CYP19A1 genetic variation in relation to prostate cancer risk and circulating sex hormone concentrations in men from the Breast and Prostate Cancer Cohort Consortium

The Harvard community has made this article openly available. Please share how this access benefits you. Your story matters

Citation	Travis, R. C., F. Schumacher, J. N. Hirschhorn, P. Kraft, N. E. Allen, D. Albanes, G. Berglund, et al. 2009. "CYP19A1 Genetic Variation in Relation to Prostate Cancer Risk and Circulating Sex Hormone Concentrations in Men from the Breast and Prostate Cancer Cohort Consortium." Cancer Epidemiology Biomarkers & Prevention 18 (10): 2734–44. https://doi.org/10.1158/1055-9965.epi-09-0496.
Citable link	http://nrs.harvard.edu/urn-3:HUL.InstRepos:41292879
Terms of Use	This article was downloaded from Harvard University's DASH repository, and is made available under the terms and conditions applicable to Other Posted Material, as set forth at http://nrs.harvard.edu/urn-3:HUL.InstRepos:dash.current.terms-of-use#LAA

Published in final edited form as:
Cancer Epidemiol Biomarkers Prev. 2009 October ; 18(10): 2734. doi:10.1158/1055-9965.EPI-09-0496.

CYP19A1 genetic variation in relation to prostate cancer risk and circulating sex hormone concentrations in men from the Breast and Prostate Cancer Cohort Consortium

Ruth C. Travis[1], **Fredrick Schumacher**[2], **Joel N. Hirschhorn**[3], **Peter Kraft**[4], **Naomi E. Allen**[1], **Demetrius Albanes**[5], **Goran Berglund**[6], **Sonja I. Berndt**[5], **Heiner Boeing**[7], **H. Bas Bueno-de-Mesquita**[8], **Eugenia E. Calle**[9], **Stephen Chanock**[10], **Alison M. Dunning**[11], **Richard Hayes**[12], **Heather Spencer Feigelson**[13], **J. Michael Gaziano**[14], **Edward Giovannucci**[4,14,15], **Christopher A. Haiman**[2], **Brian E. Henderson**[2], **Rudolf Kaaks**[16], **Laurence N. Kolonel**[17], **Jing Ma**[14,15], **Laudina Rodriguez**[18], **Elio Riboli**[19], **Meir Stampfer**[4,14,15], **Daniel O. Stram**[2], **Michael J. Thun**[9], **Anne Tjønneland**[20], **Dimitrios Trichopoulos**[4], **Paolo Vineis**[19], **Jarmo Virtamo**[21], **Loïc Le Marchand**[6], and **David J. Hunter**[4,14,15]

[1]Cancer Epidemiology Unit, Nuffield Department of Clinical Medicine, University of Oxford, Richard Doll Building, Oxford, OX3 7LF United Kingdom [2]Department of Preventive Medicine, Keck School of Medicine, University of Southern California, Los Angeles, California, United States of America [3]Broad Institute of Harvard and MIT; Divisions of Genetics and Endocrinology and Program in Genomics, Children's Hospital; Department of Genetics, Harvard Medical School, Massachusetts, United States of America [4]Department of Epidemiology, Harvard School of Public Health, Boston, Massachusetts, United States of America [5]Division of Cancer Epidemiology and Genetics, National Cancer Institute, National Institutes of Health, Bethesda, Maryland [6]Department of Clinical Sciences, Malmö University Hospital, Lund University, Malmö, Sweden [7]German Institute of Human Nutrition, Department of Epidemiology Potsdam-Rehbűcke, Germany [8]National Institute of Public Health and Environment (RIVM), Bilthoven, The Netherlands [9]Department of Epidemiology, American Cancer Society, Williams Street, NW, Atlanta, Georgia 30303-1002, United States of America [10]Core Genotyping Facility, National Cancer Institute, Gaithersburg, Maryland, United States of America [11]Department of Oncology, University of Cambridge, Cambridge, United Kingdom [12]Division of Epidemiology, NYU Langone Medical Center, New York, New York 10016, United States of America [13]Institute for Health Research, Kaiser Permanente Colorado, PO Box 378066, Denver, CO 80237-8066, United States of America [14]Department of Medicine, Brigham and Women's Hospital, VA Boston and Harvard Medical School, Boston, Massachusetts, United States of America [15]Channing Laboratory, Harvard Medical School, Boston, Massachusetts, United States of America [16]Division of Cancer Epidemiology, German Cancer Research Center (DKFZ), Heidelberg, Germany [17]Epidemiology Program, Cancer Research Center, University of Hawaii, Honolulu, Hawaii, United States of America [18]Public Health and Participation Directorate, Health and Health Care Services Council, Asturias, Spain [19]Department of Epidemiology and Public Health, Imperial College, London, UK [20]Institute of Cancer Epidemiology, Danish Cancer Society, Copenhagen, Denmark [21]Department of Chronic Disease Prevention, National Institute for Health and Welfare, Helsinki, Finland

Abstract

Author for reprint requests: Ruth C. Travis, Cancer Epidemiology Unit, Nuffield Department of Clinical Medicine, University of Oxford, Richard Doll Building, Oxford, United Kingdom. ruth.travis@ceu.ox.ac.uk..

Sex hormones, in particular the androgens, are important for the growth of the prostate gland and have been implicated in prostate cancer carcinogenesis, yet the determinants of endogenous steroid hormone levels remain poorly understood. Twin studies suggest a heritable component for circulating concentrations of sex hormones, although epidemiological evidence linking steroid hormone gene variants to prostate cancer is limited. Here we report on findings from a comprehensive study of genetic variation at the *CYP19A1* locus in relation to prostate cancer risk and to circulating steroid hormone concentrations in men by the Breast and Prostate Cancer Cohort Consortium (BPC3), a large collaborative prospective study. The BPC3 systematically characterised variation in *CYP19A1* by targeted resequencing and dense genotyping; selected haplotype-tagging single nucleotide polymorphisms (htSNPs) that efficiently predict common variants in U.S. and European whites, Latinos, Japanese Americans, and Native Hawaiians; and genotyped these htSNPs in 8,166 prostate cancer cases and 9,079 study-, age-, and ethnicity-matched controls. *CYP19A1* htSNPs, two common missense variants and common haplotypes were not significantly associated with risk of prostate cancer. However, several htSNPs in linkage disequilibrium blocks 3 and 4 were significantly associated with a 5–10% difference in estradiol concentrations in men (association per copy of the two-SNP haplotype rs749292–rs727479 (*A*–*A*) versus noncarriers; $P=1 \times 10^{-5}$), and withinverse, although less marked changes, in free testosterone concentrations. These results suggest that although germline variation in *CYP19A1* characterised by the htSNPs produces measurable differences in sex hormone concentrations in men, they do not substantially influence risk for prostate cancer.

Keywords

prostate; cancer; *CYP19A1*; estradiol; testosterone

Introduction

Prostate cancer is the most commonly diagnosed cancer in males in developed countries yet aetiological risk factors for the disease are not well understood. The only established risk factors for the disease are age, family history of prostate cancer and ethnicity. There has been considerable interest in the potential role of sex hormones in prostate cancer carcinogenesis, with a particular focus on androgens, which are important for the development, growth and maintenance of the prostate gland. Finasteride, which blocks the metabolism of testosterone within the prostate, has been found to reduce the risk of prostate cancer (although the increase incidence of high grade tumors on biopsy has led to controversy) (1) and prostate tumours can be induced when testosterone, either alone or together with estradiol, is administered to laboratory animals (2). Estrogens have also been implicated in prostate biology and in the development of prostate cancer, via direct estrogen-receptor mediated effects and indirect effects, although data suggest the role of estrogens may vary with disease progression (3,4). However, a recent re-analysis of the worldwide prospective data found no large associations between circulating androgen and estrogen concentrations in humans and prostate cancer risk (5).

Twin studies suggest a heritable component for circulating concentrations of sex hormones in men (6). The *CYP19A1* gene has been identified as a candidate locus that may influence circulating sex hormone concentrations and risk for hormone-related cancers (7). *CYP19A1* encodes aromatase, an enzyme which catalyses the conversion of the C_{19} androgens, androstenedione and testosterone, to the C_{18} estrogens, estrone and estradiol, respectively. This cytochrome P450 enzyme is expressed primarily in the gonads, as well as in peripheral sites including the prostate (8). In postmenopausal women, common variants in the *CYP19A1* gene have been found to be associated with a 10% to 20% difference in circulating estrogen levels and a number of studies have assessed the relationship of variants in *CYP19A1* with risk for cancers of the breast and endometrium (9–12). However, published data on *CYP19A1* in men

in relation to sex hormones and prostate cancer are relatively limited, partly due to incomplete characterisation of genetic variation at the locus of interest and small sample sizes (13–23).

