Germ-Line Mutation of \textit{NKK3.1} Cosegregates with Hereditary Prostate Cancer and Alters the Homeodomain Structure and Function

S. Lilly Zheng,\textsuperscript{1} Jeong-ho Ju,\textsuperscript{2} Bao-li Chang,\textsuperscript{1} Elizabeth Ortner,\textsuperscript{2} Jielin Sun,\textsuperscript{1} Sarah D. Isaacs,\textsuperscript{3} Jishang Sun,\textsuperscript{1} Kathy E. Wiley,\textsuperscript{1} Wennuan Liu,\textsuperscript{1} Micheas Zemedkun,\textsuperscript{2} Patrick C. Walsh,\textsuperscript{3} James Ferretti,\textsuperscript{1} James Gruschus,\textsuperscript{1} William B. Isaacs,\textsuperscript{3} Edward P. Gelmann,\textsuperscript{2} and Jianfeng Xu\textsuperscript{1}

\textsuperscript{1}Center for Human Genomics, Wake Forest University School of Medicine, Winston-Salem, North Carolina; \textsuperscript{2}Departments of Oncology and Medicine, Lombardi Comprehensive Cancer Center, Georgetown University, Washington, District of Columbia; \textsuperscript{3}Department of Urology, Johns Hopkins Medical Institutions, Baltimore, Maryland; and \textsuperscript{4}Laboratory of Biophysical Chemistry, National Heart, Lung, and Blood Institute, Bethesda, Maryland

Abstract

\textit{NKK3.1}, a gene mapped to 8p21, is a member of the NK class of homeodomain proteins and is expressed primarily in the prostate. \textit{NKK3.1} exerts a growth-suppressive and differentiating effect on prostate epithelial cells. Because of its known functions and its location within a chromosomal region where evidence for prostate cancer linkage and somatic loss of heterozygosity is found, we hypothesize that sequence variants in the \textit{NKK3.1} gene increase prostate cancer risk. To address this, we first resequenced the \textit{NKK3.1} gene in 159 probands of hereditary prostate cancer families recruited at Johns Hopkins Hospital; each family has at least three first-degree relatives affected with prostate cancer. Twenty-one germ-line variants were identified in this analysis, including one previously described common nonsynonymous change (R52C), two novel rare nonsynonymous changes (A17T and T164A), and a novel common 18-bp deletion in the promoter. Overall, the germ-line variants were significantly linked to prostate cancer, with a peak heterogeneity logarithm of odds of 2.04 ($P = 0.002$) at the \textit{NKK3.1} gene. The rare nonsynonymous change, T164A, located in the homeobox domain of the gene, segregated with prostate cancer in a family with three affected brothers and one unaffected parent. Importantly, nuclear magnetic resonance solution structure analysis and circular dichroism studies showed this specific mutation to affect the stability of the homeodomain of the \textit{NKK3.1} protein and decreased binding to its cognate DNA recognition sequence. These results suggest that germ-line sequence variants in \textit{NKK3.1} may play a role in susceptibility to hereditary prostate cancer and underscore a role for \textit{NKK3.1} as a prostate cancer gatekeeper. (Cancer Res 2006; 66(1): 69-77)

Introduction

Prostate cancer is the most common cancer and the second leading cause of cancer death among men in the United States. In 2003, an estimated 220,900 men were diagnosed with prostate cancer in the United States and 28,900 died of the disease (1). Although the etiology of prostate cancer is unknown, age, race/ethnicity, and family history are three well-established risk factors. Furthermore, evidence for a major prostate cancer susceptibility gene has been consistently provided by multiple segregation studies (2). Several chromosomal regions likely harbor prostate cancer susceptibility genes, and several candidate genes in these linkage regions have been reported (2, 3).

Suggestive evidence for prostate cancer linkage at chromosome 8p has been reported independently in multiple study populations of familial prostate cancer (4–9). The likelihood of a gene playing an important role in prostate carcinogenesis in this region is strengthened by the accumulated evidence that 8p is the site of the most frequent loss of heterozygosity (LOH) in prostate cancer tumors (10). \textit{Macrophage scavenger receptor 1} (\textit{MSR1}) is a candidate prostate cancer susceptibility gene (11) that was identified during systematic evaluations of mutations in at 8p22-23 by screening for mutations in probands of 190 hereditary prostate cancer families ascertained at Johns Hopkins Hospital and by family-based linkage and association tests in hereditary prostate cancer families. Seven rare nonsynonymous mutations, including one nonsense mutation (R293X), were identified in 13 of the 190 (6.8%) hereditary prostate cancer families studied. These mutations cosegregate with prostate cancer in nuclear families where one parent is an obligate carrier. However, the evidence for linkage at the 8p region remains after removing the families with \textit{MSR1} mutations, suggesting that there are other major prostate cancer susceptibility genes in the region. Evaluation of additional candidate genes at 8p22-23 region in hereditary prostate cancer families found a weak association between prostate cancer risk and sequence variants in \textit{leucine zipper tumor suppressor} (\textit{LZTS1}), originally termed \textit{FEZ1}; ref. 12) and no association between prostate cancer risk and sequence variants in \textit{deleted in liver cancer} (\textit{DLC1}; ref. 13) and \textit{PINX1} (14).

