ORIGINAL INVESTIGATION

Fine mapping of familial prostate cancer families narrows the interval for a susceptibility locus on chromosome 22q12.3 to 1.36 Mb

Bo Johanneson · Shannon K. McDonnell · Danielle M. Karyadi · Scott J. Hebbring ·
Liang Wang · Kerry Deutsch · Laura McIntosh · Erika M. Kwon · Miia Suuriniemi ·
Janet L. Stanford · Daniel J. Schaid · Elaine A. Ostrander · Stephen N. Thibodeau

Abstract Genetic studies suggest that hereditary prostate cancer is a genetically heterogeneous disease with multiple contributing loci. Studies of high-risk prostate cancer families selected for aggressive disease, analysis of large multi-generational families, and a meta-analysis from the International Consortium for Prostate Cancer Genetics (ICPCG), all highlight chromosome 22q12.3 as a susceptibility locus with strong statistical significance. Recently, two publications have narrowed the 22q12.3 locus to a 2.18 Mb interval using 54 high-risk families from the ICPCG collaboration, as defined by three recombination events on either side of the locus. In this paper, we present the results from fine mapping studies at 22q12.3 using both haplotype and recombination data from 42 high-risk families contributed from the Mayo Clinic and the Prostate Cancer Genetic Research Study (PROGRESS) mapping studies. No clear consensus interval is present when all families are used. However, in the subset of 14 families with ≥5 affected men per family, a 2.53-Mb shared consensus segment that overlaps with the previously published interval is identified. Combining these results with data from the earlier ICPCG study reduces the three-recombination interval at 22q12.3 to approximately 1.36 Mb.

B. Johanneson · D. M. Karyadi · E. M. Kwon · M. Suuriniemi ·
E. A. Ostrander
Cancer Genetics Branch,
National Human Genome Research Institute,
National Institutes of Health, Bethesda, MD, USA

S. K. McDonnell · D. J. Schaid
Department of Health Sciences Research,
Mayo Clinic, Rochester, MN, USA

S. J. Hebbring · L. Wang · S. N. Thibodeau (✉)
Department of Laboratory Medicine and Pathology,
Mayo Clinic, 200 First Street SW, 920 Hilton Building,
Rochester, MN 55905, USA
e-mail: sthibodeau@mayo.edu

K. Deutsch
Institute for Systems Biology, Seattle, WA, USA

L. McIntosh · J. L. Stanford
Division of Public Health Sciences,
Fred Hutchinson Cancer Research Center, Seattle, WA, USA

J. L. Stanford
Department of Epidemiology,
School of Public Health and Community Medicine,
University of Washington, Seattle, WA, USA

Introduction

Prostate cancer (PC) is the second most common cancer and the second leading cause of cancer deaths for men in the United States, with 218,890 new cases and 27,050 cancer deaths expected during 2007 (Jemal et al. 2007). There are differences in both the incidence and mortality rates of PC across populations. From 1999 to 2003, the highest PC rates within the US occurred in African Americans, followed by men of European ancestry, and then Asian–American men (Jemal et al. 2007). In addition, a family history of disease is one of the strongest known risk factors for PC. An unaffected man with one or more affected first-degree relatives or an affected brother diagnosed before age 60 has a 2- to 4-fold increased risk of developing this malignancy (Goldgar et al. 1994; Keetch et al. 1995). Furthermore, 5–10% of all PC cases and up to 40% of those ≤55 years of age may have a hereditary basis (Bratt 2002; Carter et al.

1992; Zhang et al. 2002). Segregation analyses support an autosomal dominant, multifactorial mode of inheritance as well as recessive or X-linked inheritance. Although the studies are in diverse populations, seven analyses report evidence for a single major gene with a rare autosomal dominant effect, especially in families with early ages of onset (Carter et al. 1992; Conlon et al. 2003; Cui et al. 2001; Gronberg et al. 1997; Schaid et al. 1998; Valeri et al. 2003; Verhage et al. 2001). Two studies provide evidence for recessive inheritance (Cui et al. 2001; Pakkanen et al. 2007), while Gong et al. (2002) suggest that a multifactorial model best fits the available data.