In the present study, we examined the contribution of common genetic variation at the *CYP19A1* locus to prostate cancer risk and to concentrations of serum sex hormones and sex hormone binding globulin (SHBG) in a large, collaborative investigation (The Breast and Prostate Cancer Cohort Consortium, (BPC3)) (24).

Materials and Methods

Study Population

The BPC3 has been described in detail elsewhere (24). Briefly, the consortium includes large well-established cohorts assembled in the United States and Europe that have DNA for genotyping and extensive questionnaire data from cohort members. For prostate cancer, analyses include men from seven cohort studies: the American Cancer Society Cancer Prevention Study II (ACS CPS-II) (25), the Alpha-Tocopherol, Beta-Carotene Cancer Prevention (ATBC) Study (26), the European Prospective Investigation into Cancer and Nutrition (EPIC) Cohort (itself comprising cohorts from Denmark, Germany, Great Britain, Greece, Italy, the Netherlands, Spain, and Sweden) (27), the Health Professionals Follow-up Study (HPFS) (28), the Hawaii/Los Angeles Multi-ethnic Cohort Study (MEC) (29), the Physicians' Health Study (PHS) (30), and the Prostate, Lung, Colorectal, and Ovarian (PLCO) Cancer Screening Trial (31). With the exception of the MEC and PLCO, these cohorts are composed predominantly of whites of European descent. The MEC contributes African American, Latino, Japanese American and Native Hawaiian cases and controls recruited from Los Angeles and Hawaii. The PLCO includes over 400 African American participants. Cases of prostate cancer were identified through population-based cancer registries or self report confirmed by medical records. The BPC3 data for prostate cancer consist of a series of matched nested case-control studies from each cohort; controls were matched within each cohort to cases on a number of potential confounding factors, including age (within five years), ethnicity, and in some cohorts, additional criteria, such as region of recruitment in EPIC (for further details on the selection of controls in each cohort see Supplementary Methods). The study protocol was approved by all institutional review boards.

SNP discovery and selection of haplotype-tagging single nucleotide polymorphisms (htSNPs)

We used a multi-stage approach to characterize genetic variation by initially resequencing the coding exons and exon-intron boundaries of *CYP19A1* in a multiethnic panel of 95 advanced prostate cancer cases to identify putative functional alleles directly. The linkage disequilibrium (LD) structure of the locus was determined empirically in a multiethnic panel of 349 individuals from the MEC by genotyping a dense network of SNPs, including SNPs selected from dbSNP (www.ncbi.nlm.nih.gov/SNP/) and the Celera SNP database (Celera SNP database (www.Celera.com) and common missense SNPs found during resequencing (although not novel and can also be found in dbSNP) using the Sequenom and Illumina genotyping platforms at the Broad Institute/MIT Center for Genome Research. Haplotype-tagging SNPs (htSNPs) for each haplotype block, determined by the confidence interval method of Gabriel *et al.*(32, 33), were chosen based on R_h^2, a measure of the correlation between observed and predicted haplotypes based on the htSNP genotypes(34), to select a minimum set of SNPs that would achieve an $R_h^2 \geq 0.7$ for all common haplotypes among whites with an estimated frequency of ≥5%.

The LD pattern across *CYP19A1* has been previously shown (9). Briefly, the genomic structure of the region consists of four LD blocks spanning 181kb which was determined by 107 SNPs

with MAF ≥ 0.05 among whites, and 2 missense SNPs with a frequency of >1% among whites in the multiethnic panel. In total 19 htSNPs (Supplementary Table 1) were selected to provide high predictability ($R^2_H \geq 0.7$) (35) of 27 common haplotypes (≥ 0.05 frequency in at least one ethnic group among the five ethnic groups in the multiethnic panel) across the 4 LD block regions (33,36) with inter-block distances < 6 kb. Haplotype frequencies were similar for whites across cohorts (data not shown), while some differences in haplotype frequencies were seen between whites, African Americans, Japanese Americans, and Native Hawaiians.

Genotyping

Genotyping of the htSNPs and two missense SNPs in the 17,245 case patient and control participant samples was conducted using the TaqMan assay (Applied Biosystems) in four BPC3 laboratories. Initial quality control checks of the SNP assays were performed at the manufacturer (ABI); an additional 500 test reactions were run by the BPC3 on the multiethnic reference panel; greater than 99.5% concordance was observed across genotyping platforms. Assay characteristics for the htSNPs for *CYP19A1* are available on the public website http://www.uscnorris.com/mecgenetics/CohortGCKView.aspx. Sequence validation for each SNP assay was performed and 100% concordance observed (http://snp500cancer.nci.nih.gov) (37). To assess inter-laboratory variation, each centre ran assays on a designated set of 94 samples from the SNP 500 cancer panel, showing completion and concordance rates of greater than 99% (37). The internal quality of genotype data at each center was assessed by 5–10% blinded samples in duplicate or triplicate (depending on study) and intra-laboratory concordance rates of greater than 99.5% were observed. Empty water wells were also included on each plate and positioned according to the individual laboratory's protocols. Hardy-Weinberg Equilibrium (HWE) checks have been performed among the controls in each study and stratified by ethnicity in the MEC and by country for EPIC. No deviation in HWE was observed ($p<0.01$) across more than one study for any given assay.

Of the 8,248 prostate cancer cases and 9,312 controls sent for genotyping, at least 1 SNP was successfully genotyped for 8,166 (94%) cases and 9,079 (94%) controls. This study therefore comprises 8,166 case patients and 9,079 control participants. Among these men, we evaluated the relationship of prostate cancer risk to the htSNPs, to the common *CYP19A1* haplotypes predicted among Whites by the 19 htSNPs, and to the two common missense SNPs at the *CYP19A1* locus, including SNPs R264C (rs700519) and T201M (rs28757184). In a subset of control participants from ATBC, EPIC, HPFS, PHS and PLCO, we also investigated whether variation at this locus contributes to inter-individual differences in circulating levels of sex hormones and SHBG, using measurements of steroid hormones made previously by individual cohorts (38–41).

Statistical analyses

We used conditional logistic regression to estimate odds ratios for prostate cancer associated with carrying either 1 or 2 versus 0 copies of the minor allele for each SNP. We estimated haplotype-specific odds ratios using an expectation-substitution approach to account for haplotype uncertainty given unphased genotype data (42,43). Haplotype frequencies and subject-specific expected haplotype indicators were calculated separately for each cohort (and country within EPIC or ethnicity in the MEC and PLCO). To test the global null hypothesis of no association between variation in common *CYP19A1* haplotypes and risk of prostate cancer and to control for type-I error over all the SNPs and haplotypes considered, we used a likelihood ratio test comparing a model with additive effects for each common haplotype (treating the most common haplotype as the referent) to the intercept-only model. For all cancer risk analyses, we test for significance associations at the 0.01 level to minimize the chance of both false positive and false negative results.

We tested for heterogeneity in odds ratio estimates across cohorts among white participants and by ethnicity. In addition, we calculated risk-stratum-specific odds ratios and tested for departures from a multiplicative interaction model to assess whether other risk factors for prostate cancer modify the association with htSNPs, missense SNPs or common haplotypes, including age (at diagnosis), body mass index and family history of prostate cancer. To assess the influence of genetic variation in *CYP19A1* on prostate cancer diagnosed with an aggressive phenotype, we calculated stratum specific odds ratios for high grade prostate cancer (defined as poorly differentiated or Gleason score>=8), advanced stage prostate cancer (Stage C or D), aggressive disease (defined as high grade, advanced stage prostate cancer or death from prostate cancer).

We used fixed effect models to evaluate associations of circulating steroid hormone and sex hormone binding globulin (SHBG) concentrations with *CYP19A1* htSNPs, missense SNPs and common haplotypes in cases and controls, with study cohort included as a fixed effect with a suitable nesting of batches within study where appropriate, and adjustment for age (in five-year age-groups). For these analyses hormone and SHBG concentrations were logarithmically transformed for statistical analyses to approximately normalise their distributions. We report the *P*-value from the test of trend across genotype groups.

Results

Characteristics of prostate cancer case patients and controls

The demographic and other characteristics of cases and controls from the seven cohorts are shown in Table 1. Most study participants were U.S. or European whites (75%), followed by African Americans (11%), Latinos (7%), Japanese Americans (5%), and Native Hawaiians (1%). Among participants, 14% of cases and 9% of controls reported a father or a brother with prostate cancer. Cases and controls were similar with respect to age, BMI, and height. Data on tumour stage and grade were available from six of the seven participating cohorts (all but PLCO). Stage information was available on 70% of genotyped prostate cancer cases, and of these, 19% had advanced disease (defined as stage C or D disease at diagnosis). Gleason score was recorded for 65% of genotyped cases, and 18% of cases had a Gleason sum of 8 or greater.