To continue the search for major genes at this region by fine mapping linkage analysis and evaluation of candidate genes, we found the evidence for linkage at 8p extends to 8p21. The evidence for linkage at 8p21 was also observed in the 139 hereditary prostate cancer families ascertained in Germany (9) and the 57 hereditary prostate cancer families ascertained in Sweden (6). An excellent candidate gene at 8p21 is \textit{NKK3.1}, a homeodomain-containing transcription factor that is expressed in a largely prostate-specific and androgen-regulated manner.

Note: S.L. Zheng, J-h. Ju, and B-L Chang contributed equally to this work. Supplementary data for this article are available at Cancer Research Online (http://cancerres.aacrjournals.org/).

Requests for reprints: William B. Isaacs, Department of Urology, Johns Hopkins University School of Medicine, 600 North Wolfe Street, Marburg 115, Baltimore, MD 21287. Phone: 410-955-2518; Fax: 410-955-0833; E-mail: wisaacs@jhmi.edu.

©2006 American Association for Cancer Research.
NYKX3.1 is expressed specifically in prostate luminal epithelial cells and undergoes progressive loss of protein expression with prostate cancer progression to hormone independence and metastases (15). The NYKX3.1 gene is not subject to somatic point mutation in prostate cancer (16, 17), although it does undergo frequent LOH and nonconventional promoter methylation (18). Importantly, gene-targeting studies in mice showed that NYKX3.1 haploinsufficiency can predispose to prostate epithelial dysplasia and can cooperate with other oncogenic mutations to augment carcinogenesis (19, 20). In gene-targeted mice, NYKX3.1 haploinsufficiency is accompanied by decreased expression of genes under the regulation of the NYKX3.1 homeoprotein (21). The data are highly suggestive that NYKX3.1 both is a tumor suppressor protein and is inactivated at the earliest stages of human prostate cancer.

In this study, we tested the hypothesis that mutations and sequence variants in NYKX3.1 might increase prostate cancer risk and partially account for the linkage evidence at 8p. We first resequenced coding and potential regulatory regions of the NYKX3.1 gene for mutations and sequence variants in 159 hereditary prostate cancer probands. We then tested for cosegregation between a subset of identified mutations and sequence variants of NYKX3.1 and prostate cancer among 188 hereditary prostate cancer families. Among the variants identified, one mutant that cosegregated with prostate cancer and was predicted to affect the structure of the protein was selected for detailed structural and functional analyses. These analyses revealed this mutation to affect the stability of the homeodomain and to decrease the DNA-binding affinity of the encoded NYKX3.1 protein.

Materials and Methods

Subjects. All 188 hereditary prostate cancer families, each having at least three individuals affected with prostate cancer and who were first-degree relatives, were recruited and studied at the Brady Urology Institute at Johns Hopkins Hospital (Baltimore, MD). A detailed description of this series of hereditary prostate cancer families has been presented elsewhere (22). Briefly, prostate cancer diagnosis was verified by medical records for each affected male studied. The mean age at diagnosis was 64.4 years for the cases in these families. The number of families with three, four, and five or more affected individuals were 28, 47, and 113, respectively. There were 169 European American families and 17 African American families. Informed consent was obtained from all participants, and study protocols were reviewed and approved by the institutional review boards at each institution.

Sequencing methods and single nucleotide polymorphism genotyping. We directly sequenced the PCR products of ~500 bp of the promoter, all exons, and exon-intron junctions in 159 hereditary prostate cancer probands. The primers used to amplify the PCR products, the sizes of amplified PCR fragments, and the annealing temperatures for each pair of primers are available at the author’s Web site. All PCR reactions were done in a 10-μL volume consisting of 10 ng genomic DNA, 0.2 mmol/L of each primer, 0.2 mmol/L of each deoxynucleotide triphosphate, 1.5 mmol/L MgCl₂, 20 mmol/L Tris-HCl, 50 mmol/L KCl, and 0.5 units Taq polymerase (Invitrogen Corporation, Carlsbad, CA). PCR cycling conditions were as follows: 94°C for 4 minutes followed by 30 cycles of 94°C for 30 seconds, specified annealing temperature for 30 seconds, and 72°C for 30 seconds, with a final extension of 72°C for 6 minutes. All PCR products were purified using the QuickStep PCR Purification kit (Edge BioSystems, Gaithersburg, MD) to remove deoxynucleotide triphosphates and excess primers. All sequencing reactions were done using dye terminator chemistry (BigDye, ABI, Foster City, CA) and then precipitated using 60% ethanol. Samples were loaded onto an ABI 3700 DNA analyzer after adding 8 μL formamide. Single nucleotide polymorphisms (SNP) were identified using Sequencher software version 4.0.5 (Gene Codes Corp., Ann Arbor, MI).