Building on this evidence, large collections of hereditary prostate cancer (HPC) families have been analyzed using genome-wide linkage scans in order to identify molecular factors or pathways involved in disease susceptibility. While highlighting many loci, the results from these studies have not been well replicated in independent efforts (reviewed in: (Easton et al. 2003; Ostrander et al. 2004; Ostrander and Stanford 2000; Schaid 2004)). There are several possible explanations for this lack of replication. Since PC is a common disorder (one in six men will be diagnosed with this disease over their lifetime), collections of high-risk families are likely to include a moderate number of phenocopies, which can confound analyses. Also, HPC demonstrates significant locus heterogeneity. This is supported by a recent segregation analysis, which suggests a multifactorial model with multiple lower penetrance genes (Gong et al. 2002).

One approach to unambiguously identifying PC genes is to focus on regions, such as 22q12.3, where multiple independent studies have identified "suggestive" evidence for linkage. Putative linkage to chromosome 22 was first reported by three of the eight genome-wide scans published simultaneously in *The Prostate*, (Cunningham et al. 2003; Janer et al. 2003; Lange et al. 2003). In these studies, the best evidence for linkage was found in the subset of families with ≥4 affected (University of Michigan, LOD = 1.87), ≥5 affected (PROGRESS, HLOD = 2.21), or with a median age at diagnosis ≥66 years (Mayo Clinic, LOD = 1.59). Subsequently, Camp et al. reported a TLOD (theta LOD; a robust multipoint linkage statistic) of 2.42 at 22q12.3 using data from 436 three-generation pedigrees (Camp et al. 2005). Further evaluation of previously reported HPC families provides suggestive linkage for 22q. Specifically, by modeling two-locus gene-gene interactions, Chang et al. (2006) detected suggestive evidence for an epistatic interaction between 22q13 and 21q22 in 426 families from Johns Hopkins University, University of Michigan, University of Umeå, and University of Tampere. By analyzing HPC families according to disease aggressiveness, two studies reported evidence for linkage to chromosome 22q with a dominant HLOD >2 (Chang et al. 2006; Stanford et al. 2006). Finally, the most convincing

evidence comes from an analysis of 1,233 HPC families by the International Consortium for Prostate Cancer Genetics (ICPCG). Using data from 11 independent but collaborating research centers, Xu et al. reported that the only significant evidence for linkage was on chromosome 22q12.3 (dominant HLOD = 3.57) in 269 pedigrees with ≥5 affected per family (Xu et al. 2005).

The chromosome 22q region has been further analyzed using two datasets, one from the University of Utah and the other combined data from the ICPCG (excluding the eight previously reported Utah pedigrees). Fourteen University of Utah HPC families with evidence for linkage to 22q12.3 were selected either by having a pedigree LOD score ≥0.58 (corresponding to a nominal *P*-value ≤ 0.05) or at least 5 affected men that shared the same haplotype (Camp et al. 2006). After genotyping additional fine-mapping markers, a minimal recombination interval of 881-Kb was identified. However, using the more rigorous standard of three recombination events at each end, a 3.20-Mb interval was defined (Camp et al. 2006). The ICPCG analysis was able to further refine the map of the 22q12 region using 40 of the initial 1,233 families (Camp et al. 2007). The critical families all had ≥4 affected men and individual pedigree LOD scores ≥0.58 inside the defined susceptibility region (Camp et al. 2007). A 3.78-Mb consensus interval, defined by three recombination events on each side, was delineated. Remarkably, none of the ICPCG pedigrees was in conflict with the consensus interval. When combining the 40 ICPCG and 14 University of Utah pedigrees, the initial 12 cM ICPCG LOD-1 support interval was reduced to a minimal 882-Kb consensus interval, or a 2.18-Mb interval as defined by three recombination events at both ends of the region (Camp et al. 2007).

While highly suggestive, the ICPCG combined analysis was based on sparsely placed genome-wide scan markers (6–8 markers). In this new analysis, we have further refined the region of interest at 22q12.3 by using a high-density set of fine mapping markers, providing maximal pedigree informativeness on a set of 42 families from the Mayo Clinic and Prostate Cancer Genetic Research Study (PROGRESS) mapping studies.

Results

In this study, we have refined the location of a putative prostate cancer locus on chromosome 22 using significantly more markers and family information. Previous analyses of both the Mayo Clinic and PROGRESS HPC families suggested evidence for a susceptibility locus at 22q (Cunningham et al. 2003; Janer et al. 2003; Stanford et al. 2006). The current Mayo Clinic study is comprised of 173 HPC families with 482 affected men and 18 relatives with genotypes. The PROGRESS study encompasses 254 HPC families with

858 sampled affected men and 498 relatives. On average, there are 2.8 (2–7) and 3.5 (2–12) affected men genotyped per family from the Mayo Clinic and PROGRESS studies, respectively.