Associations of CYP19A1 htSNPs, missense SNPs and common haplotypes with prostate cancer risk

Genotype specific odds ratios for the 21 SNPs tested, including the two missense SNPs, are shown in Table 2 for analyses pooling all participants. There was no evidence of an association of the 2 missense SNPs with risk for prostate cancer ($P > 0.50$ for both for analyses). None of the htSNPs showed a significant association with prostate cancer risk at the 0.01 level. The SNP-specific odds ratios did not show significant heterogeneity across ethnicities or across cohorts among whites at the 0.01 level. Global tests of association between *CYP19A1* common haplotypes and prostate cancer were also not significant at the 0.01 level (see Table 3), and no statistically significant associations were observed between individual common haplotypes in each block and risk for prostate cancer ($P > 0.01$ for all).

Tumour stage and grade

We calculated odds ratios for risk of advanced disease by *CYP19A1* htSNPs, missense SNPs and common haplotypes, classifying disease severity by prostate cancer stage, grade or a combined score of prostate tumour stage and histological grade. For aggressive disease in relation to *CYP19A1* htSNPs, rs2445762 in LD block 1 was significantly associated with high grade disease with a significant decrease in risk being observed in heterozygotes ($P_{2\ \text{d.f.test}} = 0.0002$, OR in heterozygotes = 0.76, 99% CI = 0.62–0.93, OR in homozygotes = 1.15, 99% CI = 0.83–1.59), and similarly with the composite variable for aggressive disease ($P = 0.0003$).

We found no significant association between other *CYP19A1* htSNPs or missense SNPs and risk for tumour stage or grade or the composite score of aggressive disease at the 0.01 level. There was no evidence of any association between *CYP19A1* common haplotypes and risk for aggressive prostate cancer, defined as high grade, advanced stage or a composite variable.

Family history, age at diagnosis and body mass index

Tests for departures from multiplicative interaction models were null when we examined statistical interaction between *CYP19A1* htSNPs, missense SNPs or common haplotypes and prostate cancer risk with the following risk factors: family history (at least one first degree relative diagnosed with prostate cancer versus none), age at diagnosis (≤ 65, > 65 years), and BMI (< 25, ≥ 25 < 30, ≥ 30 kg/m^2).

Associations of CYP19A1 htSNPs and missense SNPs with sex hormone concentrations

In analyses of genetic variation in *CYP19A1* in relation to circulating sex hormone and SHBG concentrations, there were significant associations at the 0.01 level for a number of htSNPs with concentrations of estradiol, free estradiol, free testosterone and androstanediol glucuronide. No significant associations were observed with concentrations of testosterone or SHBG, and neither of the missense SNPs were associated with concentrations of sex hormones or SHBG.

For estradiol, the most significant associations were observed for htSNPs in LD blocks 3 and 4: P_{trend} <0.005 for every htSNP in block 3, and P≤0.006 for rs727479 and rs10046 in block 4 (Table 4). These associations did not differ significantly among the three cohorts that measured estradiol, nor did they differ between cases and controls or by ethnicity, although only approximately 5% of the hormone data were from non-white participants. Percentage change in estradiol between homozygotes for the wild-type and the variant allele for these SNPs ranged from approximately 5 to 10%. There is a high degree of LD between previously defined LD blocks 3 and 4, and therefore many of the htSNPs are highly correlated ($r^2 \geq 0.83$ for SNP pairs rs749292 and rs6493494, and, rs2414096 and rs10046) (9). Haplotypes in blocks 3 and 4 were also strongly associated with estrogen levels and the magnitude of the associations was similar to the independent tagging SNPs in these blocks (data not shown).

The two SNPs most significantly associated with estradiol concentrations, rs749292 and rs727479, were only modestly correlated with each other ($r^2 = 0.46$) and each remained significantly associated with circulating levels when modelled concurrently ($P_{trend} = 0.002$; $P_{trend} = 0.006$, respectively). A 2-SNP haplotype (A–A) comprised of these SNPs was found to be a more significant predictor of estradiol concentrations ($P_{trend} = 1 \times 10^{-5}$) (Table 5); however, it accounted for only 0.75% of the variation in estradiol concentrations and when we examined risk for prostate cancer in relation to this two snp-haplotype we observed no significant association with risk for disease at the 0.01 significance level ($P = 0.028$, OR in men with one copy of the A–A haplotype = 0.96, 99% CI = 0.88–1.06, OR for two copies = 1.08, 99% CI = 0.96–1.22).

Full results for all sex hormones and SHBG in relation to *CYP19A1* htSNPs are shown in Supplementary Table 2. Findings for free estradiol were broadly similar to those observed for estradiol, both with respect to level of statistical significance and the percentage change in hormone level by genotype. Significant associations were also observed between *CYP19A1* htSNPs in LD blocks 3 and 4 and the ratio of free estradiol to free testosterone, as an index of aromatase activity, with trends mirroring those observed for estradiol concentrations.

For testosterone, there were no significant associations with htSNPs at the 0.01 significance level, although weak associations that did not reach statistical significance were observed with

the 3 htSNPs in LD block 3 (P_{trend} <0.09), and with rs727479 in LD block 4 (P = 0.04); changes in testosterone concentrations by genotype were the inverse of those seen for estradiol concentrations and percentage changes in hormone concentrations were approximately 2–3% for htSNPs in LD block 3 and 4% for rs727479. For free testosterone concentrations, however, significant or borderline significant associations were observed with a number of htSNPs in LD blocks 3 and 4 (P < 0.01 for rs6493494, rs727479 and rs10046), with trends in free testosterone concentrations being the inverse of those observed for estradiol and free estradiol (Table 4) and percentage changes in free testosterone being approximately 3–4%.

For androstranediol glucuronide, significant associations were observed between androstanediol glucuronide and htSNPs rs28566535/CV1664178 and rs1902586 in LD block 2 ($P_{trend} = 3 \times 10^{-4}$ and 0.001, respectively) and with rs727479 in LD block 4 (P_{trend} = 0.009) (Supplementary Table 2). The percentage changes in androstranediol glucuronide concentration between the wild-type and the variant allele were 18.7%, 13.9% and 6.3%, respectively. We observed no significant associations between *CYP19A1* htSNPs and concentrations of SHBG (Supplementary Table 2).

Discussion

We investigated common genetic variation at the *CYP19A1* locus in relation to prostate cancer risk and hormone levels in a large, collaborative investigation (BPC3) (24). We evaluated both haplotype patterns in four well-characterized LD blocks and two common missense variants and found no evidence that this gene harbours a prostate cancer susceptibility allele. In addition, this is the first large study that has comprehensively assessed genetic variation in *CYP19A1* in relation to sex hormones and SHBG concentrations in men. Our results provide evidence for significant associations of a number of htSNPs in *CYP19A1* with circulating concentrations of estradiol, free estradiol, free testosterone and androstranediol glucuronide.

Ten studies of the relationship between genetic variation in *CYP19A1* and risk for prostate cancer have been published to date (13–18,20–23), with inconsistent findings. These studies only considered a limited number of variants across the locus, and with few exceptions (15, 20,21) were small, being underpowered to detect the modest magnitude of effect anticipated for a common low-penetrance susceptibility allele (RR<1.5. These variants include a tetranucleotide repeat (14,16,21–23), 2 nonsynonymous mutations (13,15,17,18,20), a G/A SNP in the 3`untranslated region (rs10046) (20), and 6 other genetic variants identified in screening of prostate cancer patients (15). The null association with the missense mutation R264C in the current study contrasts with those from three small previous studies (13,17,18), but is in agreement with null results from two large recent studies (15,20), as is our null finding for rs10046 (20). Our results suggesting no strong association with T201M, however, contrast with those from the one previous study, which reported an increase in risk, especially in association with low grade organ confined disease (15). The genome-wide association studies of prostate cancer also provide data on the *CYP19A1* locus but have not implicated the region. For example, in a genome-wide association study in the Cancer Genetic Markers of Susceptibility (CGEMS) project of 1,172 prostate cases and 1,157 controls, 46 SNPs that met quality control metrics were evaluated across CYP19A1 (from 30 kb upstream of the ATG and 30 kb downstream of the polyA tail)(44). Using HapMap Phase 2 CEU samples, the 46 SNPs genotyped tagged (using r2 > 0.8) a total 164 SNPs in 40 bins, which represents 74.9% of all CEU HapMap SNPs with MAF > 5%. In CGEMS, no SNPs had a p value < 0.01 in the adjusted trend test; rs2124873 displayed the smallest p value (p = > 0.03) in CGEMS.

Findings from the current study for prostate cancer risk in relation to genetic variation in *CYP19A1* provide no evidence that risks differ by ethnicity, family history, age or BMI. Our findings also did not support a strong association with severity of disease, with one exception.