Genotyping was done using the MassARRAY system (Sequenom, Inc., San Diego, CA). For the MassARRAY assay, PCR and extension primers for sequence variants were designed using SpectroDesigner software (Sequenom). The primer information is available at the author’s Web site. PCR and extension reactions were done according to the manufacturer’s instructions, and extension product sizes were determined by mass spectrometry.

Bioinformatics analysis. To predict whether promoter sequence variants alter putative transcription factor–binding sites, we used the Transcription Element Search System program to search the TRANSFAC database.

The program ESEfinder was employed to identify potential exonic splicing enhancer motifs, which could affect the splicing efficiency of pre-mRNA, in the exonic sequence. For the nonsynonymous changes, two algorithms were used to predict the effect of amino acid substitutions on protein activity. The Sorting Intolerant from Tolerant (SIFT) algorithm emphasizes sequence homology among related genes and domains over evolutionary time as well as the characteristics of the amino acid residues (23). The Polymorphism Phenotyping (PolyPhen) algorithm incorporates sequence conservation and the nature of the amino acid residues involved and accounts for the location of the substitution within identified functional domains and known structures and structural features of the protein based on information available in the SwissProt annotated database (24).

Statistical methods. Multipoint linkage analyses were done using both parametric and nonparametric methods implemented by the computer program GENEHUNTER-PLUS (25, 26). For the parametric analysis, the same autosomal dominant model that was used by Smith et al. (27) was assumed. Linkage in the presence of heterogeneity was assessed by use of an admixture test for heterogeneity (28).

Family-based association tests were done for the SNPs genotyped in the 188 hereditary prostate cancer families using the FBAT software package (29). FBAT uses data from nuclear families, sibships, or a combination of the two to test for linkage and linkage disequilibrium (association) between traits and genotypes. The test for linkage is valid when multiple affected members per pedigree are used, and the power to detect linkage is increased if there is an association. The test for association is valid if one affected member per pedigree is used (the genotypes of all the affected members can be included) or if the empirical variance is used to account for correlation between transmissions in families when linkage is present.

Plasmid construction. The pET32a vector (Novagen, Madison, WI) constructs expressed the NYKX3.1 homeodomain plus 10 NH₂-terminal amino acids, NYKX3.1(114-184), and COOH-terminal truncated NYKX3.1(1-184) as NH₂-terminal fusion peptides with the TRX B protein and hexahistidine sequences. The T164A mutation was generated using a QuickChange Site-Directed Mutagenesis kit (Stratagene, La Jolla, CA) according to the manufacturer’s protocol. All NYKX3.1 expression constructs were fully sequenced to confirm the presence of mutations and ensure that no additional nucleotide changes had been introduced.

Protein expression and purification. Plasmid constructs were used to transform competent Escherichia coli BL21 (DE3) cell. The proteins were overexpressed in E. coli BL21 (DE3) cell in LB medium. Uniformly ¹⁵N-labeled proteins were produced by culturing the cells in minimal medium containing [¹⁵N]NH₄Cl as the sole source of nitrogen for the nuclear magnetic resonance (NMR) experiments. Cells were grown at 37°C to an absorbance (A₅₇₀nm) of ~0.6 and then induced with 1 mmol/L...
isopropyl-β-D-thio-β-D-galactopyranoside for 4 hours. The cells were suspended in 20 mmol/L Tris buffer (pH 8.0), 500 mmol/L NaCl, and 0.125% Triton X-100 and then lysed with sonication. The proteins were isolated by Ni²⁺ affinity column chromatography and subsequently purified using gel filtration (HiLoad 16/60) and cation-exchange (HiTrap SP HP) columns with the AKTA fast protein liquid chromatography system (Amersham Biosciences, Piscataway, NJ). For NMR and circular dichroism studies, all the fused protein and tag sequences were removed by enterokinase treatment after protein purification. The homogeneity of purified proteins was confirmed by gel electrophoresis and estimated to be ≥95% in each case. The molecular mass of each purified proteins was checked by mass spectroscopy to confirm protein identity and size.

**NMR spectroscopy.** Four ¹⁵N-labeled samples were prepared, the wild-type and the T164A mutant NKX3.1 homodomains, both free and bound to the NK2 consensus DNA 16-mer duplex 5'-TGTTGCAAGTGCTGT-3' at a 1.0:1.1 protein/DNA molar ratio (30).

All samples had concentrations between 250 and 300 μmol/L (pH 6.5), 80 mmol/L NaCl, and 10 mmol/L sodium phosphate. NMR spectra were obtained at 285 K for the free samples and 308 K for the DNA-bound samples. Two-dimensional ¹⁵N heteronuclear spin quantum correlation (HSQC) and three-dimensional ¹⁵N-separated HSQC nuclear Overhauser effects spectroscopy (NOESY) spectra (31) were collected for each sample on a Bruker Avance 800 spectrometer equipped with a triple-axis gradient TXI probe (31). Spectra were processed and analyzed with the nmrPipe package (32).