In the initial PROGRESS genome-wide scan, six microsatellite markers with an average spacing of 8.45 cM were genotyped for chromosome 22 in 254 HPC families (Janer et al. 2003). Thirty-two families (12.6%) with evidence for linkage (LOD >0.5) were selected for additional genotyping. Each family was analyzed with a core panel of 24 microsatellites with an average spacing of 2.01 cM. Additional markers (microsatellites or SNPs) were genotyped to refine family-specific recombinants. The final set of 52 markers genotyped on the PROGRESS families includes the initial six microsatellites used in the genome-wide scan, and an additional 32 fine-mapping microsatellites and 14 SNPs.

The majority of the Mayo Clinic HPC families (167 of 173) were previously genotyped with 40 SNPs on chromosome 22 from the Early Access Affymetrix Mapping 10 K SNP array (Schaid et al. 2004). Additional association-based studies of chromosome 22, in the region of 25.7–37.4 Mb, were performed with a custom set of 738 tagSNPs with Illumina's GoldenGate assay on all families (173 families with a total of 482 sampled affected men). For the fine mapping linkage studies, a final set of 174 independently informative SNPs was selected from the merged set of 670 high-quality Illumina tagSNPs and 40 Early Access Affymetrix SNPs to remove LD among the SNP's ($r^2 < 0.10$).
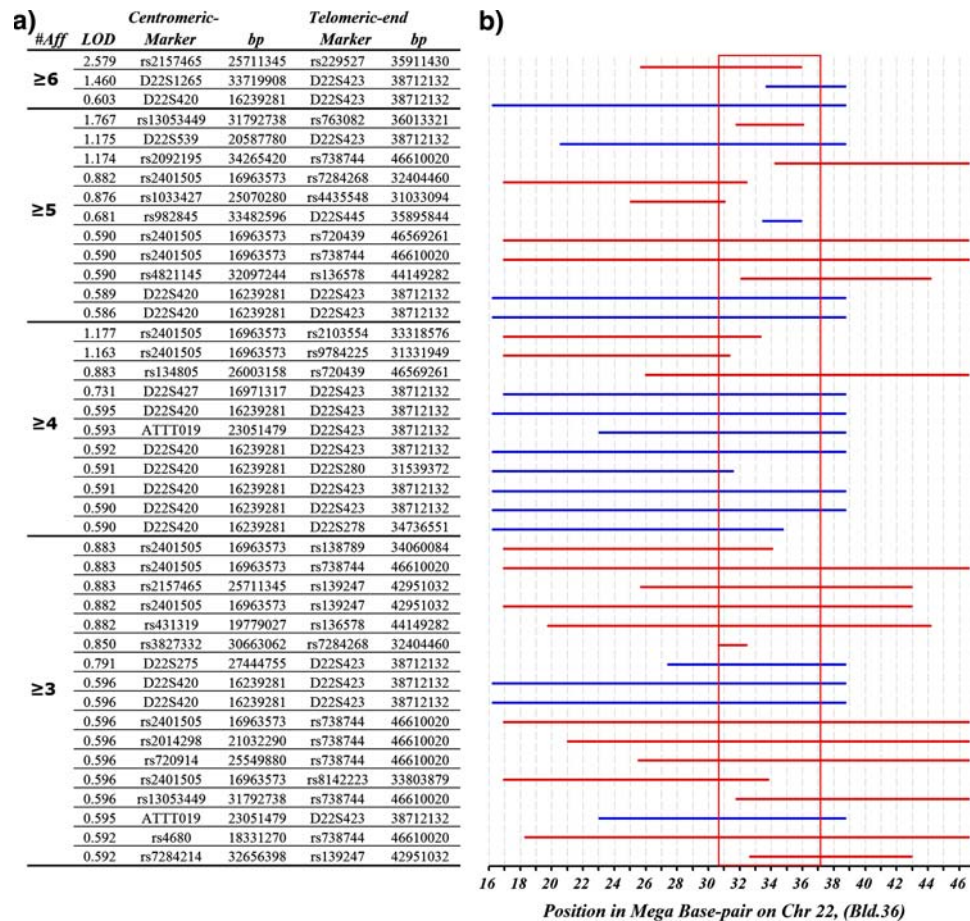
Recombination analysis was carried out only on the subset of Mayo and PROGRESS families that provided evidence for linkage inside the LOD-1 support interval identified in the combined ICPCG study (Xu et al. 2005). To determine the approximate physical position for this interval, known to be located roughly between 35 and 47 cM, a linear interpolation between microsatellite markers with known genetic and physical position on both sides of the borders was used. The physical boundaries defining the ICPCG LOD-1 support interval extend from 30.62 Mb (35 cM) to 37.20 Mb (47 cM). Any family with a pedigree-LOD score ≥ 0.58 inside this interval was selected for additional analysis. Although some families may have a LOD score ≥ 0.58 in this region simply by chance, the selection process is conservative and was intended to increase the proportion of truly linked families to this region. A parametric multi-point linkage analysis was performed with a dominant "affected only" model using a disease allele frequency of 0.003 and 100% penetrance. This analysis, therefore, highlighted only families where all affected men shared the same haplotype.

