasonable estimate of the day of ovulation (11). Satisfactory identification of the drop in salivary 17β-estradiol concentration could not be made for 14 women. For the remaining 189 women, a "late-follicular" index (mean value aligned cycle days -5 to -1), a "luteal" index (mean value aligned cycle days 0 to $+6$), and a "mid-menstrual" index (mean value aligned cycle days -7 to $+6$) were calculated. Maximum peak level refers to the highest measured hormone value during the menstrual cycle (aligned cycle day -1).

SNP selection and genotyping

To represent the variability in *CYP17* in Caucasians, eight of its SNP (rs1004467, rs743575, rs4919687, rs3781286, rs3824755, rs10786712, rs743572, and rs2486758) were selected using the Genome Variation Server (LDSelect). Criteria for SNP selection were an r^2 threshold of 0.8 and a minor allele frequency greater than 5%. Genotyping was performed using the Illumina Golden Gate Platform (Illumina Inc., San Diego, CA).

Homo- and heterozygote genotypes of the minor allele were combined into one category, using the homozygote genotype of the major allele as the reference category (dominant model). Five SNPs in *CYP17* were not associated with ovarian hormones, and interactions with metabolic risk factors were not observed in crude analyses. Another two (rs743575, rs4919687) were weakly associated with overall mean salivary 17β-estradiol concentration in crude analyses, but no interactions with metabolic risk factors were observed. Therefore, these seven SNPs were considered less suitable for the main aim of this report and were discarded. One SNP in *CYP17* (rs2486758) was associated with overall mean salivary 17β-estradiol concentrations in crude analyses and was therefore used in further analyses. Further information about SNP selection, as well as details about genotyping and quality control including allele frequencies and a linkage disequilibrium plot, is reported in the Supplemental Materials and Methods, Supplemental Table 1, and Supplemental Fig. 1.

Statistical analyses

Statistical analyses were performed using STATA version SE 11.0 (StataCorp., College Station, TX). The statistical significance level was defined as $P < 0.05$. More details are available in the Supplemental Materials and Methods and Supplemental Table 3.

Descriptive analyses

Student's *t* test and Pearson's χ^2 test were used to compare means and proportions of selected characteristics of the participants by genotype. The homeostatic model assessment (HOMA) score [fasting glucose (mmol/liter) × fasting insulin (μIU/ml)]/22.5 was used as an indicator of insulin resistance. We constructed a clustered metabolic score by summarizing z-scores [(individual raw values − sample mean)/sample SD] of the following variables for each woman: waist circumference, fasting triglycerides, total cholesterol/high-density lipoprotein cholesterol (HDL-C) ratio, HOMA score, and mean arterial pressure [(diastolic blood pressure × 2) + systolic blood pressure]/3. Metabolic score was considered a continuous variable with a mean score of 0, so that lower values corresponded to a more favorable profile.

Analyses of single SNP

Linear regression and generalized estimating equation models were used to evaluate associations between *CYP17* SNP and salivary ovarian hormone concentrations. All hormone values were log-transformed before the statistical analyses. Possible interactions between *CYP17* SNP were evaluated by multiplicative terms in the models with age as the only covariate. Wald χ^2 test statistics was used to assess the associations between *CYP17* rs2486758 genotypes, metabolic risk factors, and daily salivary 17β-estradiol concentrations. The study population was homogeneously Caucasian; thus, ethnicity was not included in the model as a covariate. Participants were stratified by tertiles of metabolic risk factors (individual and clustered).

TABLE 1. Characteristics of the Norwegian EBBA-I Study population by *CYP17* rs2486758 genotypes (n = 203)[a]

	Major allele[b]	Minor allele[b]	*P* value[c]
n	121	82	
Age (yr)	30.7 (30.1, 31.2)	30.8 (30.1, 31.4)	0.85
Age at menarche (yr)	13.0 (12.8, 13.3)	13.3 (12.9, 13.6)	0.25
Cycle length (d)	28.5 (28.0, 29.1)	27.8 (27.0, 28.5)	0.087
Energy intake (kJ/d)	8192 (7853, 8531)	7915 (7494, 8336)	0.31
Alcohol (units/wk)	3.26 (3.03, 3.49)	3.14 (2.85, 3.42)	0.50
Current smokers (%)	23.1	20.7	0.69
Sedentary activity in leisure time (%)	59.4	40.6	0.98
Body composition			
BMI (kg/m^2)	24.2 (23.6, 24.9)	24.5 (23.7, 25.4)	0.59
Waist circumference (cm)	79.0 (77.3, 80.7)	79.9 (77.8, 82.0)	0.54
Clinical measurements			
Systolic blood pressure (mm Hg)	112.5 (110.4, 114.6)	114.4 (112.1, 116.7)	0.23
Diastolic blood pressure (mm Hg)	70.5 (68.9, 72.0)	71.5 (70.0, 73.1)	0.37
Serum concentrations[d]			
Total cholesterol (mmol/liter)	4.48 (4.34, 4.63)	4.39 (4.23, 4.55)	0.41
HDL-C (mmol/liter)	1.55 (1.49, 1.61)	1.53 (1.45, 1.60)	0.64
Total cholesterol/HDL-C ratio	3.00 (2.86, 3.15)	3.00 (2.82, 3.18)	0.97
Triglycerides (mmol/liter)	0.78 (0.70, 0.86)	0.99 (0.64, 1.34)	0.17
Glucose (mmol/liter)	4.97 (4.87, 5.07)	5.09 (4.96, 5.22)	0.15
Insulin (pmol/liter)	85.2 (72.9, 97.4)	84.1 (75.9, 92.3)	0.90
SHBG (nmol/liter)[e]	52.6 (49.1, 56.2)	51.0 (46.7, 55.3)	0.57
Clustered metabolic score[f]	2.07 (1.44, 2.70)	2.91 (1.97, 3.85)	0.13
Saliva concentrations (pmol/liter)[e]			
Overall 17β-estradiol[g]	13.0 (11.8, 14.3)	15.4 (13.7, 17.2)	0.026
Late follicular index 17β-estradiol[h]	16.7 (15.0, 18.6)	18.4 (16.2, 21.0)	0.11
Luteal index 17β-estradiol[i]	13.9 (12.3, 15.6)	16.4 (14.3, 18.8)	0.035
Overall progesterone[g]	90.0 (78.8, 102.8)	101.4 (87.3, 117.7)	0.24

Data are expressed as mean (95% confidence interval) or percentage unless otherwise specified. BMI, Body mass index.

[a] Number may vary due to missing information.

[b] Major allele, homozygote; minor allele, hetero- and homozygote.

[c] Student's *t* test or Pearson's χ^2 test, linear regression, or generalized estimating equation with log-transformed hormones as dependent variable.