We observed a statistically significant association between rs2445762 and risk for high grade disease and a composite score of aggressive disease, with a decreased risk being observed in heterozygotes, but not in homozygotes, for the variant allele. Given the multiple-testing in this study and the small number of cases with two copies of the rs2445762 variant allele (91 and 191 cases with high grade and aggressive disease, respectively), these associations with aggressive prostate cancer are likely due to chance; replication is warranted in further large prostate cancer studies which have detailed information on prostate cancer phenotype.

Despite the lack of evidence for an association between genetic variants and prostate cancer risk, we found evidence for an association between *CYP19A1* variants and circulating sex hormone concentrations. The strongest associations between *CYP19A1* variation and hormone concentrations were observed for htSNPs in LD blocks 3 and 4. These findings are consistent with those from a recent report on 5531 men from Sweden and the US that found SNP rs2470152 in intron 1 of *CYP19A1* was associated with serum estradiol (8% to 13% difference between AA and GG homozygotes in the three cohorts studied; $P = 2 \times 10^{-14}$ for all cohorts combined) (45). For the two most strongly associated SNPs in the current study, the LD relationship with rs2470152 was strong for rs749292 (D' = 0.96, $r^2 = 0.65$) and weaker LD for rs727479 (D' = 0.60, $r^2 = 0.19$), and given these allelic associations, the findings with estradiol for these SNPs (rs749292, 7% difference between homozygotes, $P = 2 \times 10^{-4}$; rs727479, 9% difference, $P = 1 \times 10^{-5}$) are comparable with the Swedish data. Other published studies on sex hormones and *CYP19A1* variation in men have had small sample sizes and results have been conflicting (19,46–48). Data from the current study for estradiol and free estradiol in relation to *CYP19A1* variants are also broadly consistent with those previously reported for women (9–11,49,50), including findings from 3,400 postmenopausal women participating in BPC3 (9). For five out of seven variants in *CYP19A1* LD blocks 3 and 4 found to be significantly associated with estradiol concentrations in women in BPC3, we observed significant associations in men. However, the strength and the magnitude of these associations among men in the current study is somewhat weaker than that observed among postmenopausal women; in men we found variants to be significantly associated with a 5% to 10% difference in estradiol concentrations, whereas in postmenopausal women in BPC3 these variants were significantly associated with a 10% to 20% difference in endogenous estrogen levels (9). It is possible that while *CYP19A1* variants influence aromatase activity, the impact on circulating hormone concentrations in men might be relatively small because of homeostasis of free testosterone by the hypothalamic-pituitary-gonadal feedback loop.

We found no evidence for a significant association of total testosterone with *CYP19A1* variation; however, we found several variants in LD blocks 3 and 4 to be significantly associated with a 2–4% difference in levels of free testosterone, with the trend in hormone concentration being the reverse of that seen for estradiol. An explanation for these findings may lie in the fact that free testosterone is the substrate for conversion to estradiol by aromatase. The association found between androstanediol glucuronide and genetic variation in LD blocks 2 and 4 of *CYP19A1* is also compatible with an influence of *CYP19A1* genetic variants on aromatase activity as androstanediol glucuronide is an end metabolite of testosterone (51).

Overall, our findings that variants in *CYP19A1* were associated with small differences in circulating hormone concentrations but no detectable effect on risk for prostate cancer are consistent with what is known on the association of hormones with risk. A recent collaborative re-analysis of the world-wide data reported no strong association between serum concentrations of sex hormones and risk of prostate cancer (5), with differences in serum sex hormone concentrations between the highest and lowest fourths of the distribution being in the order of two-fold (38). Thus, the small differences (5–10%) in sex hormones in relation to *CYP19A1* variants may be of insufficient magnitude to have a detectable influence on risk for prostate cancer. The relationship between circulating hormone concentrations and intraprostatic

hormone levels, however, remains unclear, and the authors of the collaborative reanalysis concluded that any biological interpretation of their results for serum hormones must be viewed with caution (5). Aromatase is expressed in the prostate and local estrogen production may lead to intraprostatic levels exceeding those in the circulation, with prostate cells being exposed to estradiol from the blood plus locally produced estradiol (3,52,53). However, the null association in the current study between CYP19A1 variants and prostate cancer risk does not lend support to the hypothesis that intraprostatic sex hormone metabolism by aromatase is strongly associated with risk since it would be anticipated that *CYP19A1* variation might have a similar effect on intraprostatic sex hormone concentrations as on circulating hormones.

Strengths of this study lie in the scale of the analysis, both with respect to the comprehensive screening for prostate cancer susceptibility alleles across the entire gene region facilitated by the consortium's cost-effective haplotype-tagging approach, and with respect to the large sample size, which makes it possible to conduct adequately powered subgroup analyses by potential prostate cancer risk factors and by tumour characteristics. In this study, with over 8,000 case patients and 9,000 control participants, we have greater than 95% power to detect a dominant effect or log-additive odds ratio of 1.3 for an allele with 5% minor allele frequency at the 0.001 level. Furthermore, with respect to subgroup analyses we still have, for example, greater than 95% power to detect a stratum-specific dominant odds ratio of 1.7 for a 5% frequency variant at the 0.001 level when the stratum consists of only 20% of the sample. The current study, however, had less power to assess whether variation in *CYP19A1* is associated with risk for prostate cancer in non-white ethnic groups. Further analyses may clarify associations in non-white populations as sample sizes are increased with longer follow-up or the addition of new cohorts.

We have studied the importance of variation of *CYP19A1* to steroid hormone concentrations and to prostate cancer risk. Ultimately, however, the importance of variation in *CYP19A1* on endogenous hormone concentrations should be assessed in the context of variation in the many other genes involved in steroid hormone biosynthesis and metabolism pathways, as well as in genes involved in steroid hormone-binding and -receptor pathways. Whilst variation in individual genes that encode enzymes in the hormone metabolism pathways may individually only result in modest differences in hormone concentrations, a combination of several genes with functional variants could result in a cumulative large effect which would endure over the course of a human lifespan (54). Indeed, twin studies suggest that additive genetic factors may account for up to 57% and 25% of the total variation in circulating testosterone and estradiol concentrations, respectively, in men (6). Thus, the development of a multi-factorial score for the prediction of hormone concentrations within large-scale studies, such as is planned within BPC3, may provide important insights into the determinants of endogenous hormone concentrations and risk for the diseases and disorders that they influence. Such conditions are not limited to cancer and may also include, for example, vascular disease (55), diabetes (56) and those related to cognitive decline (57,58), and to bone metabolism (19,45,46).

In summary, results from this study suggest that although germline mutations in *CYP19A1* produce measurable differences in steroid hormone concentrations in men, they do not substantially influence risk for prostate cancer. These findings are consistent with a reanalysis of worldwide data which found no large associations between serum concentrations of sex hormones and prostate cancer risk. Our results for steroid hormone concentrations in relation to variants in *CYP19A1*, however, may have wider relevance for other conditions of public health importance in men.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

The authors gratefully acknowledge the participants in the component cohort studies.

Grant Support: National Cancer Institute cooperative agreements UO1-CA98233, UO1-CA98710, UO1-CA98216, and UO1-CA98758 and Cancer Research UK (R.C. Travis).

