INTRODUCTION

In the United States, some 180,400 men were diagnosed with prostate cancer in 2000 alone, and 31,900 men died of the disease during that same year (1). In fact, prostate cancer is the most commonly diagnosed malignancy among men in the United States (1).

Prostate cancer is rare before the age of 30 years, but the rate of increase thereafter is such that by the time a man is 90 years old, he has almost a 100 percent chance of having the disease (2). There is a large variation in prostate cancer incidence rates between racial/ethnic groups in the United States, highest among African-Americans, intermediate among Caucasian-Americans and Hispanic-Americans, and lowest among Asian-Americans (3). These epidemiologic data provide support for a major genetic component to prostate cancer risk, although environmental risk factors could also be partially involved. In addition, epidemiologic and biochemical evidence suggests that variations on androgen loci may play a role in the risk for prostate cancer. For example, a few studies (albeit using mostly small sample sizes) indicate that adult dihydrotestosterone metabolite level variation among distinct racial/ethnic groups parallels their prostate cancer risk (e.g., Ross et al. (4)).

ANDBOGENS AND PROSTATE CANCER

Androgens play an important role in normal and abnormal prostate development. Studies of androgens and prostate cancer go back nearly 60 years and were rewarded with a Nobel Prize (5, 6). Both testosterone and dihydrotestosterone have been shown to induce prostate adenocarcinomas in experimental rat models (7). Normal prostate growth and development are both induced by dihydrotestosterone. In men, testosterone is produced in large amounts by the testes (8). Testosterone is then irreversibly metabolized to dihydrotestosterone, mainly in the prostate gland. Dihydrotestosterone (or, much less efficiently, testosterone) is bound by the androgen receptor (figure 1). This complex then translocates to the cell nucleus where it activates transcription of several genes with androgen-responsive elements in their promoters (8). This process ultimately results in increased DNA synthesis (figure 1). Dihydrotestosterone is known to promote DNA synthesis and cell replication in the prostate (8). Thus, increased prostatic dihydrotestosterone levels may increase the chance for prostate cancer.

The above data suggest that any one of the androgen metabolic loci may be a good candidate gene for prostate cancer predisposition. We chose to review epidemiologic data on four androgen metabolic loci, CYP17, SRD5A2, HSD3B2, and AR (figure 1) because they are the only androgen metabolic genes as of yet for which at least one study has demonstrated a positive association with prostate cancer predisposition.

17α-HYDROXYLASE/17,20 LYASE (CYP17) GENE

Testosterone is synthesized from cholesterol in a series of enzymatic steps involving several of the cytochrome p450 enzymes (9). The enzyme 17α-hydroxylase/17,20 lyase (cytochrome p450c17; EC 1.14.99.9) catalyzes two sequential reactions in the biosynthesis of testosterone, in both the gonads and in the adrenals (figure 1). The first step is the conversion of pregnenolone to 17-hydroxyprogrenenolone (hydroxylase activity), and the second step is its subsequent conversion to C19 steroid dehydroepiandrosterone (lyase activity), a steroid with androgenic activity (9).

The enzyme cytochrome p450c17 is the product of the CYP17 gene (10) (figure 1). The CYP17 gene is located on chromosome 10 and is comprised of eight exons (11). A T to C nucleotide transition (*A2 allele) in the 5’ untranslated region of the CYP17 gene creates a potential Sp1 binding site (CCACC box), 34 bp upstream from the initiation of translation but downstream from the transcription start site (11). Since it is thought that the number of Sp1 binding sites correlates with promoter activity (12), it is conceivable that this polymorphism results in increased transcription of the CYP17 gene. Kristensen et al. (13), however, showed that neither the CYP17*A2 nor the CYP17*A1 allele could form a complex with the Sp1 protein in electrophoretic mobility shift assays. Nevertheless, Lunn et al. (10) demonstrated that the CYP17*A2 allele occurred at a relatively higher frequency in Caucasian-American prostate cancer patients than in matched controls (odds ratio (OR) = 1.7; 95 percent confidence interval (CI): 1.0, 3.0). This finding supports the hypothesis that the CYP17*A2 allele may predispose men to
Molecular Epidemiology of Hormone-Metabolic Loci

Androgen metabolism leading to increased cell division in the prostate. Androgen metabolic genes analyzed in this review are presented in bold type. Steroid abbreviations: Preg, pregnenolone; DHEA, dehydroepiandrosterone; 4-Dione, 4-androstenedione; T, testosterone; DHT, dihydrotestosterone; 3α-Diol, 3α-androstanediol; 3β-Diol, 3β-androstanediol; AR, androgen receptor.