[d] Fasting serum measurements at d 1–5 after onset of menstrual cycle.

[e] Data represent age-adjusted geometric means (95% confidence interval).

[f] Estimated from waist circumference, fasting triglycerides, total cholesterol/HDL-C ratio, mean arterial pressure [(diastolic blood pressure × 2) + systolic blood pressure]/3, and HOMA score [fasting glucose (mmol/liter) × fasting insulin (μIU/ml)]/22.5.

[g] Genotypes equally distributed among women with and without identified drop day.

[h] Aligned cycle day −5, −1 (n = 189).

[i] Aligned cycle day 0, +6 (n = 189).

Haplotype analyses of multiple SNP

Haplotype analyses were performed by using SNPStats software (Catalan Institute of Oncology, IDIBELL, Epidemiology and Cancer Registry L'Hospitalet, Barcelona, Spain) (12). The eight selected SNP formed nine common haplotypes (frequency $> 0.5\%$) in our study population. Likelihood ratio test was used to assess the association between the haplotypes and salivary 17β-estradiol concentration.

Results

Analyses of a single SNP

CYP17 rs2486758 genotype and ovarian hormones

Participants with the *CYP17* rs2486758 minor allele had an 18.5% higher overall salivary 17β-estradiol concentration ($P = 0.026$), and an 18.0% higher luteal index ($P = 0.035$), compared with women with the homozygote genotype of the major allele (Table 1). Analyses of salivary 17β-estradiol concentration by aligned cycle day are reported in Supplemental Table 2 and Supplemental Fig. 2. *CYP17* rs2486758 genotype was not associated with salivary progesterone and serum SHBG concentrations (Table 1).

Comparison across tertiles of metabolic risk factors: CYP17 rs2486758 genotype and 17β-estradiol

For participants in the upper tertile of metabolic risk factors with the minor allele, we observed an increase in daily salivary 17β-estradiol concentration throughout the menstrual cycle: 24% for HOMA score ($P = 0.021$), 32% for fasting triglycerides ($P = 0.004$), 44% for total cholesterol/HDL-C ratio ($P = 0.005$), and 53% for clustered metabolic score ($P < 0.001$), compared with all other women combined (Fig. 1). *CYP17* rs2486758 was not associated with serum SHBG concentration among women in the upper tertile of the clustered metabolic score ($P = 0.19$).

Comparison within the upper tertile of metabolic risk factors: CYP17 rs2486758 genotype and 17β-estradiol

For women in the upper tertile of metabolic risk factors, the minor allele was associated with an increase in daily salivary 17β-estradiol concentration of 29% for fasting triglycerides ($P = 0.040$), 35% for total cholesterol/HDL-C ratio ($P = 0.039$), and 38% for clustered metabolic score ($P = 0.004$) compared with women in the upper tertile of metabolic risk factors carrying the major allele (Fig. 1). The increase in peak salivary 17β-estradiol concentration was 36% ($P = 0.024$) for clustered metabolic score (Fig. 1F).

Haplotype analyses

One specific haplotype was associated with increased salivary 17β-estradiol concentration ($P = 0.0045$). There was no significant difference in salivary 17β-estradiol concentration between haplotypes (global haplotype association P value = 0.65) (Table 2).

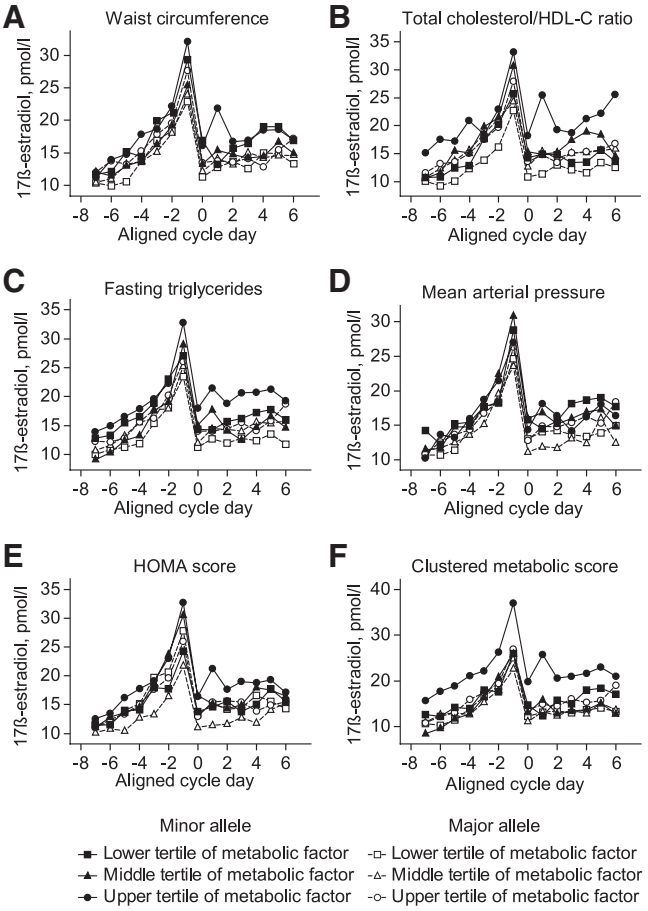


FIG. 1. Daily salivary 17β-estradiol concentrations (geometric means) in midmenstrual cycle (aligned cycle day, −7, 6) for women categorized by *CYP17* rs2486758 genotype and tertiles of metabolic risk factors. Major allele, homozygote; minor allele, hetero- and homozygote. A, Waist circumference. Major allele and lower tertile (n = 32), middle tertile (n = 41), upper tertile (n = 38); minor allele and lower tertile (n = 23), middle tertile (n = 27), upper tertile (n = 28). B, Total cholesterol/HDL-C ratio. Major allele and lower tertile (n = 36), middle tertile (n = 34), upper tertile (n = 38); minor allele and lower tertile (n = 27), middle tertile (n = 27), upper tertile (n = 22). C, Fasting triglycerides. Major allele and lower tertile (n = 37), middle tertile (n = 38), upper tertile (n = 34); minor allele and lower tertile (n = 27), middle tertile (n = 24), upper tertile (n = 25). D, Mean arterial pressure calculated by the equation: [(diastolic blood pressure × 2) + systolic blood pressure]/3. Major allele and lower tertile (n = 41), middle tertile (n = 34), upper tertile (n = 36); minor allele and lower tertile (n = 21), middle tertile (n = 28), upper tertile (n = 29). E, HOMA score calculated by the equation: [fasting glucose (mmol/liter) × fasting insulin (μIU/ml)]/22.5. Major allele and lower tertile (n = 39), middle tertile (n = 41), upper tertile (n = 29); minor allele and lower tertile (n = 22), middle tertile (n = 21), upper tertile (n = 33). F, Clustered metabolic score estimated from: waist circumference, fasting triglycerides, total cholesterol/HDL-C ratio, blood pressure as per the equation above, and HOMA score as per the equation above. Major allele and lower tertile (n = 43), middle tertile (n = 31), upper tertile (n = 34); minor allele and lower tertile (n = 22), middle tertile (n = 27), upper tertile (n = 27).