References

1. Thompson IM, Goodman PJ, Tangen CM, et al. The influence of finasteride on the development of prostate cancer. N Engl J Med 2003;349:215–24. [PubMed: 12824459]
2. Noble RL. The development of prostatic adenocarcinoma in Nb rats following prolonged sex hormone administration. Cancer Res 1977;37:1929–33. [PubMed: 858144]
3. Ricke WA, Wang Y, Cunha GR. Steroid hormones and carcinogenesis of the prostate: the role of estrogens. Differentiation 2007;75:871–82. [PubMed: 17924963]
4. Prins GS, Korach KS. The role of estrogens and estrogen receptors in normal prostate growth and disease. Steroids 2008;73:233–44. [PubMed: 18093629]
5. Roddam AW, Allen NE, Appleby P, Key TJ. Endogenous sex hormones and prostate cancer: a collaborative analysis of 18 prospective studies. J Natl Cancer Inst 2008;100:170–83. [PubMed: 18230794]
6. Ring HZ, Lessov CN, Reed T, et al. Heritability of plasma sex hormones and hormone binding globulin in adult male twins. J Clin Endocrinol Metab 2005;90:3653–8. [PubMed: 15755867]
7. Kristensen VN, Andersen TI, Lindblom A, Erikstein B, Magnus P, Borresen-Dale AL. A rare CYP19 (aromatase) variant may increase the risk of breast cancer. Pharmacogenetics 1998;8:43–8. [PubMed: 9511180]
8. Hiramatsu M, Maehara I, Ozaki M, Harada N, Orikasa S, Sasano H. Aromatase in hyperplasia and carcinoma of the human prostate. Prostate 1997;31:118–24. [PubMed: 9140125]
9. Haiman CA, Dossus L, Setiawan VW, et al. Genetic variation at the CYP19A1 locus predicts circulating estrogen levels but not breast cancer risk in postmenopausal women. Cancer Res 2007;67:1893–7. [PubMed: 17325027]
10. Dunning AM, Dowsett M, Healey CS, et al. Polymorphisms associated with circulating sex hormone levels in postmenopausal women. J Natl Cancer Inst 2004;96:936–45. [PubMed: 15199113]
11. Berstein LM, Imyanitov EN, Suspitsin EN, et al. CYP19 gene polymorphism in endometrial cancer patients. J Cancer Res Clin Oncol 2001;127:135–8. [PubMed: 11216915]
12. Setiawan VW, Doherty JA, Shu XO, et al. Two estrogen-related variants in CYP19A1 and endometrial cancer risk: a pooled analysis in the Epidemiology of Endometrial Cancer Consortium. Cancer Epidemiol Biomarkers Prev 2009;18:242–7. [PubMed: 19124504]
13. Modugno F, Weissfeld JL, Trump DL, et al. Allelic variants of aromatase and the androgen and estrogen receptors: toward a multigenic model of prostate cancer risk. Clin Cancer Res 2001;7:3092–6. [PubMed: 11595700]
14. Latil AG, Azzouzi R, Cancel GS, et al. Prostate carcinoma risk and allelic variants of genes involved in androgen biosynthesis and metabolism pathways. Cancer 2001;92:1130–7. [PubMed: 11571725]
15. Mononen N, Seppala EH, Duggal P, et al. Profiling genetic variation along the androgen biosynthesis and metabolism pathways implicates several single nucleotide polymorphisms and their combinations as prostate cancer risk factors. Cancer Res 2006;66:743–7. [PubMed: 16424004]
16. Suzuki K, Nakazato H, Matsui H, et al. Association of the genetic polymorphism of the CYP19 intron 4[TTTA]n repeat with familial prostate cancer risk in a Japanese population. Anticancer Res 2003;23:4941–6. [PubMed: 14981949]
17. Suzuki K, Nakazato H, Matsui H, et al. Genetic polymorphisms of estrogen receptor alpha, CYP19, catechol-O-methyltransferase are associated with familial prostate carcinoma risk in a Japanese population. Cancer 2003;98:1411–6. [PubMed: 14508827]
18. Onsory K, Sobti RC, Al-Badran AI, et al. Hormone receptor-related gene polymorphisms and prostate cancer risk in North Indian population. Mol Cell Biochem 2008;314:25–35. [PubMed: 18483761]

19. Gennari L, Masi L, Merlotti D, et al. A polymorphic CYP19 TTTA repeat influences aromatase activity and estrogen levels in elderly men: effects on bone metabolism. J Clin Endocrinol Metab 2004;89:2803–10. [PubMed: 15181061]
20. Cunningham JM, Hebbring SJ, McDonnell SK, et al. Evaluation of genetic variations in the androgen and estrogen metabolic pathways as risk factors for sporadic and familial prostate cancer. Cancer Epidemiol Biomarkers Prev 2007;16:969–78. [PubMed: 17507624]
21. Cussenot O, Azzouzi AR, Nicolaiew N, et al. Combination of polymorphisms from genes related to estrogen metabolism and risk of prostate cancers: the hidden face of estrogens. J Clin Oncol 2007;25:3596–602. [PubMed: 17704407]
22. dos Santos RM, de Jesus CM, Trindade Filho JC, et al. PSA and androgen-related gene (AR, CYP17, and CYP19) polymorphisms and the risk of adenocarcinoma at prostate biopsy. DNA Cell Biol 2008;27:497–503. [PubMed: 18491956]
23. Huang YC, Chen M, Lin MW, et al. CYP19 TCT tri-nucleotide Del/Del genotype is a susceptibility marker for prostate cancer in a Taiwanese population. Urology 2007;69:996–1000. [PubMed: 17482958]
24. Hunter DJ, Riboli E, Haiman CA, et al. A candidate gene approach to searching for low-penetrance breast and prostate cancer genes. Nat Rev Cancer 2005;5:977–85. [PubMed: 16341085]
25. Calle EE, Rodriguez C, Jacobs EJ, et al. The American Cancer Society Cancer Prevention Study II Nutrition Cohort: rationale, study design, and baseline characteristics. Cancer 2002;94:500–11. [PubMed: 11900235]
26. The ATBC Cancer Prevention Study Group. The alpha-tocopherol, beta-carotene lung cancer prevention study: design, methods, participant characteristics, and compliance. Ann Epidemiol 1994;4:1–10. [PubMed: 8205268]
27. Riboli E, Hunt KJ, Slimani N, et al. European Prospective Investigation into Cancer and Nutrition (EPIC): study populations and data collection. Public Health Nutr 2002;5:1113–24. [PubMed: 12639222]
28. Giovannucci E, Pollak M, Liu Y, et al. Nutritional predictors of insulin-like growth factor I and their relationships to cancer in men. Cancer Epidemiol Biomarkers Prev 2003;12:84–9. [PubMed: 12582016]
29. Kolonel LN, Altshuler D, Henderson BE. The multiethnic cohort study: exploring genes, lifestyle and cancer risk. Nat Rev Cancer 2004;4:519–27. [PubMed: 15229477]
30. Chan JM, Stampfer MJ, Ma J, et al. Insulin-like growth factor-I (IGF-I) and IGF binding protein-3 as predictors of advanced-stage prostate cancer. J Natl Cancer Inst 2002;94:1099–106. [PubMed: 12122101]
31. Hayes RB, Reding D, Kopp W, et al. Etiologic and early marker studies in the prostate, lung, colorectal and ovarian (PLCO) cancer screening trial. Control Clin Trials 2000;21:349S–355S. [PubMed: 11189687]
32. Barrett JC, Fry B, Maller J, Daly MJ. Haploview: analysis and visualization of LD and haplotype maps. Bioinformatics 2005;21:263–5. [PubMed: 15297300]
33. Gabriel SB, Schaffner SF, Nguyen H, et al. The structure of haplotype blocks in the human genome. Science 2002;296:2225–9. [PubMed: 12029063]
34. Stram DO, Haiman CA, Hirschhorn JN, et al. Choosing haplotype-tagging SNPS based on unphased genotype data using a preliminary sample of unrelated subjects with an example from the Multiethnic Cohort Study. Hum Hered 2003;55:27–36. [PubMed: 12890923]
35. Stram DO, Leigh Pearce C, Bretsky P, et al. Modeling and E-M estimation of haplotype-specific relative risks from genotype data for a case-control study of unrelated individuals. Hum Hered 2003;55:179–90. [PubMed: 14566096]
36. Kristensen VN, Harada N, Yoshimura N, et al. Genetic variants of CYP19 (aromatase) and breast cancer risk. Oncogene 2000;19:1329–33. [PubMed: 10713674]
37. Packer BR, Yeager M, Staats B, et al. SNP500Cancer: a public resource for sequence validation and assay development for genetic variation in candidate genes. Nucleic Acids Res 2004;32(Database issue):D528–32. [PubMed: 14681474]
38. Gann PH, Hennekens CH, Ma J, Longcope C, Stampfer MJ. Prospective study of sex hormone levels and risk of prostate cancer. J Natl Cancer Inst 1996;88:1118–26. [PubMed: 8757191]