A total of 42 families, 24 from Mayo Clinic and 18 from PROGRESS, achieved a pedigree-LOD score ≥ 0.58 inside the ICPCG LOD-1 support interval. Haplotypes for each family were then constructed using MERLIN software

(Abecasis et al. 2002; Cook Jr 2002). Within each family, any haplotype that was shared among all affected men was identified. The recombinant boundaries for the shared-affected haplotypes in each linked pedigree are given in Fig. 1a and are illustrated in Fig. 1b. The ICPCG LOD-1 support interval is indicated by a red square in Figs. 1b, 2a and b. When all shared-affected haplotypes from each family are aligned, a tally of the number of families that have a shared-affected haplotype at any position across the region is indicated (Fig. 2a). The data are presented from Mayo Clinic and PROGRESS families combined, as well as each group independently. Also included are the data reported in the recent chromosome 22 analyses from Camp et al. (2006) and the ICPCG study (Camp et al. 2007).

No single interval was identified in the recombination mapping that is shared by all 42 families, either as a combined Mayo Clinic and PROGRESS analysis or in each data set separately. In the combined analysis, there are three peaks with 35 families (83%) contributing to a consensus interval (Figs. 2a, 3). If the three intervals were to be defined by only one recombination event at each end, the first is a 630-Kb region (30,663,062–31,033,094) overlapping four known genes [YWHAH (OMIM: 113508), SLC5A1 (OMIM: 182380), RFPL2 (OMIM: 605969) and SLC5A4, see Fig. 3]. The second is a 307-Kb region (32,097,224–32,404,460 bp) positioned over the *LARGE* gene [OMIM: 603590]. The third interval is only 84-Kb (33,719,908–33,803,879 bp) long, partly overlapping an intestine-specific homeobox gene, ISX (RAXLX). If the intervals are defined more rigorously by three recombinant events, the first interval increases to a 5.54-Mb centromeric interval (26,003,158–31,539,372 bp). The other two intervals overlap and form a 2.94-Mb telomeric region (31,792,738–34,736,551 bp). This interval corresponds to the three-recombination interval reported in the ICPCG analysis (Camp et al. 2007).

When analyzing the Mayo Clinic and PROGRESS families separately, a shared consensus interval was not observed (Fig. 2a). The two most centromeric one-recombination consensus intervals in the Mayo Clinic data (red line) show 19 (79%) and 20 (83%) families overlapping with the two centromeric intervals from the combined analysis of 42 families described above, indicating that Mayo Clinic families contribute significantly to the centromeric boundaries. In the PROGRESS pedigrees (blue line), a 1.02-Mb one-recombination consensus interval (33,719,908–34,736,551 bp) with a 11.27-Mb three-recombination interval (27,444,755–38,712,132 bp) is detected in all but one family (17 out of 18; 94%). This region corresponds to the three peaks from the combined analysis, and overlaps precisely with the interval identified by the combined ICPCG analysis (Camp et al. 2007) and the University of Utah study (Camp et al. 2006).

a)