Discussion

To our knowledge, this is the first study to evaluate a set of SNPs, including *CYP17* rs2486758, in relation to ovarian

TABLE 2. Haplotype association with salivary 17β-estradiol concentrations for the eight selected SNP in *CYP17* in the Norwegian EBBA-I Study (n = 203)

		17β-estradiol		
Haplotype[a]	Frequency	Difference	95% CI	*P* value
AAAGCGAA	0.28	0.00	Ref.	
ACGGCGAA	0.19	−0.06	−0.23, 0.12	0.53
AAAACAGA	0.16	0.03	−0.14, 0.21	0.73
AAAGCGAG	0.16	0.25	0.08, 0.43	0.0045
AAAACAGG	0.06	−0.17	−0.47, 0.13	0.26
GAAAGAGA	0.05	0.06	−0.18, 0.31	0.62
ACGACAGA	0.05	−0.04	−0.3, 0.23	0.79
GCGAGAGA	0.03	0.19	−0.15, 0.53	0.28
GAAAGAGG	0.02	0.1	−0.33, 0.53	0.65

Test of equality between haplotypes (global haplotype association), *P* value = 0.65.

[a] Haplotype order, rs1004467 (A>G), rs743575 (A>C), rs4919687 (G>A), rs3781286 (G>A), rs3824755 (C>G), rs10786712 (G>A), rs743572 (A>G), and rs2486758(A>G).

hormone concentrations and metabolic risk factors. We observed a novel association between the *CYP17* rs2486758 minor allele and elevated daily levels of free biologically active estradiol among premenopausal Norwegian women in the upper tertile of metabolic risk factors.

Some previous studies have reported increased levels of salivary and serum estradiol among healthy premenopausal women with the hetero- and homozygote genotype of the *CYP17* rs743572 minor allele (1–3), whereas others reported inconsistent results (4), as has also been the case with serum progesterone (1, 4).

CYP17 rs2486758 is localized in the intergenic section near the 5′ of *CYP17*, and approximately 40% of trait-associated SNPs have been found in intergenic regions (13). Based on current understanding in this field, we can predict that the *CYP17* rs2486758 minor allele may increase *CYP17* expression by effecting gene splicing, transcription factor binding, or the sequence of noncoding RNA (13). Changes in the feedback sensitivity and adjustment of the estradiol set-point driven by *CYP17* could be a possible explanation for the observed increase in circulating levels of estradiol, and this will have to be tested in further studies.

In the present analysis, *CYP17* rs2486758 was not associated with metabolic risk factors and levels of SHBG. Thus, our data support a true gene-environment interaction, in which only women with the *CYP17* rs2486758 minor allele are susceptible to the possible estrogen-enhancing effects of a high clustered metabolic score. A similar interaction was observed between *CYP17* rs743572 and body mass index in a study of serum estradiol (2). Nutrition and physiological stress have also been reported to change gene expression through epigenetic mechanisms, and there is indirect evidence of epigenetic regulation of *CYP17* from *in vitro* studies (14). *CYP17* expression in the ovaries depends on complex interactions between multiple molecular pathways (15). Moreover, differential tissue-specific regulation of *CYP17* in the ovary and adipose tissue (16) may also have contributed to the interactions observed in our study.

Physiological studies have shown that nutritional status (*i.e.* overweight) and energy metabolism may regulate estrogen levels through different mechanisms; increased energy availability up-regulates ovarian function and hormone secretion in premenopausal years (17). Furthermore, estrogens are formed from androgens in adipose tissue, and adiposity-associated insulin resistance causes reduction in the hepatic production of SHBG (18).

The observed interaction between *CYP17* and metabolic risk factors in the present study may be particularly relevant for breast cancer. Indeed, because the disease seems to cluster in a subset of the female population (19), genetic variation and interactions with lifestyle-related factors may contribute to breast cancer risk, but further studies are needed.

In a larger study, no association between *CYP17* rs2486758 and breast cancer risk was observed (20). On the contrary, an increased breast cancer risk has been linked to *CYP17* rs743572 minor allele carriers (21). Nevertheless, several studies have reported conflicting results and questioned breast cancer susceptibility related to *CYP17* genotypes (6).

Our study suggests that genetically susceptible women with the *CYP17* rs2486758 minor allele may benefit from the modification of metabolic risk factors for the prevention of breast cancer and other estrogen-related diseases (*i.e.* osteoporosis, cardiovascular disease). These are important public health perspectives considering the worldwide increase in the prevalence of obesity, type 2 diabetes, and metabolic syndrome. However, our findings need to be replicated by others and in populations of different ethnicities. Also, studies of breast cancer and other disease outcomes are needed to clarify whether this new knowledge should be incorporated in health care delivery.

Acknowledgments

Address all correspondence and requests for reprints to: Anita Iversen, Department of Community Medicine, Faculty of Health Sciences, University of Tromsø, 9037 Tromsø, Norway. E-mail: anita.iversen@uit.no.

Funding for this study was provided by Norwegian Foundation for Health and Rehabilitation Grants 59010-2000/2001/2002, Norwegian Cancer Society Grants 05087 and TP 49 258, Aakre Foundation Grants 5695-2000 and 5754-2002, Northern Norway Regional Health Authority Grant SFP-563-06, and a University of Tromsø Ph.D. grant.

Disclosure Summary: There were no conflicts of interest.