39. Platz EA, Leitzmann MF, Rifai N, et al. Sex steroid hormones and the androgen receptor gene CAG repeat and subsequent risk of prostate cancer in the prostate-specific antigen era. Cancer Epidemiol Biomarkers Prev 2005;14:1262–9. [PubMed: 15894683]
40. Travis RC, Key TJ, Allen NE, et al. Serum androgens and prostate cancer among 643 cases and 643 controls in the European Prospective Investigation into Cancer and Nutrition. Int J Cancer 2007;121:1331–8. [PubMed: 17514649]
41. Dorgan JF, Albanes D, Virtamo J, et al. Relationships of serum androgens and estrogens to prostate cancer risk: results from a prospective study in Finland. Cancer Epidemiol Biomarkers Prev 1998;7:1069–74. [PubMed: 9865423]
42. Kraft P, Cox D, Paynter R, Hunter D, De Vivo I. Accounting for haplotype uncertainty in association studies: A comparison of simple and flexible techniques. Genet Epidemiol. In press.
43. Zaykin D, Westfall P, Young S, Karnoub M, Wagner M, Ehm M. Testing association of statistically inferred haplotypes with discrete and continuous traits in samples of unrelated individuals. Hum Hered 2002;53:79–91. [PubMed: 12037407]
44. Yeager M, Orr N, Hayes RB, et al. Genome-wide association study of prostate cancer identifies a second risk locus at 8q24. Nat Genet 2007;39:645–9. [PubMed: 17401363]
45. Eriksson AL, Lorentzon M, Vandenput L, et al. Genetic variations in sex steroid-related genes as predictors of serum estrogen levels in men. J Clin Endocrinol Metab 2009;94:1033–41. [PubMed: 19116238]
46. Van Pottelbergh I, Goemaere S, Kaufman JM. Bioavailable estradiol and an aromatase gene polymorphism are determinants of bone mineral density changes in men over 70 years of age. J Clin Endocrinol Metab 2003;88:3075–81. [PubMed: 12843146]
47. Low YL, Taylor JI, Grace PB, et al. Phytoestrogen exposure, polymorphisms in COMT, CYP19, ESR1, and SHBG genes, and their associations with prostate cancer risk. Nutr Cancer 2006;56:31–9. [PubMed: 17176215]
48. Low YL, Taylor JI, Grace PB, et al. Polymorphisms in the CYP19 gene may affect the positive correlations between serum and urine phytoestrogen metabolites and plasma androgen concentrations in men. J Nutr 2005;135:2680–6. [PubMed: 16251630]
49. Haiman CA, Hankinson SE, Spiegelman D, et al. A tetranucleotide repeat polymorphism in CYP19 and breast cancer risk. Int J Cancer 2000;87:204–10. [PubMed: 10861475]
50. Tworoger SS, Chubak J, Aiello EJ, et al. Association of CYP17, CYP19, CYP1B1, and COMT polymorphisms with serum and urinary sex hormone concentrations in postmenopausal women. Cancer Epidemiol Biomarkers Prev 2004;13:94–101. [PubMed: 14744739]
51. Horton R, Hawks D, Lobo R. 3 alpha, 17 beta-androstanediol glucuronide in plasma. A marker of androgen action in idiopathic hirsutism. J Clin Invest 1982;69:1203–6. [PubMed: 7068855]
52. Ellem SJ, Schmitt JF, Pedersen JS, Frydenberg M, Risbridger GP. Local aromatase expression in human prostate is altered in malignancy. J Clin Endocrinol Metab 2004;89:2434–41. [PubMed: 15126575]
53. Ellem SJ, Risbridger GP. Aromatase and prostate cancer. Minerva Endocrinol 2006;31:1–12. [PubMed: 16498360]
54. Feigelson HS, Henderson BE. Future possibilities in the prevention of breast cancer: role of genetic variation in breast cancer prevention. Breast Cancer Res 2000;2:277–82. [PubMed: 11250721]
55. Klouche M. Estrogens in human vascular diseases. Ann N Y Acad Sci 2006;1089:431–43. [PubMed: 17261786]
56. Ding EL, Song Y, Malik VS, Liu S. Sex differences of endogenous sex hormones and risk of type 2 diabetes: a systematic review and meta-analysis. JAMA 2006;295:1288–99. [PubMed: 16537739]
57. Geerlings MI, Strozyk D, Masaki K, et al. Endogenous sex hormones, cognitive decline, and future dementia in old men. Ann Neurol 2006;60:346–55. [PubMed: 16865684]
58. Huang R, Poduslo SE. CYP19 haplotypes increase risk for Alzheimer's disease. J Med Genet 2006;43:e42. [PubMed: 16882736]

Table 1

Characteristics of the study population by study, BPC3.

	ACS CPS-II		ATBC		EPIC		HPFS	
	Cases	**Controls**	**Cases**	**Controls**	**Cases**	**Controls**	**Cases**	**Controls**
Number	1176	1177	1048	1055	760	1179	700	698
Ethnicity (%)*								
White	99	99	100	100	100	100	94	94
African-American								
Native Hawaiian								
Japanese Americans								
Latino								
Age at diagnosis (mean)	70	70	70	69	65	65	69	69
BMI (mean)	26	26	26	26	27	27	25	26
Family history available (n)	1176	1177	914	927	0	0	700	698
Family history (% yes)	21	11	6	3	n/a	n/a	20	15
For cases:								
Years of diagnoses (range)	1992–2002		1986–2003		1991–2003		1994–2000	
Stage info available (n)	1142		651		424		607	
Stage (% ≥ C)	11		31		16		15	
Gleason score available	1009		632		101		618	
Gleason score (% ≥ 8)	11		25		16		10	

	MEC		PHS		PLCO		TOTAL	
	Cases	**Controls**	**Cases**	**Controls**	**Cases**	**Controls**	**Cases**	**Controls**
Number	2320	2290	887	1052	1275	1628	8166	9079
Ethnicity (%)*								
White	20	20	95	94	92	80	75	75
African-American	29	28			8	20	10	11
Native Hawaiian	3	3					1	1
Japanese Americans	20	21					6	5
Latino	28	28					8	7
Age at diagnosis (mean)	68	66	70	70	67	67	68	68

	MEC		PHS		PLCO		TOTAL	
	Cases	**Controls**	**Cases**	**Controls**	**Cases**	**Controls**	**Cases**	**Controls**
Mean BMI	27	27	25	25	27	28	26	27
Family history available (n)	2111	2096	0	0	1275	1628	6176	6526
Family history (% yes)	12	8	n/a	n/a	11	7	14	9
For cases:								
Years of diagnoses (range)	1995–2002		1982–2000		1994–2001		1982–2003	
Stage info avail	2180		698		0		5702	
Stage (% ≥ C)	18		31		n/a		19	
Gleason score available	2234		703		0		5279	
Gleason score (% ≥ 8)	24		13		n/a		18	

*May not add to 100% due to missing data.

Table 2

CYP19A1 htSNPs, missense SNPs and prostate cancer risk among all participants, BPC3.

SNP (minor allele)		Homozygote major allele	Heterozygote	Homozygote minor allele	*P*-value for association*	*P*-value for heterogeneity †
All subjects						
rs2446405 (T)	Cases/Controls	5364/4889	2981/2652	632/539		0.234
	OR (99% CI)	1.00 (ref)	0.98 (0.90,1.07)	0.97 (0.81,1.15)	0.782	
rs2445765 (C)	Cases/Controls	5932/5373	2705/2451	321/264		0.408
	OR (99% CI)	1.00 (ref)	1.01 (0.92,1.10)	0.92 (0.74,1.16)	0.616	
rs2470144 (C)	Cases/Controls	2082/1941	4236/3794	2615/2315		0.092
	OR (99% CI)	1.00 (ref)	0.96 (0.87,1.06)	0.98 (0.87,1.10)	0.558	
rs2445762 (C)	Cases/Controls	4501/4140	3641/3105	702/681		0.162
	OR (99% CI)	1.00 (ref)	0.93 (0.85,1.01)	1.08 (0.92,1.25)	0.014	
rs1004984 (A)	Cases/Controls	3437/3202	4208/3709	1285/1121		0.458
	OR (99% CI)	1.00 (ref)	0.95 (0.87,1.04)	0.96 (0.85,1.09)	0.361	
rs1902584 (T)	Cases/Controls	7523/6801	1413/1214	68/65		0.174
	OR (99% CI)	1.00 (ref)	0.93 (0.83,1.04)	1.10 (0.70,1.74)	0.222	
rs3751591 (G)	Cases/Controls	6280/5659	2348/2096	262/212		0.667
	OR (99% CI)	1.00 (ref)	0.97 (0.88,1.06)	0.90 (0.70,1.15)	0.388	
rs28566535/CV1664178 (C)	Cases/Controls	7116/6450	1572/1351	307/279		0.083
	OR (99% CI)	1.00 (ref)	0.96 (0.85,1.08)	1.02 (0.80,1.31)	0.564	
rs2445759 (T)	Cases/Controls	7742/6970	1176/996	52/53		0.934
	OR (99% CI)	1.00 (ref)	0.95 (0.85,1.08)	1.14 (0.68,1.89)	0.487	
rs936306 (T)	Cases/Controls	5685/5178	2694/2369	583/520		0.060
	OR (99% CI)	1.00 (ref)	0.97 (0.89,1.07)	1.06 (0.88,1.28)	0.469	
rs1902586 (A)	Cases/Controls	7162/6509	1549/1317	292/253		0.041
	OR (99% CI)	1.00 (ref)	0.95 (0.84,1.07)	1.00 (0.77,1.29)	0.536	
rs749292 (A)	Cases/Controls	2899/2623	4434/3889	1598/1535		0.356
	OR (99% CI)	1.00 (ref)	0.97 (0.88,1.06)	1.07 (0.95,1.21)	0.055	
rs6493494 (A)	Cases/Controls	3247/2895	4268/3728	1332/1318		0.381
	OR (99% CI)	1.00 (ref)	0.97 (0.89,1.07)	1.11 (0.98,1.25)	0.017	