**Circular dichroism spectroscopy.** Circular dichroism measurements were done in a Jasco J-715 spectropolarimeter equipped with PTC-348WI temperature control unit interfaced to a computer. The 11.3 \( g \) gels at 175 V for 2 hours at ambient temperature in 0.5 M EDTA buffer after pre-electrophoresis of the gel at 200 V for 30 minutes. The reactions were analyzed on native 8% polyacrylamide gels at 175 V for 2 hours at ambient temperature in 0.5 \( g \) Tris-borate EDTA buffer after pre-electrophoresis of the gel at 200 V for 30 minutes. Kodak BioMax XAR film was exposed to the dried gel and the bands were quantified using Scion Image software and confirmed with density analysis in Corel Photopaint.

**Results**

**Identification of mutations and sequence variants in NKX3.1.** The NKX3.1 gene consists of two exons and spans ~4.2-kb genomic DNA (NC_000008). The full length of mRNA transcript of this gene is 3,266 bp (NM_006167), with a short
Table 1. Identified sequence variants in the NKX3.1 gene

<table>
<thead>
<tr>
<th>Variant ID</th>
<th>Position</th>
<th>Amino acid change</th>
<th>Minor allele</th>
<th>Frequency</th>
</tr>
</thead>
<tbody>
<tr>
<td>−507AGn</td>
<td>Promoter</td>
<td>AGn</td>
<td>0.474</td>
<td></td>
</tr>
<tr>
<td>−360G/A</td>
<td>Promoter</td>
<td>A</td>
<td>0.2</td>
<td></td>
</tr>
<tr>
<td>−276 18-bp INDEL</td>
<td>Promoter</td>
<td>del</td>
<td>0.167</td>
<td></td>
</tr>
<tr>
<td>−46G/T</td>
<td>5′-UTR</td>
<td>T</td>
<td>0.005</td>
<td></td>
</tr>
<tr>
<td>−15A/G</td>
<td>5′-UTR</td>
<td>G</td>
<td>0.185</td>
<td></td>
</tr>
<tr>
<td>49G/A</td>
<td>Coding</td>
<td>A17T</td>
<td>0.005</td>
<td></td>
</tr>
<tr>
<td>154C/T</td>
<td>Coding</td>
<td>R52C</td>
<td>0.05</td>
<td></td>
</tr>
<tr>
<td>1454A/G</td>
<td>Coding</td>
<td>T164A</td>
<td>0.003</td>
<td></td>
</tr>
<tr>
<td>1495T/C</td>
<td>Coding</td>
<td>C</td>
<td>0.01</td>
<td></td>
</tr>
<tr>
<td>1558G/A</td>
<td>Coding</td>
<td>P198P</td>
<td>0.003</td>
<td></td>
</tr>
<tr>
<td>1870T/C</td>
<td>3′-UTR</td>
<td>T</td>
<td>0.308</td>
<td></td>
</tr>
<tr>
<td>1977T/C</td>
<td>3′-UTR</td>
<td>T</td>
<td>0.141</td>
<td></td>
</tr>
<tr>
<td>2179AGn</td>
<td>3′-UTR</td>
<td>AGn</td>
<td>0.495</td>
<td></td>
</tr>
<tr>
<td>2395G/A</td>
<td>3′-UTR</td>
<td>A</td>
<td>0.46</td>
<td></td>
</tr>
<tr>
<td>2553T/C</td>
<td>3′-UTR</td>
<td>C</td>
<td>0.011</td>
<td></td>
</tr>
<tr>
<td>2813C/A</td>
<td>3′-UTR</td>
<td>A</td>
<td>0.008</td>
<td></td>
</tr>
<tr>
<td>3318G/T</td>
<td>3′-UTR</td>
<td>T</td>
<td>0.007</td>
<td></td>
</tr>
<tr>
<td>3397G/C</td>
<td>3′-UTR</td>
<td>C</td>
<td>0.003</td>
<td></td>
</tr>
<tr>
<td>3992G/C</td>
<td>3′-UTR</td>
<td>G</td>
<td>0.004</td>
<td></td>
</tr>
<tr>
<td>4071A/G</td>
<td>3′-UTR</td>
<td>G</td>
<td>0.004</td>
<td></td>
</tr>
<tr>
<td>4237T/C</td>
<td>3′-UTR</td>
<td>C</td>
<td>0.007</td>
<td></td>
</tr>
</tbody>
</table>