References

1. **Feigelson HS, Shames LS, Pike MC, Coetzee GA, Stanczyk FZ, Henderson BE** 1998 Cytochrome P450c17α gene (CYP17) polymorphism is associated with serum estrogen and progesterone concentrations. Cancer Res 58:585–587
2. **Small CM, Marcus M, Sherman SL, Sullivan AK, Manatunga AK, Feigelson HS** 2005 CYP17 genotype predicts serum hormone levels among pre-menopausal women. Hum Reprod 20:2162–2167
3. **Jasienska G, Kapiszewska M, Ellison PT, Kalemba-Drozdz M, Nenko I, Thune I, Ziomkiewicz A** 2006 CYP17 genotypes differ in salivary 17β-estradiol levels: a study based on hormonal profiles from entire menstrual cycles. Cancer Epidemiol Biomarkers Prev 15:2131–2135
4. **García-Closas M, Herbstman J, Schiffman M, Glass A, Dorgan JF** 2002 Relationship between serum hormone concentrations, reproductive history, alcohol consumption and genetic polymorphisms in pre-menopausal women. Int J Cancer 102:172–178
5. **Chen Y, Pei J** 2010 Factors influencing the association between CYP17 T34C polymorphism and the risk of breast cancer: meta-regression and subgroup analysis. Breast Cancer Res Treat 122:471–481
6. **Mitrunen K, Jourenkova N, Kataja V, Eskelinen M, Kosma VM, Benhamou S, Vainio H, Uusitupa M, Hirvonen A** 2000 Steroid metabolism gene CYP17 polymorphism and the development of breast cancer. Cancer Epidemiol Biomarkers Prev 9:1343–1348
7. **Furberg AS, Jasienska G, Bjurstam N, Torjesen PA, Emaus A, Lipson SF, Ellison PT, Thune I** 2005 Metabolic and hormonal profiles: HDL cholesterol as a plausible biomarker of breast cancer risk. The Norwegian EBBA Study. Cancer Epidemiol Biomarkers Prev 14: 33–40
8. **Iversen A, Thune I, McTiernan A, Emaus A, Finstad SE, Flote V, Wilsgaard T, Lipson SF, Ellison PT, Jasienska G, Furberg AS** 2011 Ovarian hormones and reproductive risk factors for breast cancer in premenopausal women: the Norwegian EBBA-I study. Hum Reprod 26:1519–1529
9. **Finstad SE, Emaus A, Tretli S, Jasienska G, Ellison PT, Furberg AS, Wist EA, Thune I** 2009 Adult height, insulin, and 17β-estradiol in young women. Cancer Epidemiol Biomarkers Prev 18:1477–1483
10. **Emaus A, Veierød MB, Furberg AS, Espetvedt S, Friedenreich C, Ellison PT, Jasienska G, Andersen LB, Thune I** 2008 Physical activity, heart rate, metabolic profile, and estradiol in premenopausal women. Med Sci Sports Exerc 40:1022–1030
11. **Ellison PT, Lipson SF** 1999 Salivary estradiol—a viable alternative? Fertil Steril 72:951–952
12. **Solé X, Guinó E, Valls J, Iniesta R, Moreno V** 2006 SNPStats: a web tool for the analysis of association studies. Bioinformatics 22:1928–1929
13. **Manolio TA** 2010 Genomewide association studies and assessment of the risk of disease. N Engl J Med 363:166–176
14. **Flück CE, Miller WL** 2004 GATA-4 and GATA-6 modulate tissue-specific transcription of the human gene for P450c17 by direct interaction with Sp1. Mol Endocrinol 18:1144–1157
15. **Patel SS, Beshay VE, Escobar JC, Carr BR** 2010 17α-Hydroxylase (CYP17) expression and subsequent androstenedione production in the human ovary. Reprod Sci 17:978–986
16. **Puche C, José M, Cabero A, Meseguer A** 2002 Expression and enzymatic activity of the P450c17 gene in human adipose tissue. Eur J Endocrinol 146:223–229
17. **Ziomkiewicz A, Ellison PT, Lipson SF, Thune I, Jasienska G** 2008 Body fat, energy balance and estradiol levels: a study based on hormonal profiles from complete menstrual cycles. Hum Reprod 23: 2555–2563
18. **Key TJ, Allen NE, Verkasalo PK, Banks E** 2001 Energy balance and cancer: the role of sex hormones. Proc Nutr Soc 60:81–89
19. **Balmain A, Gray J, Ponder B** 2003 The genetics and genomics of cancer. Nat Genet 33:238–244
20. **Setiawan VW, Schumacher FR, Haiman CA, Stram DO, Albanes D, Altshuler D, Berglund G, Buring J, Calle EE, Clavel-Chapelon F, Cox DG, Gaziano JM, Hankinson SE, Hayes RB, Henderson BE, Hirschhorn J, Hoover R, Hunter DJ, Kaaks R, Kolonel LN, Kraft P, Ma J, Le Marchand L, Linseisen J, Lund E, Navarro C, Overvad K, Palli D, Peeters PH, Pike MC, Riboli E, Stampfer MJ, Thun MJ, Travis R, Trichopoulos D, Yeager M, Ziegler RG, Spencer Feigelson H, Chanock SJ** 2007 CYP17 genetic variation and risk of breast and prostate cancer from the National Cancer Institute Breast and Prostate Cancer Cohort Consortium (BPC3). Cancer Epidemiol Biomarkers Prev 16:2237–2246
21. **Chakraborty A, Murthy NS, Chintamani C, Bhatnagar D, Mohil RS, Sharma PC, Saxena S** 2007 CYP17 gene polymorphism and its association with high-risk north Indian breast cancer patients. J Hum Genet 52:159–165

Supplemental data

Materials and methods

Clinical parameters

All clinical measurements were conducted by two trained nurses at the Clinical Research Center, University Hospital of North Norway (UNN), Tromsø, Norway (1). Each participant came to the UNN three times during their menstrual cycle for clinical examination, between days 1–5, days 7–12 and days 21–25. The first visit was on the first day possible after the onset of menstrual bleeding, and was the only visit taken into account in the present report. During this first visit anthropometric measurements were taken with participants wearing light clothing and no footwear. Height was measured to the nearest 0.5 cm, and weight to the nearest 0.1 kg on an electronic scale. Body mass index was measured in kg/m^2. Waist circumference (cm) was measured 2.5 cm above the umbilicus. Blood pressure was measured three times subsequent to a 5-minute resting period in a sitting position, and the mean of the final two measurements was used in the analysis.

Collection and analysis of blood samples

Morning blood samples were taken at each of the three scheduled visits, after a fasting period that started at midnight and included abstaining from smoking and exercise. The present report considered data from the first visit only. Glucose, sex- hormone-binding globulin (SHBG) and lipid levels were measured in fresh serum at the Department of Clinical Chemistry, UNN (1). Serum glucose was measured enzymatically by the hexokinase method. SHBG was measured by an immunometric method (Diagnostic Products Corporation (DPC), Bierman GmbH, Bad Nauheim, Germany). Fasting triglycerides were assayed by enzymatic hydrolysis with lipase. Total cholesterol was determined enzymatically using cholesterol esterase and cholesterol oxidase. High-density lipoprotein cholesterol (HDL-C) was quantified by direct assay using PEG-modified enzymes and dextran sulfate. Insulin was measured at the Hormone Laboratory, Aker

University Hospital, Oslo, Norway, in sera stored at -70°C until analysis by radioimmunoassay (RIA) (Linco Research Inc., St. Charles, Missouri, USA).

Collection and analysis of saliva samples

Participants self-collected daily morning saliva samples at home for one entire menstrual cycle starting on the first day of bleeding (1, 4). Hormone assays were run in the Reproductive Ecology Laboratory, Harvard University, Cambridge, Massachusettes, USA.