SNP (minor allele)		Homozygote major allele	Heterozygote	Homozygote minor allele	*P*-value for association*	*P*-value for heterogeneity †
All subjects						
rs1008805 (G)	Cases/Controls	3210/2953	4250/3745	1491/1346	0.195	0.246
	OR (99% CI)	1.00 (ref)	0.94 (0.86,1.03)	0.95 (0.84,1.07)		
rs727479 (C)	Cases/Controls	3896/3504	3943/3592	1078/930	0.467	0.139
	OR (99% CI)	1.00 (ref)	1.00 (0.92,1.10)	0.94 (0.83,1.08)		
rs2414096 (G)	Cases/Controls	1875/1705	4338/3870	2707/2443	0.731	0.491
	OR (99% CI)	1.00 (ref)	0.98 (0.88,1.08)	1.00 (0.89,1.12)		
rs28757184(A) (Thr201Met)	Cases/Controls	8253/7487	569/439	8/7	0.075	0.989
	OR (99% CI)	1.00 (ref)	0.86 (0.72,1.02)	0.99 (0.26,3.84)		
rs700519 (A) (Arg264Cys)	Cases/Controls	7900/7112	934/840	85/86	0.670	0.331
	OR (99% CI)	1.00 (ref)	1.00 (0.87,1.14)	1.15 (0.76,1.75)		
rs17601241(A)	Cases/Controls	7548/6741	1353/1254	65/78	0.180	0.703
	OR (99% CI)	1.00 (ref)	1.04 (0.93,1.16)	1.33 (0.86,2.07)		
rs10046 (G)	Cases/Controls	2070/1845	4364/3915	2466/2233	0.842	0.424
	OR (99% CI)	1.00 (ref)	1.00 (0.90,1.11)	1.02 (0.91,1.15)		
rs4646 (A)	Cases/Controls	4570/4156	3626/3193	762/705	0.520	0.065
	OR (99% CI)	1.00 (ref)	0.96 (0.89,1.05)	1.00 (0.86,1.16)		

*P-value from 2 d.f. likelihood ratio test for association.

†P-value for likelihood ratio test of heterogeneity of odds ratios across ethnicity (all subjects).

Table 3

CYP19A1 common haplotypes in LD blocks 1, 2, 3 & 4 and prostate cancer risk among all participants, BPC3

Haplotype*		0 copies	1 copy	2 copies	*P*-value for association†	*P*-value for heterogeneity‡
Block 1: global χ^2 = 8.10 on 8 d.f., p = 0.423						**0.057**
AGTTGA	Case/Controls	2370/2656	3835/4287	1934/2073		
	OR (99% CI)	1.00 (ref)	0.97 (0.88,1.07)	1.01 (0.90,1.13)	0.501	
AGCTGA	Case/Controls	7271/8098	839/884	29/34		
	OR (99% CI)	1.00 (ref)	1.09 (0.95,1.25)	1.00 (0.50,1.99)	0.298	
AGCTAA	Case/Controls	7354/8169	758/817	27/30		
	OR (99% CI)	1.00 (ref)	1.03 (0.89,1.19)	0.98 (0.49,2.00)	0.865	
AGCCAA	Case/Controls	5451/5902	2375/2811	312/303		
	OR (99% CI)	1.00 (ref)	0.92 (0.84/1.00)	1.12 (0.90/1.40)	0.011	
TGCCGA	Case/Controls	7586/8366	501/593	52/56		
	OR (99% CI)	1.00 (ref)	0.91 (0.76,1.10)	1.05 (0.62,1.77)	0.431	
TCCTAA	Case/Controls	7432/8186	686/807	21/22		
	OR (99% CI)	1.00 (ref)	0.97 (0.84,1.13)	1.15 (0.51,2.58)	0.817	
TCCTAT	Case/Controls	6871/7567	1208/1385	60/64		
	OR (99% CI)	1.00 (ref)	0.94 (0.84,1.05)	1.07 (0.67,1.72)	0.333	
TCCCGA	Case/Controls	7269/8106	848/877	22/33		
	OR (99% CI)	1.00 (ref)	1.09 (0.95,1.24)	0.74 (0.36,1.53)	0.175	
TGCTAA	Case/Controls	8020/8881	114/134	5/1		
	OR (99% CI)	1.00 (ref)	1.14 (0.78,1.67)	5.42 (0.26,111.40)	0.175	
Block 2: global χ^2 = 2.24 on 4 d.f., p = 0.691						**0.132**
AAGCG	Case/Controls	791/894	3120/3475	4228/4647		
	OR (99% CI)	1.00 (ref)	0.96 (0.82,1.13)	0.97 (0.83,1.14)	0.842	
AATTG	Case/Controls	7653/8473	473/530	13/13		
	OR (99% CI)	1.00 (ref)	1.01 (0.84,1.22)	1.04 (0.37,2.94)	0.979	
ACGTA	Case/Controls	7271/7993	727/851	141/172		
	OR (99% CI)	1.00 (ref)	0.99 (0.83,1.17)	1.00 (0.71,1.40)	0.982	

Haplotype*		0 copies	1 copy	2 copies	*P*-value for association†	*P*-value for heterogeneity‡
Block 1: global χ^2 = 8.10 on 8 d.f., p = 0.423						**0.057**
GAGCG	Case/Controls	7001/7766	1104/1202	34/47		
	OR (99% CI)	1.00 (ref)	1.00 (0.89,1.13)	0.81 (0.45,1.47)	0.664	
GAGTA	Case/Controls	7462/8241	647/738	30/37		
	OR (99% CI)	1.00 (ref)	0.92 (0.79,1.08)	0.85 (0.44,1.63)	0.351	
Block 3: global χ^2 = 2.39 on 3 d.f., p = 0.495						**0.019**
GGG	Case/Controls	3019/3270	3787/4261	1333/1485		
	OR (99% CI)	1.00 (ref)	0.94 (0.86,1.03)	0.94 (0.84,1.07)	0.213	
GGA	Case/Controls	5734/6314	2136/2402	269/300		
	OR (99% CI)	1.00 (ref)	0.98 (0.90,1.08)	1.04 (0.83,1.32)	0.788	
AGA	Case/Controls	7602/8403	503/560	34/53		
	OR (99% CI)	1.00 (ref)	1.06 (0.88,1.27)	0.76 (0.41,1.40)	0.350	
AAA	Case/Controls	3024/3362	3789/4320	1326/1335		
	OR (99% CI)	1.00 (ref)	0.97 (0.89,1.06)	1.10 (0.98,1.25)	0.017	
Block 4: global χ^2 = 9.60 on 8 d.f., p = 0.295						**0.204**
AAGGGAC	Case/Controls	2860/3201	3800/4254	1479/1561		
	OR (99% CI)	1.00 (ref)	0.99 (0.90,1.08)	1.05 (0.93,1.19)	0.326	
AAAGGAC	Case/Controls	7694/8454	438/554	6/8		
	OR (99% CI)	1.00 (ref)	0.87 (0.73,1.04)	0.85 (0.21,3.50)	0.116	
AGGGGAC	Case/Controls	7719/8517	414/491	6/8		
	OR (99% CI)	1.00 (ref)	0.93 (0.77,1.12)	0.81 (0.20,3.33)	0.560	
AGGGGGC	Case/Controls	7745/8550	351/420	43/46		
	OR (99% CI)	1.00 (ref)	1.06 (0.83,1.34)	1.00 (0.55,1.82)	0.843	
AGGGGGA	Case/Controls	8008/8847	127/164	4/4		
	OR (99% CI)	1.00 (ref)	0.93 (0.65,1.32)	1.70 (0.26,11.29)	0.668	
AGGGAGA	Case/Controls	6814/7598	1250/1351	75/66		

Haplotype*		0 copies	1 copy	2 copies	*P*-value for association†	*P*-value for heterogeneity‡
Block 1: global χ^2 = 8.10 on 8 d.f., p = 0.423						**0.057**
AGGAGGC	OR (99% CI)	1.00 (ref)	1.03 (0.92,1.15)	1.24 (0.79,1.93)	0.376	
	Case/Controls	7221/8009	849/938	70/69		
CGGGGGC	OR (99% CI)	1.00 (ref)	1.00 (0.87,1.15)	1.11 (0.69,1.76)	0.855	
	Case/Controls	5994/6679	1959/2131	185/206		
CGGGGGA	OR (99% CI)	1.00 (ref)	1.03 (0.93,1.13)	0.99 (0.75,1.29)	0.765	
	Case/Controls	5397/5913	2429/2745	313/358		
	OR (99% CI)	1.00 (ref)	0.95 (0.87,1.04)	0.94 (0.76,1.16)	0.308	

*Alleles listed for htSNPs in 5' to 3' order: : block 1 rs2446405, rs2445765, rs2470144, rs2445762, rs1004984, rs1902584; block 2 rs3751591, hCV1664178, rs2445759, rs936306, rs1902586; block 3, rs749292, rs64993494, rs1008805; block 4, rs727479, rs2414096, rs28757184 (Thr20Met), rs700519 (Arg264Cys), rs17601241, rs10046, rs4646.