5′-untranslated region (UTR) of 48 nucleotides and a long 3′-UTR of 2,514 bp. In addition to this annotated mRNA sequence in National Center for Biotechnology Information, there are four additional isoforms of NKX3.1 (AF_249669, AF_249670, AF_238671, and AF_249672) that are a result of alternative splicing (33) as shown in Fig. 1. These four isoforms encode proteins that vary only in the NH2-terminal region upstream of the homeobox. The in-frame amino acid deletions of the various isoforms range from 44 to 75 residues compared with the full-length protein product (NP_006158). A total of 21 sequence variants were identified in the sequenced region of NKX3.1 among 159 hereditary prostate cancer probands, including 18 SNPs and 3 insertion/deletion variants. The frequency and location of each SNP is listed in Table 1 and Fig. 1. Nine of the 21 variants are commonly observed, with minor allele frequencies of ≥5%. Three of five SNPs in the coding regions result in amino acid substitutions, 49G/A (A17T), 154C/T (R52C), and 1454T/C (T164A). Among these nonsynonymous changes, R52C has been reported previously in both European Americans and African Americans (16) and was associated with aggressive prostate cancer risk (34). The other two nonsynonymous changes are novel. The substitution A17T was observed in three probands, all of which are African American. The substitution T164A is in the homeobox domain of the gene and was only observed once in a European American proband. This nonsynonymous change is predicted to damage protein function by both SIFT and PolyPhen algorithms. T164 is well conserved among other NK-like homeodomain sequences (35). Two other coding SNPs are synonymous changes. The SNP 1495T/C (Y177Y) was observed in three African American families. The SNP 1558G/A (P198P) was observed once in a European American family. Neither of these synonymous SNPs is located in exonic splicing enhancer and may not have significant effect on pre-mRNA splicing. Among the three variants in the promoter region, the −276 18-bp INDEL variant (GAAGGAGAGAAATTTGGG) contains a consensus 5′-GGAA-3′ core sequence of PU.1-binding domain (36). No consensus-binding sites of transcription factors were found for the other two promoter variants. However, the SNP −360 G/A is located in a 180-bp fragment that was reported to contain some negative regulatory elements (37). For the 10 SNPs located in the 3′-UTR, none of these are located in the pentamer AUUUA or nonamer UUAUUUAU of AU-rich regulatory sequence motifs.

Cosegregation of mutations and sequence variants with prostate cancer in hereditary prostate cancer families. Nine of the 21 identified sequence variants of NKX3.I, including the three nonsynonymous changes in the coding region and the 18-bp INDEL in the promoter, were selected for genotyping among all members of 188 hereditary prostate cancer families for which DNA samples were available. Three approaches were used to assess the cosegregation of these sequence variants with prostate cancer risk. A fine mapping linkage analyses at 8p21 was done using a subset of families. The SNP 1558G/A (P198P) was observed once in a European American proband. This nonsynonymous change is predicted to damage protein function by both SIFT and PolyPhen algorithms. T164 is well conserved among other NK-like homeodomain sequences (35). Two other coding SNPs are synonymous changes. The SNP 1495T/C (Y177Y) was observed in three African American families. The SNP 1558G/A (P198P) was observed once in a European American family. Neither of these synonymous SNPs is located in exonic splicing enhancer and may not have significant effect on pre-mRNA splicing. Among the three variants in the promoter region, the −276 18-bp INDEL variant (GAAGGAGAGAAATTTGGG) contains a consensus 5′-GGAA-3′ core sequence of PU.1-binding domain (36). No consensus-binding sites of transcription factors were found for the other two promoter variants. However, the SNP −360 G/A is located in a 180-bp fragment that was reported to contain some negative regulatory elements (37). For the 10 SNPs located in the 3′-UTR, none of these are located in the pentamer AUUUA or nonamer UUAUUUAU of AU-rich regulatory sequence motifs.

Cosegregation of mutations and sequence variants with prostate cancer in hereditary prostate cancer families. Nine of the 21 identified sequence variants of NKX3.I, including the three nonsynonymous changes in the coding region and the 18-bp INDEL in the promoter, were selected for genotyping among all members of 188 hereditary prostate cancer families for which DNA samples were available. Three approaches were used to assess the cosegregation of these sequence variants with prostate cancer risk. A fine mapping linkage analyses at 8p21 was done using a subset of sequence variants within NKX3.I and microsatellite markers (D8S1752 and D8S1771), which lie on either side of the gene, spanning an interval of ~3.3 Mb. Because fine mapping linkage analyses using markers that are in strong linkage disequilibrium may inflate evidence for linkage, only the two most informative NKX3.I variants (−507AGn and 2179AGn) that are not in significant pair-wise linkage disequilibrium (P = 0.09) were included in the analysis to reduce the effect of marker linkage disequilibrium on linkage analysis. Evidence for linkage was observed and the...
The diagnostic part of $^{1}H-^{15}N$ HSQC spectra of free and DNA-bound homeodomains are superimposed for wild-type (A) and T164A mutant (B) NKX3.1. Red, free protein resonances; black, DNA-bound protein resonances. All the resonances of the NKX3.1 homeodomain protein were assigned by comparison with the NK2 homeodomain combined with the three-dimensional $^{15}N$-edited HSQC NOESY experiments. The chemical shift change of side chain Trp $^{111}$ (homeodomain residue 48) and appearance of side chain Asn $^{174}$ (homeodomain residue 51) resonances are indicative of specific DNA binding. Note: The complete $^{1}H-^{15}N$ HSQC spectra of free and T164A mutant NKX3.1 homeodomains are shown in Supplementary Fig. S1.