Salivary 17β-estradiol concentration was assayed for 20 days (reverse cycle days −5 to −24; with -1 implicating the last day of the menstrual cycle) and progesterone for 14 days (reverse cycle days −1 to −14) of the menstrual cycle using I-125-based RIA kits (Diagnostic Systems Laboratories, Webster, TX, USA). All samples were run in duplicate, and all samples from a single participant were run together in the same assay, with women randomly assigned to assay batches. The sensitivity of the 17β-estradiol assay was 4 pmol/l. Average intra-assay variability was 9%, and inter-assay variability ranged from 23% for low pools to 13% for high pools. For progesterone, the sensitivity of the essay was 13 pmol/l. Average intra-assay variability was 10%, inter-assay variability ranged from 19% for low pools to 12% for high pools.

Prior to statistical analysis of daily ovarian hormone levels, the cycles of the participants were aligned at mid-cycle following published methods (4). Alignment was based on the identification of the mid-cycle drop in salivary 17β-estradiol concentration (aligned cycle day 0), which provides a reasonable estimate of the day of ovulation. Satisfactory identification of the mid-cycle drop in salivary 17β-estradiol concentration could not be made for 14 women. For the remaining 189 women with aligned cycles, a “late-follicular” index (mean value aligned cycle days -5 to -1), a “luteal” index (mean value aligned cycle days 0 to + 6), a “Mid-menstrual” index (mean value aligned cycle days -7 to + 6) were calculated. Maximum peak level refers to the highest measured hormone value during the menstrual cycle (aligned cycle day -1). Genotypes were equally distributed among women with and without an identified drop day.

SNP selection and genotyping

DNA was extracted from whole blood using MagAttract DNA Blood Mini M48 kit (Qiagen, Oslo, Norway) by the Department of Medical Genetics, UNN. To represent the variability of *CYP17* in Caucasians, eight single nucleotide polymorphisms (SNPs) (rs1004467, rs743575, rs4919687, rs3781286, rs3824755, rs10786712, rs743572, rs2486758) were selected using the Genome Variation Server (LD select) (2-3). The SNPs were selected at an r^2 threshold of 0.8 and a minor allele frequency of > 5%. TagSNP coverage extends 2 kb upstream and 1kb downstream of the gene. Genotyping was performed at the Fred Hutchinson Cancer Research Center, Seattle, Washington, USA, using the Illumina Golden Gate platform. Included in the assay were 22 blinded replicates, and genotype control for the Caucasian population (NA07034). One sample failed, with a call frequency < 85% and none of the selected SNPs in *CYP17* were monomorphic or significantly out of the Hardy-Weinberg Equilibrium. Thus, 203 women were included in the present study.

The eight selected SNPs in *CYP17* were organized into four groups, or bins according to r^2, with one highly correlated tagSNP in each bin to ensure the genetic diversity of that bin in the statistical analysis. One SNP from each bin (rs743575, rs3824755, rs743572 and rs2486758) was then used in the crude analysis. Homo- and heterozygote genotypes of the minor allele were combined into one category using the homozygote genotype of the major allele as the reference category (dominant model).

Ethical considerations

All participating women signed an informed consent form and the study was approved by the Regional Committee for Medical Research Ethics and the Norwegian Data Inspectorate.

Statistical analysis

All hormone values were log transformed prior to regression analysis and back-transformed for presentation (geometric means and 95% confidence intervals). Inclusion of height, the only covariate associated with genotype, and body mass index in the models did not change the regression coefficients of genotypes to any meaningful extent.

Based on metabolic scores associated with physical activity (5, 6) and breast cancer risk (7) in Norwegian women, and the definition of metabolic syndrome by the International Diabetes Federation, which includes waist circumference as the primary measure of central obesity (8), we constructed a clustered metabolic score by summarizing z-scores [(individual raw values − sample mean) / sample standard deviation] of waist circumference, fasting triglycerides, total cholesterol/HDL-C ratio, homeostatic model assessement (HOMA) score (9), and mean arterial pressure [(2 diastolic blood pressure + 1 systolic blood pressure) / 3]. HOMA score and fasting triglycerides were log transformed prior to the calculation of z-scores. This metabolic score was considered as a continuous variable with a mean of 0, so that lower values corresponded to a more favorable profile. Exploratory factor analysis was used to describe the correlations among the variables in the clustered metabolic score, and to determine the number of factors describing the covariance structure (10). Kaiser's criterion (eigenvalue > 1) and Cattell's scree plot were used for factor extraction. The exploratory factor analysis indicated a satisfactory loading (> 0.400) for all variables except fasting triglycerides (< 0.400). Fasting triglycerides were nevertheless kept in the clustered metabolic score because of biological plausibility (Table III).

Generalized estimating equation models, adjusted for age, with Wald chi-square test statistics were used to assess the associations between *CYP17* rs2486758 genotypes and daily salivary 17β-estradiol concentrations.

References

1. **Furberg AS, Jasienska G, Bjurstam N, Torjesen PA, Emaus A, Lipson SF, Ellison PT, Thune I** 2005 Metabolic and hormonal profiles: HDL cholesterol as a plausible biomarker of breast cancer risk. The Norwegian EBBA Study. Cancer Epidemiol Biomarkers Prev 14:33-40.

2. **Carlson CS**, **Eberle MA**, **Rieder MJ**, **Yi Q**, **Kruglyak L**, **Nickerson DA** 2004 Selecting a maximally informative set of single-nucleotide polymorphisms for association analyses using linkage disequilibrium. American Journal of Human Genetics 74:106-120.

3. http://gvs.gs.washington.edu/GVS/

4. **Ellison PT and Lipson SF** 1999 Salivary estradiol- a viable alternative? Fertility and Sterility 72:951-952.

5. **Emaus A**, **Veierød MB**, **Furberg AS**, **Espetvedt S**, **Friedenreich C**, **Ellison PT**, **Jasienska G**, **Andersen LB**, **Thune I** 2008 Physical activity, heart rate, metabolic profile, and estradiol in premenopausal women. Med Sci Sports Exerc 40:1022-1030.

6. **Rizzo NS, Ruiz JR, Hurtig-Wennløf A, Ortega FB, Sjøstrøm M** 2007 Relationship of physical activity, fitness, and fatness with clustered metabolic risk in children and adolescents:the european youth heart study. J Pediatr 150:388-394.

7. **Bjørge T, Lukanova A, Jonsson H, Tretli S, Ulmer H, Manjer J, Stocks T, Selmer R, Nagel G, Almquist M, Concin H, Hallmans G, Häggström C, Stattin P, and Engeland A** 2010 Metabolic syndrome and breast cancer in the me-can (metabolic syndrome and cancer) project. Cancer Epidemiol Biomarkers Prev 19:1737-1745.