†P-value from 2 d.f. likelihood ratio test for association.

‡P-value for global test of heterogeneity of odds ratios across ethnicity (all subjects).

Table 4

Geometric mean serum estradiol (E2) and testosterone (T) concentrations by htSNPs in LD blocks 3 and 4 and common nonsynonomous SNPs in CYP19A1 among cases and controls

RS Number	Genotype	Estradiol, pmol/L			Free Estradiol, pmol/L			Testosterone, nmol/L			Free testosterone, nmol/L		
		N	Mean (95% CI)	*P* trend	N	Mean (95% CI)	*P* trend	N	Mean (95% CI)	*P* trend	N	Mean (95% CI)	*P* trend
rs749292	GG	683	105.3 (98.0– 113.1)	2.1×10^{-4}	681	4.88 (4.48– 5.32)	1.8×10^{-4}	1542	17.5 (16.0– 19.3)	0.072	1524	7.90 (7.35– 8.49)	0.015
	GA	1047	107.2 (99.9– 115.1)		1046	4.92 (4.55– 5.36)		2341	17.4 (15.8– 19.1)		2304	7.80 (7.28– 8.39)	
	AA	421	113.0 (104.9– 121.7)		420	5.32 (4.88– 5.80)		917	17.0 (15.4– 18.7)		905	7.63 (7.115– 8.25)	
rs6493494 / hCV8234971	GG	695	105.5 (98.2– 113.4)	3.8×10^{-3}	692	4.88 (4.48– 5.32)	2.1×10^{-3}	1611	17.6 (16.0– 19.3)	0.056	1590	7.90 (7.35– 8.53)	6.7×10^{-3}
	GA	1043	107.5 (100.2– 115.3)		1042	4.95 (4.55– 5.36)		2307	17.3 (15.8– 19.0)		2273	7.80 (7.25– 8.39)	
	AA	388	111.6 (103.6– 120.3)		387	5.25 (4.81– 5.73)		846	17.0 (15.4– 18.7)		834	7.63 (7.07– 8.25)	
rs1008805	AA	743	109.8 (102.2– 118.0)	3.3×10^{-3}	742	5.14 (4.70– 5.58)	3.1×10^{-3}	1602	17.1 (15.6– 18.8)	0.089	1581	7.73 (7.18– 8.32)	0.011
	GA	1041	106.8 (99.42– 114.7)		1040	4.92 (4.51– 5.32)		2351	17.6 (16.1– 19.4)		2316	7.94 (7.35– 8.53)	
	GG	356	104.1 (96.5– 112.3)		353	4.84 (4.40– 5.28)		841	17.5 (15.9– 19.3)		830	7.94 (7.38– 8.56)	
rs727479	AA	916	110.5 (103.0– 118.6)	1.2×10^{-5}	913	5.14 (4.70– 5.58)	1.9×10^{-4}	2025	17.1 (15.5– 18.7)	0.037	1996	7.63 (7.07– 8.22)	1.8×10^{-4}
	AC	950	107.3 (100.0– 115.3)		949	4.95 (4.55– 5.39)		2175	17.4 (15.8– 19.1)		2145	7.90 (7.35– 8.52)	
	CC	267	100.8 (93.3– 108.9)		266	4.66 (4.26– 5.14)		587	17.7 (16.0– 19.5)		579	7.94 (7.35– 8.60)	
rs2414096	AA	537	110.0 (102.2– 118.4)	0.066	534	5.10 (4.70– 5.58)	0.094	1218	17.2 (15.6– 18.9)	0.32	1198	7.63 (7.07– 8.22)	7.7×10^{-3}
	GA	1088	107.1 (99.8– 114.9)		1087	4.95 (4.55– 5.36)		2404	17.3 (15.7– 19.0)		2373	7.84 (7.28– 8.42)	
	GG	523	106.4 (98.9– 114.5)		522	4.92 (4.51– 5.39)		1179	17.5 (15.9– 19.2)		1164	7.90 (7.35– 8.53)	
rs28757184 / Thr201Met	GG	1985	107.8 (100.6– 115.5)	0.10	1981	4.99 (4.59– 5.43)	0.055	4426	17.3 (15.8– 19.0)	0.17	4362	7.80 (7.25– 8.39)	0.16
	GA	137	102.0 (93.7– 111.0)	-	137	4.59 (4.15– 5.10)		329	17.9 (16.1– 19.8)		326	7.94 (7.32– 8.60)	
	AA	3	135.8 (96.7– 190.5)	-	3	7.67 (5.10– 11.49)		5	18.4 (12.7– 26.6)		5	10.37 (7.77– 13.80)	
rs700519/Arg264Cys	GG	2000	107.5 (100.2– 115.3)	0.60	1995	4.99 (4.59– 5.39)	0.55	4438	17.5 (15.9– 19.2)	0.13	4374	7.87 (7.32– 8.46)	0.079
	GA	151	108.1 (99.8– 117.1)		151	4.99 (4.51– 5.47)		283	16.8 (15.2– 18.6)		282	7.59 (7.00– 8.22)	
	AA	3	67.5 (48.0– 95.0)		3	3.19 (2.13– 4.77)		22	17.1 (14.0– 20.7)		21	7.73 (6.62– 9.01)	
rs17601241	GG	1799	107.0 (99.8– 114.8)	0.042	1795	4.95 (4.55– 5.36)	0.044	4006	17.4 (15.8– 19.1)	0.22	3950	7.80 (7.28– 8.42)	0.22
	GA	356	110.2 (102.2– 118.8)		356	5.10 (4.66– 5.58)		786	17.2 (15.6– 18.9)		779	7.73 (7.18– 8.36)	
	AA	16	117.0 (997– 137.3)		16	5.65 (4.70– 6.86)		37	16.1 (13.7– 18.8)		35	7.25 (6.38– 8.25)	
rs10046	AA	572	111.5 (103.7– 120.0)	5.7×10^{-3}	569	5.21 (4.77– 5.65)	0.014	1300	17.2 (15.6– 18.9)	0.45	1279	7.70 (7.14– 8.29)	0.033
	GA	1070	106.5 (99.3– 114.3)		1069	4.92 (4.51– 5.36)		2380	17.3 (15.8– 19.0)		2347	7.80 (7.28– 8.42)	
	GG	485	106.3 (98.8– 114.4)		484	4.95 (4.51– 5.39)		1082	17.4 (15.8– 19.1)		1070	7.90 (7.35– 8.53)	

RS Number	Genotype	Estradiol, pmol/L			Free Estradiol, pmol/L			Testosterone, nmol/L			Free testosterone, nmol/L		
		N	Mean (95% CI)	*P* trend	N	Mean (95% CI)	*P* trend	N	Mean (95% CI)	*P* trend	N	Mean (95% CI)	*P* trend
rs4646	CC	1113	107.4 (100.0– 115.2)	0.930	1110	4.95 (4.55– 5.39)	0.46	2567	17.3 (15.8– 19.0)	0.88	2531	7.77 (7.21– 8.36)	0.51
	AC	850	107.7 (100.2– 115.8)		849	5.03 (4.62– 5.47)		1870	17.2 (15.7– 18.9)		1842	7.80 (7.25– 8.39)	
	AA	171	107.2 (98.8– 116.4)		171	4.99 (4.51– 5.51)		363	17.4 (15.7– 19.2)		360	7.87 (7.25– 8.53)	

Table 5

Estradiol and testosterone concentrations* by haplotype A-A (rs749292-rs727479)

	Number of copies of A-A haplotype			*P* trend
	0	1	2	
Estradiol				
N	698	1088	427	1×10^{-5}
Mean	109.5	113.6	119.3	
(95% CI)	(57.3– 209.2)	(60.1– 214.7)	(61.7– 230.6)	
Free estradiol				
N	696	1086	425	6.6×10^{-5}
Mean	1.45	1.55	1.66	
(95% CI)	(0.46–4.55)	(0.51–4.71)	(0.55–4.97)	
Testosterone				
N	1573	2423	929	0.099
Mean	16.2	16.0	15.7	
(95% CI)	(7.4– 35.5)	(7.2– 35.7)	(6.9– 35.4)	
Free testosterone				
N	1555	2385	918	0.020
Mean	2.29	2.32	2.26	
(95% CI)	(0.97–5.39)	(1.01–5.3)	(0.96–5.29)	
Free testosterone/ free estradiol ratio				
N	690	1080	423	8×10^{-6}
Mean	0.67	0.69	0.74	
(95% CI)	(0.30–1.48)	(0.31–1.53)	(0.33–1.67)	

* Hormone values are geometric means and 95% confidence intervals.