strongest at the 2179AGn variant of NKX3.1 (heterogeneity logarithm of odds = 2.04; $P = 0.002$). This linkage evidence at this region was stronger than the linkage analysis without NKX3.1 variants (heterogeneity logarithm of odds = 0.66; $P = 0.08$). These linkage results suggest that sequence variants at this region are linked to prostate cancer either due to causal variant(s) within the NKX3.1 gene or within other genes in the region. In the second approach, we assessed the association of the six common sequence variants ($-$507AGn$-$$276$ 18-bp INDEL, 154C/T, 1870T/C, 2179AGn, and 2553T/C) with prostate cancer risk using a family-based association test. However, these results do not exclude the possibility that several of these common variants, or rare variants, confer increased prostate cancer risk.

Finally, for the three rare sequence variants where a formal family-based association test is inappropriate due to the small number of subjects that carry these variants, we examined their cosegregation with prostate cancer in each family where there is at least one subject with the variant. Whereas the nonsynonymous change of 49G/A (A17T) and the synonymous change 1495T/C (Y177Y) did not cosegregate with prostate cancer in the African American families, the nonsynonymous change 1454A/G (T164A) cosegregated completely with prostate cancer. This nonsynonymous change was observed in one European American family. As shown in Fig. 2, all three affected sons diagnosed with prostate cancer at ages 55, 59, and 67 years were heterozygous carriers of the alanine containing allele at this locus. The men diagnosed at ages 55 and 67 years had non-organ-confined disease at the time of treatment, whereas the remaining brother, diagnosed at age 59 years, had cancer in both lobes of the prostate. The only unaffected son (age 70 years) was homozygous for the wild-type allele. No DNA was available for the affected father who was diagnosed with prostate and liver cancer at age 70 years. Bioinformatics analysis indicates that this variant lies within the homeobox domain of the gene and is predicted to have a damaging effect on protein function. To further characterize this specific mutant and assess this possibility, structural and functional analyses were done.
NMR of wild-type and T164A mutant NKX3.1 homeodomains. The $^{15}$N-labeled wild-type NKX3.1 homeodomain was studied by three-dimensional $^{15}$N-separated HSQC NOESY spectra. The high degree of secondary and tertiary structural similarity with the NK2 homeodomain, already fully assigned and structurally determined, was clearly apparent from the NOESY spectrum; thus, the resonance assignment of the NKX3.1 homeodomain was straightforward. The NKX3.1 homeodomain has the same three helix bundle structure and the same hydrophobic core packing seen for all homeodomains whose structures have been determined (31).

To identify structural differences between wild-type and NKX3.1 (T164A) homeodomains, two-dimensional $^{15}$N HSQC and three-dimensional $^{15}$N-separated HSQC NOESY spectra of the T164A mutant were acquired. Again, resonance assignment was straightforward, with only those resonances of residues neighboring the T164A mutation (homeodomain residue 41) showing significant perturbations. This indicated that the overall structures of wild-type and mutant peptides were quite similar.

Two- and three-dimensional spectra were also acquired for the wild-type and mutant NKX3.1 homeodomains bound to the NK2 cognate DNA sequence. The spectrum of the DNA-bound wild-type homeodomain showed the typical chemical shift changes of the DNA recognition helix III and other residues caused by contact with the DNA. For example, the chemical shift change of the $H_1$ resonance of Trp$^{171}$ (homeodomain residue 48) as well as the appearance of the very downfield shifted Asn$^{177}$ (homeodomain residue 51) side chain amide resonances are diagnostic of the DNA-bound complex (see Fig. 3A). The spectrum of the free and DNA-bound T164A NKX3.1 mutant protein also showed similar chemical shift changes of DNA-contacting residues (see Fig. 3B). From these results, one can safely conclude that the T164A mutant has a similar overall structure and DNA contacts as the wild-type homeodomain and binds the cognate DNA sequence in a similar manner.