8. **Alberti KGMM, Zimmet P, Shaw J** 2006 Metabolic syndrome-a new world-wide definition. A consensus statement from the international diabetes federation. Diabetic Medicine 23:469-480.

9. **Wallace T, Matthews DR** 2002 The assessment of insulin resistance in man. Diabetic Medicine 19:527-534.

10. **Tinsley HEA, Tinsley DJ** 1987 Uses of Factor Analysis in Counseling Psychology Research. Journal of Counseling Psychology 4:414-424.

11. **Barrett JC, Fry B, Maller J, Daly MJ** 2004 Haploview: analysis and visualization of LD and haplotype maps. Bioinformatics 21:263-265.

Supplemental Table I. Allele frequencies and distribution of selected SNPs in *CYP17* in the Norwegian EBBA-I Study (n=203)

SNP	Location (bp)	SNP Position[a]	Alleles[b]	MAF	HWE[c]	AA[d]		Aa[d]		aa[d]	
						n	(%)	n	(%)	n	(%)
rs1004467	104484497	Intron_3	A>G	0.097	0.209	169	(83)	31	(15)	4	(2)
rs743575	104584896	Intron_2	A>C	0.27	0.910	109	(53)	81	(40)	14	(7)
rs4919687	104585238	Intron _1	G>A	0.27	0.986	109	(53)	81	(40)	14	(7)
rs3781286	104585709	Intron_1	G>A	0.37	0.236	86	(42)	86	(42)	32	(16)
rs3824755	104585839	Intron_1	C>G	0.097	0.209	169	(83)	31	(15)	4	(2)
rs10786712	104586386	Intron_1	G>A	0.37	0.236	86	(42)	86	(42)	32	(16)
rs743572	104587142	5' UTR	A>G	0.37	0.236	86	(42)	86	(42)	32	(16)
rs2486758	104587470	Intergenic (GVS) 5' near gene	A>G	0.24	0.221	121	(60)	67	(33)	15	(7)

SNP, single nucleotide polymorphism; MAF, major allele frequency; HWE, Hardy-Weinberg Equilibrium; A, major allele; a, minor allele.

[a]According to information on dbnSNP home-page: http://www.ncbi.nlm.nih.gov/projects/SNP/.

[b]Most frequent to least frequent allele.

[c]p-value for HWE (cut off 0.0010).

[d]Distribution in EBBA-I study population.

Supplemental Table II. Daily salivary 17β-estradiol concentrations by cycle day (geometric means, 95% confidence intervals, CI) for women categorized by *CYP17* rs2486758 genotypes: The Norwegian EBBA-I Study (n = 203)

Hormonal index[c]	Major allele[a] (n = 111) Mean (95% CI)	Minor allele[a] (n = 78) Mean (95% CI)	Diff (%)	P value[b]
Mid-menstrual[d]	14.46 (13.98,14.95)	16.49 (15.81,17.19)	14	0.077
Luteal[e]	15.28 (14.57,16.03)	17.28 (16.30,18.32)	19	0.035
Maximum peak level[f]	24.83 (22.52,27.38)	28.83 (25.37,32.77)	16	0.061

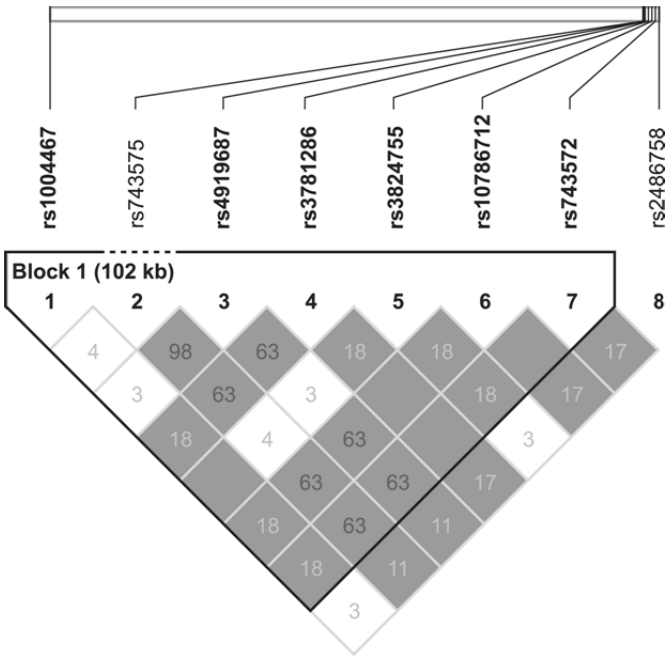
[a]Major allele: homozygote, minor allele: hetero- and homozygote.
[b]Wald chi-square test.
[c]Mean values of salivary 17β-estradiol concentration for a given number of aligned cycle days.
[d]Aligned cycle day -7,+ 6 (n = 189).
[e]Aligned cycle day 0,+ 6 (n = 189).
[f]Aligned cycle day -1 (n = 189).

Supplemental Table III. Relationships among selected risk factors in the clustered metabolic score[a]. The Norwegian EBBA-I Study (n=203)[b]

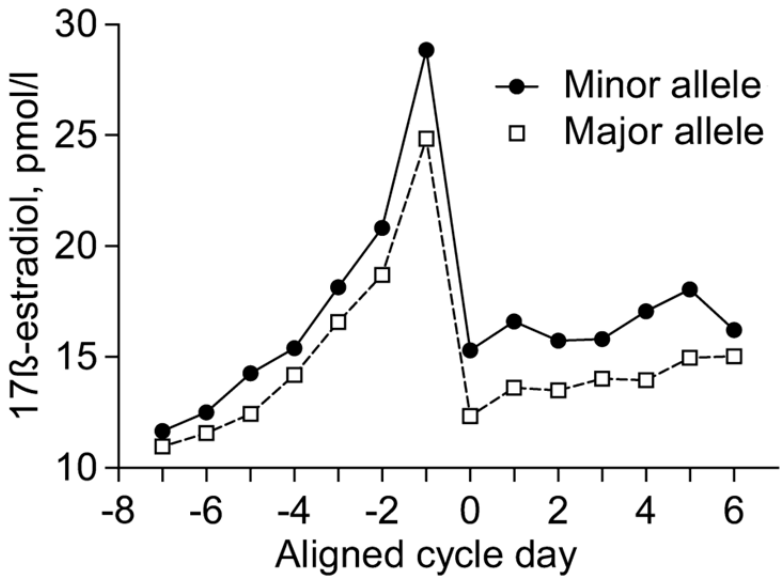
Variable, z-score[c]	Factor loadings[d]	Uniqueness[e]
Waist circumference	0.834	0.304
Fasting triglycerides	0.332	0.890
HOMA score	0.603	0.637
Mean arterial pressure	0.547	0.701
Total cholesterol/HDL-C ratio	0.752	0.435

HDL-C, high-density lipoprotein-cholesterol; HOMA, homeostatic model assessement.