Develop prostate cancer through increased testosterone biosynthesis. Wadelius et al. (14), however, reported higher incidence of the CYP17*A1 allele in Swedish prostate cancer patients than in Swedish controls (OR = 1.6; 95 percent CI: 1.0, 2.5), and cited preliminary data that linked the *A1/*A1 CYP17 genotype to increased circulating androgen levels. This apparent contradiction can perhaps be explained by the existence of distinct prostate cancer predisposing CYP17 mutations that are in linkage disequilibrium with the *A1 allele in some racial/ethnic groups and the *A2 allele in other racial/ethnic groups.

Steroid 5α-reductase type II (SRD5A2) Gene

Steroid 5α-reductase (testosterone 5α-reductase; EC 1.3.99.5) is a membrane-bound enzyme that catalyzes the irreversible conversion of testosterone to dihydrotestosterone (figure 1) with nicotinamide adenine dinucleotide phosphate, reduced form (NADPH) as a cofactor (8). Two isoforms have been characterized so far in humans: the type I enzyme with an alkaline pH optimum, encoded by the SRD5A1 gene, and the type II enzyme with an acidic pH optimum in vitro, encoded by the SRD5A2 gene (15). Thigpen et al. (15) showed that the type I isozyme is expressed primarily in newborn scalp, skin, and liver, while the type II isozyme protein is expressed primarily in genital skin and the prostate gland.

Cloning and characterization of the SRD5A2 gene has demonstrated that it is located on the short arm of human chromosome 2 (band 2p23) and that it is comprised of five exons and four introns (figure 2) (15). A rare disorder of male sexual differentiation, male pseudohermaphroditism due to 5α-reductase deficiency (originally called pseudoovaginal perineoscrotal hypospadias) (16), is caused by inactivating germline mutations in the SRD5A2 gene (17–19). Males affected with the 46 XY karyotype exhibit genital ambiguity and external female characteristics until puberty, at which point there is some development of male secondary sex characteristics, but the prostate remains highly underdeveloped (16). Thus, normal prostate development requires the normal function of the SRD5A2 gene.

Many authors have reported a series of molecular investigations regarding the contribution of the SRD5A2 gene in prostate cancer risk (e.g., Reichart et al. (20), and Makridakis et al. (21, 22). The investigations began by examining a polymorphic (TA)n dinucleotide repeat polymorphism in the 3′ untranslated region. Constitutional DNA missense substitutions are identified below the gene in the single-letter amino acid code.
African-American men (20) who are at the highest risk for prostate cancer in the world. Using mostly Caucasian-American samples, though, Kantoff et al. (23) showed no statistically significant difference in the distribution of this dinucleotide repeat marker among prostate cancer cases and controls. Subsequently, Makridakis et al. (21) reported the identification and characterization of an allelic variant that may decrease prostate cancer risk in Asian men: the V89L substitution, which replaces valine at codon 89 with leucine and reduces enzyme activity both in vitro and in vivo. Makridakis et al. (unpublished data) and others (10, 24), though, have found no statistically significant difference in the distribution of the V89L substitution between prostate cancer cases and controls. Finally, Makridakis et al. (22) reported data on a missense mutation, A49T (alanine-49 to threonine) (figure 2), which is associated with a 7.2-fold increased risk of advanced prostate cancer in African-American men (p = 0.001) and a 3.6-fold increased risk in Latino men (p = 0.04). This mutation also increases steroid 5α-reductase activity fivefold in vitro (22). Jaffe et al. (25) subsequently showed a significant association between the presence of the A49T mutation and a greater frequency of extracapsular prostate cancer as well as a higher pathologic tumor-lymph node-metastasis stage and poor disease prognosis, in mostly Caucasian patients. In conclusion, out of all the SRD5A2 markers examined so far, only the A49T mutation has been associated with significantly increased prostate cancer predisposition.