It is known that helix initiation and termination through specific backbone and side chain interactions can play a key role in a hierarchical protein folding mechanism (38–40). Several studies have shown that so-called N- and C-capping structural motifs can stabilize α-helices through hydrogen-bonding or hydrophobic interactions (41–43). Thr$^{164}$ at homeodomain position 41 is located in the N-cap position of the N-capping motif, which initiates the third helix of the homeodomain. By comparing the NKX3.1 homeodomain three-dimensional $^{15}$N-separated HSQC NOESY spectrum with that of NK2 (30, 31), we confirmed that this Thr$^{164}$ is in the correct position to form a hydrogen bond with the backbone amide of Gln$^{167}$ (homeodomain residue 44), the third residue of helix III (N$_3$ position; Fig. 4B). In addition, the Gln$^{167}$ side chain is positioned such that its carbonyl oxygen can hydrogen bond with the backbone amide of Thr$^{164}$. This $N_{amid}-C_{o}$ hydrogen-bonding conformation should help to initiate and stabilize the third helix of the homeodomain. In case of T164A mutant homeodomain protein, NMR signals are observed between the side chain methyl protons of Ala$^{164}$ (homeodomain residue 41) and amide backbone protons of Thr$^{164}$ and Gln$^{167}$ (homeodomain residues 43 and 44), indicating that the alanine methyl group occupies the same location as the wild-type threonine hydroxyl group. The alanine of the T164A mutant lacks the hydroxyl side chain group; thus, one of the N-capping hydrogen bonds cannot form. Therefore, this mutation can affect protein stability. Although the mutant is more unstable than the wild-type, comparison of the spectra shows that the second hydrogen bond, between the side chain of Gln$^{167}$ and Ala$^{164}$ backbone amide, still seems to be present, perhaps still providing some partial measure of stabilization to the structure (Fig. 4B, right).

Stability of wild-type and mutant NKX3.1 homeodomain peptides. To compare the stability of wild-type and mutant homeodomain, we measured circular dichroism spectra of both proteins. As an initial step, we measured far-UV circular dichroism spectra of wild-type and mutant protein. As expected in NMR experiment, both proteins showed similar patterns in circular dichroism spectra (Fig. 5A). From this result, we confirmed that the T164A mutation did not affect overall helicity of homeodomain protein and this result was consistent with results of NMR experiments. Thermal denaturation of wild-type and mutant proteins was probed by measuring the ellipticity at 222 nm at a temperature change rate of 1°C/min. This analysis indicated that the T164A mutation affected protein stability and folding (Fig. 5B). From these results, we conclude that this mutation can affect protein stability by disrupting the interaction between N-capping motif and α-helical backbone.

Binding of NKX3.1 to its cognate DNA recognition sequence. To analyze the effect of the T164A mutation on affinity of NKX3.1 for its cognate DNA sequence, we did electromobility shift assays (EMSA) using the cognate DNA-binding sequence $^{5'}$-TAGAGA-3' (44). We were unable to generate sufficient full-length NKX3.1(T164A) due to protein hydrophobicity, so we first tested the COOH-terminal truncated NKX3.1(1-184) that lacks the COOH-terminal inhibitory region and found that it is generally easier to isolate. The COOH-terminal truncated construct has about one-fifth the
DNA-binding affinity of the full-length protein. The T164A mutation reduced DNA-binding affinity by a factor of 20 as shown in the EMSA in Fig. 6A and quantitated in Fig. 6B. The effect seen in the EMSA was profoundly greater than implied by the NMR studies. Because the NH2-terminal region of NKX3.1 can influence the interaction of the homeodomain with the cognate DNA recognition sequence (34), we also compared wild-type and T164A mutant NKX3.1(114-184) peptides that contained homeodomain and 10 NH2-terminal amino acids. The NKX3.1(114-184) DNA binding was also compromised by the T164A mutation but only somewhat greater than 2-fold (Fig. 6C). As with the NMR studies, there was only a mild effect of the mutation on the interaction of the truncated peptide with a DNA recognition sequence.

Discussion

This is the first report in which germ-line variants of the NKX3.1 gene, a gene that is largely prostate specific and is commonly deleted in prostate tumors, were systematically evaluated in a large number of hereditary prostate cancer families. Multiple common and rare germ-line variants in the NKX3.1 gene were identified. These variants were linked to prostate cancer risk in 188 hereditary prostate cancer families. Importantly, a rare nonsynonymous change in the homeobox domain of the gene, T164A, cosegregated with prostate cancer in a family with three affected brothers and one unaffected brother. NMR solution structure analysis and circular dichroism studies showed this mutation to affect the stability of the homeodomain. The T164A mutation also decreased NKX3.1 binding to its cognate DNA recognition sequence. These results provide evidence that germ-line sequence variants in NKX3.1 may play a role in susceptibility to hereditary prostate cancer and underscore a role for NKX3.1 as a potential prostate cancer gatekeeper.

Inactivating mutations of NK family members cause autosomal dominant syndromes consistent with the notion that these genes may be haploinsufficient. For example, NKX2.5, the human cardiac-specific homologue of the Drosophila tinman gene, is subject to a variety of autosomal dominant point mutations that determine hereditary cardiac abnormalities (45). NKX2.1 is also known as thyroid transcription factor-1 and is a well-characterized determinant of lung and thyroid development (46, 47). NKX2.1 haploinsufficiency has been implicated in a heritable disorder of thyroid and pulmonary function (48). Here, we show that NKX3.1 acquired a mutation highly analogous to an inactivating mutation of NKX2.5 (see below). Moreover, NKX3.1 heterozygosity in gene-targeted mice results in prostate epithelial hyperplasia and dysplasia with somewhat longer latency than in the homozygous counterparts (19).