[a]Estimated from sum of z-scores in each women for: waist circumference (cm), total cholesterol/HDL-C ratio, fasting triglycerides mmol/l, mean arterial pressure calculated by the equation: [2 diastolic blood pressure (mmHg) + 1 systolic pressure (mmHg) / 3], and HOMA score given by the equation: [fasting glucose (mmol/l) x fasting insulin (μIU/ml)] / 22.5.
[b]Numbers may vary due to missing information.
[c]z-score calculation; sample mean subtracted from individual score/raw score, divided by the sample standard deviation.
[d]Factor 1 from principal component factor analysis; represents weights for variables and correlation between the variables and the factor.
[e]Proportion of the common variance of the variable not associated with the factor.



Supplemental Figure 1. Linkage disequilibrium structure among the eight selected single nucleotide polymorphisms of *CYP17* (11).



Supplemental Figure 2. Daily salivary 17β-estradiol levels by cycle day (geometric means) in mid-menstrual cycle (aligned cycle day -7, + 6) for women categorized by *CYP17* rs2486758 genotypes; major allele (homozygotes) (n=111) and minor allele (homo- and heterozygotes) (n=78).

Supplemental Table I. Allele frequencies and distribution of selected SNPs in *CYP17*. The Norwegian EBBA Study (n=203)

SNP	Location (bp)	SNP Position[a]	Alleles[b]	MAF	HWE[c]	AA[d]		Aa[d]		aa[d]	
						n	(%)	n	(%)	n	(%)
rs1004467	104484497	Intron_3	A>G	0.097	0.209	169	(83)	31	(15)	4	(2)
rs743575	104584896	Intron_2	A>C	0.27	0.910	109	(53)	81	(40)	14	(7)
rs4919687	104585238	Intron _1	G>A	0.27	0.986	109	(53)	81	(40)	14	(7)
rs3781286	104585709	Intron_1	G>A	0.37	0.236	86	(42)	86	(42)	32	(16)
rs3824755	104585839	Intron_1	C>G	0.097	0.209	169	(83)	31	(15)	4	(2)
rs10786712	104586386	Intron_1	G>A	0.37	0.236	86	(42)	86	(42)	32	(16)
rs743572	104587142	5' UTR	A>G	0.37	0.236	86	(42)	86	(42)	32	(16)
rs2486758	104587470	Intergenic (GVS) 5' near gene	A>G	0.24	0.221	121	(60)	67	(33)	15	(7)

SNP, single nucleotide polymorphism; MAF, major allele frequency; HWE, Hardy-Weinberg Equilibrium; A, major allele, a, minor allele.

[a]According to information on dbnSNP home-page: http://www.ncbi.nlm.nih.gov/projects/SNP/

[b]Most frequent to least frequent allele.

[c]p-value for HWE (cut off 0.0010).[d]Distribution in EBBA study population.

Supplemental Table II. Daily salivary 17β-estradiol concentrations by cycle day (geometric means, 95% confidence intervals, CI) for women categorized by *CYP17* rs2486758 genotypes: The Norwegian EBBA-I Study (n = 203)

Hormonal index[c]	Major allele[a] (n = 111) Mean (95% CI)	Minor allele[a] (n = 78) Mean (95% CI)	Diff (%)	P value[b]
Mid-menstrual[d]	14.46 (13.98,14.95)	16.49 (15.81,17.19)	14	0.077
Luteal[e]	15.28 (14.57,16.03)	17.28 (16.30,18.32)	19	0.035
Maximum peak level[f]	24.83 (22.52,27.38)	28.83 (25.37,32.77)	16	0.061

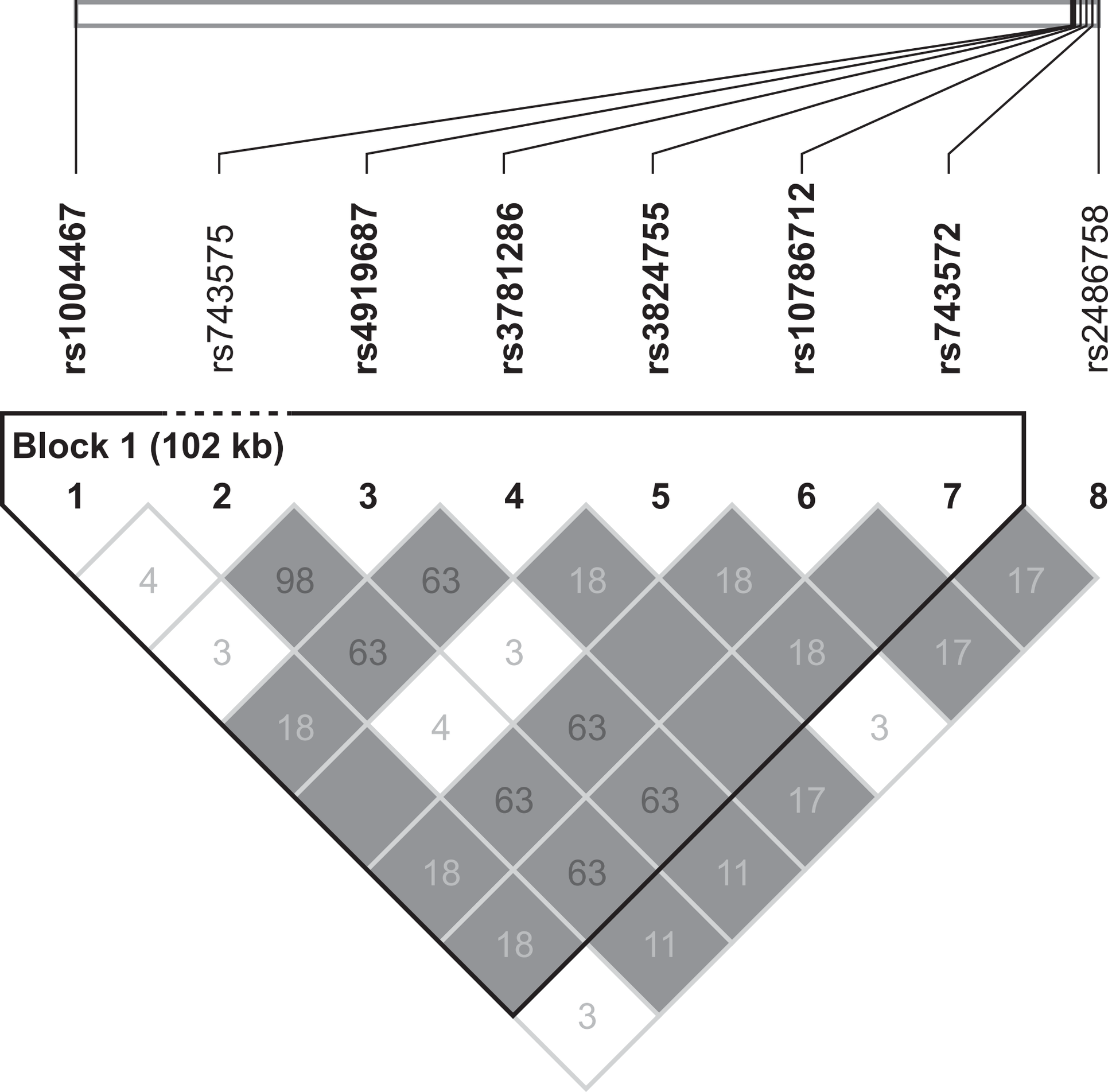
[a]Major allele: homozygote, minor allele: hetero- and homozygote.
[b]Wald chi-square test.
[c]Mean values of salivary 17β-estradiol concentration for a given number of aligned cycle days.
[d]Aligned cycle day -7,+ 6 (n = 189).
[e]Aligned cycle day 0,+ 6 (n = 189).
[f]Aligned cycle day -1 (n = 189).