#Aff	LOD	Centromeric-Marker	bp	Telomeric-end Marker	bp
≥6	2.579	rs2157465	25711345	rs229527	35911430
	1.460	D22S1265	33719908	D22S423	38712132
	0.603	D22S420	16239281	D22S423	38712132
≥5	1.767	rs13053449	31792738	rs763082	36013321
	1.175	D22S539	20587780	D22S423	38712132
	1.174	rs2092195	34265420	rs738744	46610020
	0.882	rs2401505	16963573	rs7284268	32404460
	0.876	rs1033427	25070280	rs4435548	31033094
	0.681	rs982845	33482596	D22S445	35895844
	0.590	rs2401505	16963573	rs720439	46569261
	0.590	rs2401505	16963573	rs738744	46610020
	0.590	rs4821145	32097244	rs136578	44149282
	0.589	D22S420	16239281	D22S423	38712132
	0.586	D22S420	16239281	D22S423	38712132
≥4	1.177	rs2401505	16963573	rs2103554	33318576
	1.163	rs2401505	16963573	rs9784225	31331949
	0.883	rs134805	26003158	rs720439	46569261
	0.731	D22S427	16971317	D22S423	38712132
	0.595	D22S420	16239281	D22S423	38712132
	0.593	ATTT019	23051479	D22S423	38712132
	0.592	D22S420	16239281	D22S423	38712132
	0.591	D22S420	16239281	D22S280	31539372
	0.591	D22S420	16239281	D22S423	38712132
	0.590	D22S420	16239281	D22S423	38712132
	0.590	D22S420	16239281	D22S278	34736551
≥3	0.883	rs2401505	16963573	rs138789	34060084
	0.883	rs2401505	16963573	rs738744	46610020
	0.883	rs2157465	25711345	rs139247	42951032
	0.882	rs2401505	16963573	rs139247	42951032
	0.882	rs431319	19779027	rs136578	44149282
	0.850	rs3827332	30663062	rs7284268	32404460
	0.791	D22S275	27444755	D22S423	38712132
	0.596	D22S420	16239281	D22S423	38712132
	0.596	D22S420	16239281	D22S423	38712132
	0.596	rs2401505	16963573	rs738744	46610020
	0.596	rs2014298	21032290	rs738744	46610020
	0.596	rs720914	25549880	rs738744	46610020
	0.596	rs2401505	16963573	rs8142223	33803879
	0.596	rs13053449	31792738	rs738744	46610020
	0.595	ATTT019	23051479	D22S423	38712132
	0.592	rs4680	18331270	rs738744	46610020
	0.592	rs7284214	32656398	rs139247	42951032

b)

*Position in Mega Base-pair on Chr 22, (Bld.36)*

Fig. 1 Position of shared haplotypes at 22q12.3 among all affected men for each pedigree. **a** Families were selected for this study if they had a shared haplotype among all affected men in the family and demonstrated a family-LOD score ≥0.58. Families are sorted by number of affected men and LOD score. If two haplotypes segregated in a family, they were merged into one overlapping haplotype. Physical positions are taken from the March 2006 human reference sequence (NCBI Build 36.1). Number of affected men reported in each family is indicated as is the maximum pedigree LOD score inside the LOD-1 support interval. The outermost centromeric and telomeric marker defining a position of a recombination between two markers is shown. **b** Graphic illustration of shared haplotypes described in part a. *Red bars* indicate pedigrees from the Mayo Clinic, *blue bars* indicate PROGRESS pedigrees. Physical position of the LOD-1 support interval defined in the ICPCG study (Xu et al. 2005), between 35–47 cM (deCODE genetic map) is outlined by the *red box*

The data were then stratified by the number of affected men per family (typed and untyped cases) to increase the proportion of families likely to have an inherited predisposition to PC. Depending on the number of cases reported in a family, the appearance of the shared consensus interval changes (Fig. 2b). Noticeably, a more narrow-shared consensus region on the centromeric side of the ICPCG LOD-1 susceptibility region appeared when families with ≥5 affected men are included. Although only four families had six or more members with PC, a relatively large proportion of the families (14 out of 42, 33%) have ≥5 affected men. Of these 14 families, 12 (86%) share a 2.53-Mb consensus interval (33,482,596–36,013,321 bp) defined by three recombination events on both ends. This interval overlaps with the 22q12.3 interval identified in the previous University of Utah and ICPCG studies (Camp et al. 2006, 2007).