Interestingly, the SRD5A2 locus also encodes significant pharmacogenetic variation for different competitive steroid 5α-reductase inhibitors, such as finasteride (22, 26). For example, finasteride displays 11.7-fold lower affinity for the A49T mutant enzyme than for the wild-type protein in vitro (22). This pharmacogenetic variation should be taken into account when prescribing steroid 5α-reductase inhibitors for the chemoprevention or treatment of prostatic diseases.

In addition to prostate cancer predisposition, Akalu et al. (27) have also investigated the SRD5A2 gene for its possible contribution to tumor progression. They have reported common somatic genetic alterations in the SRD5A2 gene in prostate tumors at the polymorphic (TA)n dinucleotide repeat in its 3′ untranslated region (27) (figure 2). Thus, the SRD5A2 gene may not only be involved in the predisposition to prostate cancer but also in its progression.

**3β-HYDROXysteroid Dehydrogenase Type II (HSD3B2) Gene**

Dihydrotestosterone (as discussed above) is the most active androgen in the prostate gland (8). It is synthesized from testosterone by the enzyme steroid 5α-reductase and is inactivated through a reductive reaction catalyzed by either 3-α- or 3β-hydroxysteroid dehydrogenase (figure 1) (8). Both reductions use nicotinamide adenine dinucleotide (phosphate), reduced form (NAD(P)H) as a cofactor (8). The product of either reaction, 3α- or 3β-androstanediol, is excreted outside the prostate and circulates in the blood as glucuronide conjugate (11). Thus, the 3β-hydroxysteroid dehydrogenase enzyme (EC 1.1.1.210) is one of the two enzymes that initiate the irreversible inactivation of dihydrotestosterone in the prostate. In addition, 3β-hydroxysteroid dehydrogenase is responsible for the production of androstenedione in the adrenals, which in subsequent steps is converted to dihydrotestosterone in the prostate (figure 1). Thus, 3β-hydroxysteroid dehydrogenase may affect both the synthesis and degradation of dihydrotestosterone. Finally, it has been reported that the activity of the two 3-hydroxysteroid dehydrogenase enzymes is significantly lower in abnormal prostatic tissue (28). Thus, dihydrotestosterone might accumulate in the prostate because of slower degradation (figure 1).

3β-hydroxysteroid dehydrogenase in humans can act on a number of steroid substrates, including dihydrotestosterone (29). This enzyme activity is encoded by two closely linked yet distinct loci: the HSD3B1 and HSD3B2 genes, which are both located in chromosome band 1p13 (30). The type I gene (HSD3B1) encodes the isoform present in the placenta and peripheral tissues, such as skin and mammary gland, while expression of the type II enzyme is restricted to the adrenals and gonads (29, 31).

Cloning and characterization of the human HSD3B2 gene has revealed that it spans about 7.8 kb of genomic DNA in four exons (figure 3) (32). Mutations in this gene cause male pseudohermaphroditism with congenital adrenal hyperplasia (31, 33, 34), but no mutations have yet been reported in prostate cancer patients.

A complex dinucleotide repeat polymorphism has been reported in the third intron of the HSD3B2 gene (35). Vervraet et al. (36) discovered nine alleles of a (TG)n (TA)n(CA)n short tandem repeat (figure 3). Devgan et al. (36) have subsequently identified substantial allelic differences, which include at least 25 different alleles, at the HSD3B2 locus among African-Americans, Asians, and Caucasians. Some of these alleles appear at higher frequencies in prostate cancer cases than in controls, in distinct racial/ethnic groups (36). This racial/ethnic diversity at the HSD3B2 locus suggests that it may play a role in the racial/ethnic variation in prostate cancer risk. Additional

**FIGURE 3.** Schematic representation of the HSD3B2 gene encoding the type II 3β-hydroxysteroid dehydrogenase enzyme. The translation start and stop codons are indicated above the gene. The complex (TG)n (TA)n(CA)n polymorphism in intron 3 is marked below the gene.

Epidemiol Rev Vol. 23, No. 1, 2001
screening of the \textit{HSD3B2} gene should be carried out for the identification of mutations that may predispose to prostate cancer, especially utilizing the high risk short tandem repeat alleles already characterized (36).