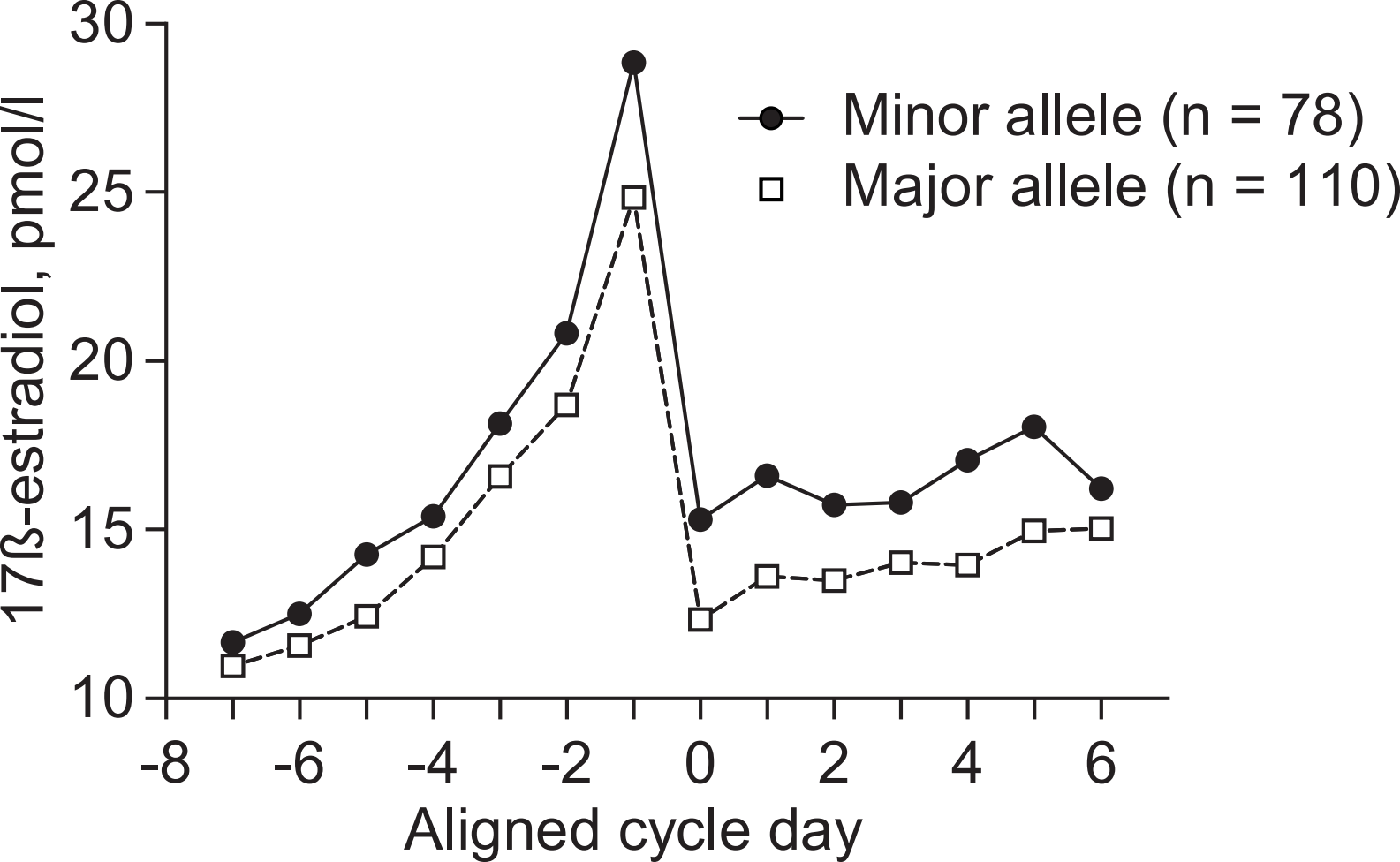
Supplemental Table III. Relationships among selected risk factors in the clustered metabolic score[a]. The Norwegian EBBA-I Study (n=203)[b]

Variable, z-score[c]	Factor loadings[d]	Uniqueness[e]
Waist circumference	0.834	0.304
Fasting triglycerides	0.332	0.890
HOMA score	0.603	0.637
Mean arterial pressure	0.547	0.701
Total cholesterol/HDL-C ratio	0.752	0.435

HDL-C, high-density lipoprotein-cholesterol; HOMA, homeostatic model assessement.

[a]Estimated from sum of z-scores in each women for: waist circumference (cm), total cholesterol/HDL-C ratio, fasting triglycerides mmol/l, mean arterial pressure calculated by the equation: [2 diastolic blood pressure (mmHg) + 1 systolic pressure (mmHg) / 3], and HOMA score given by the equation: [fasting glucose (mmol/l) x fasting insulin (μIU/ml)] / 22.5.
[b]Numbers may vary due to missing information.
[c]z-score calculation; sample mean subtracted from individual score/raw score, divided by the sample standard deviation.
[d]Factor 1 from principal component factor analysis; represents weights for variables and correlation between the variables and the factor.
[e]Proportion of the common variance of the variable not associated with the factor.


rs10044467
rs7435575
rs4919687
rs37811286
rs38247555
rs10786712
rs743572
rs2486758
Block 1 (102 kb)
1
2
3
4
5
6
7
8
4
98
63
18
18
17
3
63
3
18
17
18
4
63
3
63
63
17
18
63
11
18
11
3


30
25
20
15
10
17ß-estradiol, pmol/l
-8
-6
-4
-2
0
2
4
6
Aligned cycle day
Minor allele (n = 78)
Major allele (n = 110)