The finding that a rare nonsynonymous change (T164A) in the homeobox domain was found to cosegregate with disease in a prostate cancer family further underscores the role of NKX3.1 as the leading candidate for the prostate cancer gatekeeper. In addition to the effect of NKX3.1 loss in mouse models, the gatekeeper role for this gene is supported by the risk for prostate cancer conferred by the NKX3.1 (R52C) polymorphism that affects DNA binding (34). In addition, NKX3.1 protein expression is commonly diminished in

---

9 E. Ortner and E.P. Gelmann, unpublished observations.
10 N. Ahronovitz and E.P. Gelmann, unpublished observations.
primary prostate cancer and high-grade prostatic intraepithelial neoplasia to a degree that correlates with \( \text{NKX}3.1 \) deletion and promoter methylation (18).

The familial \( \text{NKX}3.1 \) mutation affects homeodomain position 41 changing a threonine to an alanine (T41A). Thr\(^{11} \) is well conserved among other NK-like homeodomain sequences (35) and is located at the NH\(_2\)-terminal cap position of the third helix in the \( \text{NKX}3.1 \) homeodomain. The NH\(_2\)-terminal helix cap is important for maintenance of the helical structure and its alteration may affect helix formation (41, 42). Although Thr\(^{11} \) is not involved in direct hydrogen bonding with the \( \text{NKX}3.1 \) DNA recognition sequence, it is likely to affect helix III structure and both DNA and protein interactions. The homologous Thr\(^{11} \) in the human \( \text{tinman} \) homologue \( \text{NKX}2.5 \) has been mutated to methionine in a family with hereditary atrial septal defect and atrioventricular block (49). The \( \text{NKX}2.5 \) (T178M), homeodomain T41M, mutation causes marked reduction in DNA-binding affinity of NKX2.5 but does not affect complex formation either with wild-type NKX2.5 or with GATA4. However, the T41M protein is unable to activate a reporter construct with the atrial natriuretic factor promoter. Atrial natriuretic factor is a transcriptional target for NKX2.5. T41A in \( \text{NKX}3.1 \) may not be as structurally disruptive mutation as T41M; therefore, the effect on \( \text{NKX}3.1 \) protein function may be more subtle than the effect of T41M on \( \text{NKX}2.5 \) function.

In conclusion, a resequencing analysis of the \( \text{NKX}3.1 \) gene in prostate cancer families has identified several germ-line variants that, as a group, show evidence of linkage to the disease. One specific mutation, T164A, in the homeodomain of \( \text{NKX}3.1 \), was observed to cosegregate with disease in a hereditary prostate cancer family, and structural and functional studies indicate that this change is a loss-of-function mutation. Although further studies will be necessary to fully understand the contribution that germ-line sequence variants in this gene make to prostate cancer susceptibility, these data provide additional evidence to implicate \( \text{NKX}3.1 \) as a critical gene in prostate carcinogenesis.

**Acknowledgments**

Received 6/20/2005; revised 8/19/2005; accepted 9/23/2005.

Grant support: USPHS Specialized Programs of Research Excellence grant CA58236, CA95052-01 (J. Xu), and ES09888 and CA96854 (E.P. Gelmann).

The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked advertisement in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

We thank all the study subjects who participated in this study.

**References**


Germ-Line Mutation of \textit{NKX3.1} Cosegregates with Hereditary Prostate Cancer and Alters the Homeodomain Structure and Function


\begin{tabular}{|l|l|}
\hline
\textbf{Updated version} & Access the most recent version of this article at: http://cancerres.aacrjournals.org/content/66/1/69 \\
\hline
\textbf{Supplementary Material} & Access the most recent supplemental material at: http://cancerres.aacrjournals.org/content/suppl/2006/01/17/66.1.69.DC1 \\
\hline
\end{tabular}

\begin{tabular}{|l|l|}
\hline
\textbf{Cited articles} & This article cites 45 articles, 9 of which you can access for free at: http://cancerres.aacrjournals.org/content/66/1/69.full#ref-list-1 \\
\hline
\textbf{Citing articles} & This article has been cited by 11 HighWire-hosted articles. Access the articles at: http://cancerres.aacrjournals.org/content/66/1/69.full#related-urls \\
\hline
\end{tabular}

\begin{tabular}{|l|l|}
\hline
\textbf{E-mail alerts} & Sign up to receive free email-alerts related to this article or journal. \\
\hline
\textbf{Reprints and Subscriptions} & To order reprints of this article or to subscribe to the journal, contact the AACR Publications Department at pubs@aacr.org. \\
\hline
\textbf{Permissions} & To request permission to re-use all or part of this article, use this link http://cancerres.aacrjournals.org/content/66/1/69. Click on "Request Permissions" which will take you to the Copyright Clearance Center's (CCC) Rightslink site. \\
\hline
\end{tabular}