\textbf{ANDROGEN RECEPTOR (AR) GENE}

Both testosterone and dihydrotestosterone exert their androgenic effects through the androgen receptor protein (figure 1), a transcription factor that transactivates many genes with androgen receptor elements in their promoters (8). The \textit{AR} gene (figure 4) is located on the long arm of the X-chromosome, and its large exon 1 is responsible for the transactivation activity of its mature gene product (11). Increased activity or levels of the androgen receptor protein are expected to increase prostate cancer risk.

The \textit{AR} gene has been the focus of intense molecular epidemiologic investigations. It appears that a polymorphic polyglutamine stretch also appears to have biochemical consequences in vitro. \textit{AR} receptors with expanded (CAG)$_n$ repeat alleles of the (CAG)$_n$ repeat seem to be associated with a modest (1.5- to 2.1-fold) increase in prostate cancer risk (37–39). Other studies, however, reported no association (40–42), or weak and statistically insignificant association (43), between the presence of shorter (CAG)$_n$ repeat alleles and prostate cancer. Shorter (CAG)$_n$ repeat alleles are related to an earlier age of prostate cancer diagnosis (42, 44). This polymorphic polyglutamine stretch also appears to have biochemical consequences in vitro. \textit{AR} receptors with expanded (CAG)$_n$ repeat alleles cloned from patients with Kennedy disease (an adult-onset X-linked disease associated with low virilization, decreased sperm production, testicular atrophy, and reduced fertility) (11) exhibit 50 percent lower transactivation activity but normal androgen binding activity (11, 45). Thus, it is likely that shorter (CAG)$_n$ repeats are causing increased prostate cancer risk through increased androgen receptor transactivation (in accordance with the model in figure 1).

Two additional \textit{AR} gene polymorphisms, the (GGC)$_n$ trinucleotide repeat and the G1733A (Stu I) single nucleotide polymorphism (figure 4), were evaluated for their role in prostate cancer predisposition. The \textit{Stu I} polymorphism (\textit{Stu I} alleles are designated *S1 and *S2, uncut and cut by the \textit{Stu I} restriction endonuclease, respectively) was in linkage disequilibrium with both the (CAG)$_{n}$ and the (GGC)$_{n}$ trinucleotide repeat polymorphisms, but the two polymorphisms were equilibrated with respect to each other, in African-American controls (11). The *S1 \textit{Stu I} allele was shown to be associated with a threefold higher prostate cancer risk among African-American men under the age of 65 years (11). This *S1 allele prostate cancer-associated risk did not seem to result from just shorter (CAG)$_{n}$ repeats as a function of linkage disequilibrium (11), yet its potential functional significance is unknown. On the other hand, the involvement of the (GGC)$_{n}$ \textit{AR} gene repeat in prostate cancer predisposition has been controversial. Two studies have found no association (41, 46), another study demonstrated a modest increase in prostate cancer risk for short (less than 16) repeat alleles (43), while yet another study showed a modest increase in prostate cancer relapse risk for long (more than 16) (GGC)$_{n}$ repeat alleles (40).

In conclusion, some of the genetic variants studied in the androgen receptor gene have been shown to cause increased predisposition to prostate cancer.

Finally, somatic mutations in the \textit{AR} gene, including \textit{AR} gene amplification, have been shown to play a role during prostate tumor progression as well (e.g., reviewed by Barrack et al. (47); see also Tulley et al. (48), Culig et al. (49), and Visakorpi et al. (50)).

\textbf{OTHER GENES}

Another androgen metabolism gene that may play a role in prostate cancer predisposition is \textit{HSD17B3} encoding 17\textbeta-hydroxysteroid dehydrogenase type III, the testicular 17\textbeta-ketoreductase (51). This enzyme utilizes nicotinamide adenine dinucleotide phosphate, reduced form (NADPH) as a cofactor and favors the reduction of androstenedione to testosterone, over the reverse reaction (producing androstenedione), in the testis (51). Thus, mutations in this locus may increase the production of testosterone in the testis, which can directly or indirectly (through dihydrotestosterone) activate the androgen receptor, augmenting prostate cancer risk. The \textit{HSD17B3} gene is on chromosomal region 9q22 and contains 11 exons (51). No mutations of this gene have yet been reported in prostate cancer, but several inactivating \textit{HSD17B3} mutations have been reported to cause another form of male pseudohermaphroditism called 17\textbeta-hydroxysteroid dehydrogenase deficiency (19, 51, 52).

Estrogens are also suspected to contribute to the etiology of prostate cancer and benign prostatic hypertrophy (e.g., Bosland (6), Tsugaya et al. (53), Bonkhoff et al. (54)). Aromatase cytochrome p450, the product of the \textit{CYP19} gene, is mainly localized in the ovary and the placenta, and it converts androgen to estrogen (55). Aromatase mRNA and protein have both been detected in benign prostatic hyper trophy and prostate cancer tissue, but not at significantly different levels between cancerous and noncancerous tissue (53, 56). Several \textit{CYP19} polymorphisms have been investigated in breast cancer (e.g., Probst-Hensch et al. (57), Kristensen et al. (58)), but none in prostate cancer. With respect to the estrogen receptor genes, the classic estrogen receptor (ER\textalpha), but not ER\textbeta, has been shown to be expressed in prostate cancer and premalignant prostatic lesions (e.g., Bonkhoff et al. (54)), yet again no ER\textgamma polymorphisms have been investigated in prostate cancer.
CONCLUSIONS

Despite the ample evidence linking androgens and prostate cancer, convincing molecular epidemiologic evidence for the involvement of androgen metabolic loci in prostate cancer exists for only one single point mutation so far, the A49T mutation in the SRD5A2 gene. Two independent studies confirm a significant association between the presence of the A49T mutation and the presence of advanced prostate cancer in three racial/ethnic groups. Mechanistic biochemical evidence also exists that provides a plausible etiology for this increased prostate cancer risk: enhanced dihydrotestosterone production of the mutant allele (see model in figure 1). Other polymorphisms, though, yielded mixed results. The (CAG)n repeat in the AR gene, for example, has been shown by several studies both to be associated with, and not to be associated with, prostate cancer in Caucasians, even though mechanism biochemical data supportive of an etiologic association exists. The A1 polymorphism of the CYP17 gene yielded conflicting results in two epidemiologic studies.

Some of the epidemiologic studies reviewed here are plagued by various problems in their experimental design: small sample size, absent or inadequate stratification by age between cases and controls, and inclusion of cases diagnosed before and after the availability of prostate-specific antigen testing, in the same study set. The latter study design creates variability of disease diagnosis and can be circumvented by stratification by year at diagnosis. However, even discounting all of those problematic studies does not lead to a clear understanding of the role of, for example, the (CAG)n androgen receptor repeat in prostate cancer predisposition. Any epidemiologic study can suffer from publication bias; null studies are much less likely to be submitted for publication. Yet, prostate cancer epidemiologists face additional obstacles that may explain most of the observed discrepancies. First, prostate cancer is very common. At autopsy, 64 percent of men aged 60-69 years are shown to have prostate adenocarcinoma (59). This fact increases the probability that many of the study controls may actually be undiagnosed cases. This, in turn, will reduce the “observed” relative risk for any polymorphism studied, yet it would mostly impact polymorphisms with low penetrance (low “actual” relative risk) to the point that some studies find a weak positive association while other studies find no association at all. Second, it is now recognized that prostate cancer involves more loci than other common cancers, such as breast cancer (60). This enhanced genetic heterogeneity can effectively reduce the contribution of each particular genetic modifier so that, for example, larger sample sizes or very homogeneous (inbred) populations may be required to positively identify individual genes. It also makes it more likely that distinct genes would cause disease in each racial/ethnic group, rendering comparisons between, for example, US and European Caucasians less conclusive.

Finally, prostate cancer epidemiologists studying androgen metabolic genes should not discount the effects of environmental modifiers that may be metabolized by some of the androgen metabolic enzymes. Diet, and in particular animal fat intake (which contains cholesterol and steroids), may differ significantly among distinct study populations, and, thus, may account for some of the observed discrepancies (11).

ACKNOWLEDGMENTS

This work was supported by grant nos. CA68581 and CA83112 from the National Cancer Institute, grant no. 99-0053V-10150 from the California Cancer Research Program, and from the T. J. Martell Foundation to one of the authors (J.K.V.R.).

REFERENCES