Discussion

In this report, we present a combined 22q12.3 fine-mapping analysis from two independent groups. The 24 Mayo Clinic and 18 PROGRESS families were drawn from initial data sets of 173 and 254 HPC families for which genome-wide scans had been previously undertaken. Families were eligible for this analysis if the family LOD score was ≥0.58 and all affected men had a shared haplotype somewhere within the 22q12.3 ICPCG susceptibility LOD-1 support interval. The ICPCG LOD-1 region was chosen as a reference region because it represents the most likely linked-region for chromosome 22 based on a large collection of PC families. Forty-two families were used to reduce the consensus interval where the susceptibility locus on 22q is likely to be located.
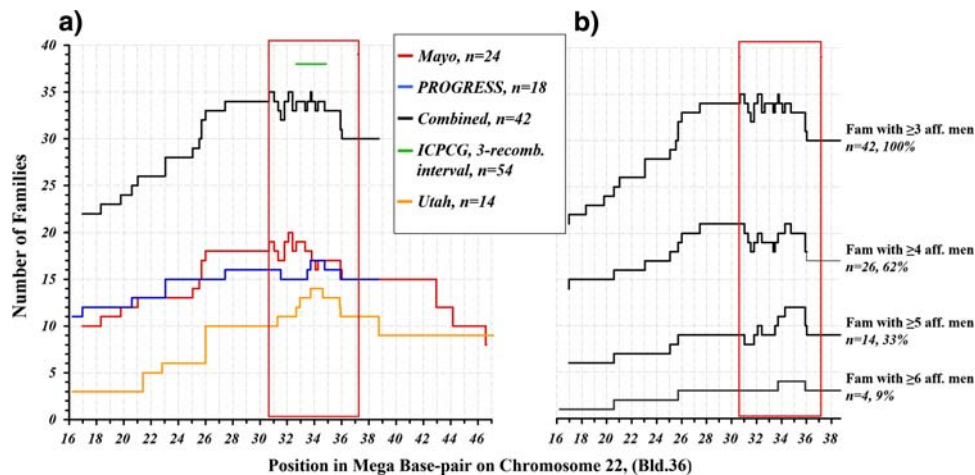
Fig. 2 Position and number of pedigrees sharing a consensus interval at 22q12.3. **a** Number of families sharing a consensus interval at each physical position along the 22q12.3 susceptibility region is shown. Mayo Clinic families are shown in *red*, PROGRESS in *blue*, and the combined in *black*. Data from previous studies based on families from the University of Utah (Camp et al. 2006), and the ICPCG (Camp et al. 2007) are included and labeled in *yellow* and *green*, respectively. Physical positions are taken from the March 2006 human reference sequence (NCBI Build 36.1). Physical position of the LOD-1 support interval defined in the ICPCG study (Xu et al. 2005), between 35–47 cM (deCODE genetic map) is outlined by the *red box*. The number of families indicating the ICPCG data is not to scale. **b** Consensus interval at each physical position along the 22q12.3 susceptibility region in the combined dataset of Mayo and PROGRESS pedigrees is shown. Each line represents the combined dataset after stratification for an increasing number of affected men reported in each family

In the combined set of 42 families, no single consensus region was identified. This is not surprising given the likely occurrence of phenocopies within the HPC families. Initially we identified two separate intervals using data from 35 families. (The recombination interval size was 5.54 and 2.94 Mb for the centromeric and telomeric intervals, respectively). Disregarding the small gap between the two intervals, an 8.74-Mb overlapping consensus region can be defined between 26.00 and 34.74 Mb. This region is bounded on either end by previously identified recombination hot spots (Tapper et al. 2001). Within this interval, there are over 100 coding genes. Of note, 35 of these genes are implicated in different cancers according to the Atlas of Genetics and Cytogenetics in Oncology and Haematology (http://AtlasGeneticsOncology.org) (Dorkeld et al. 1999), including several well-known oncogenes and tumor suppressor genes, *MN1*, *CHEK2*, *EWSR1*, *NF2* and *MYH9*. Among them, the most interesting is *CHEK2*, which has been found to increase the risk for sporadic and familial PC in studies of diverse populations from Finland, Poland and the United States (Cybulski et al. 2006; Dong et al. 2003; Seppala et al. 2003).

The study presented here represents the most detailed analysis of this region to date. It includes both additional families and a denser set of markers than was previously published in either the Mayo Clinic or PROGRESS studies alone, or in the data contributed to the recently published ICPCG analysis (Camp et al. 2007). Indeed, only nine of the 42 families presented here are in the ICPCG report (five from Mayo Clinic, and four from PROGRESS). Of these 9

families, only 1 contributed information that was important to defining the ICPCG consensus region. This one case marked the second recombination on the telomeric side. The ICPCG report used only families with ≥4 affected men, while this study includes 17 families with 3 affected men. Additionally, the ICPCG study allowed only cases with medical record or death certificate verification to be counted as affected, thus removing a number of men whose prostate cancer was self-reported or was confirmed by reports from multiple affected first-degree family members. In the PROGRESS study, 20% (231 of 1,143) of all affected men are coded as unknown in the ICPCG analysis. However, previous studies have shown that self-reporting of PC is highly accurate. Within the PROGRESS study, a self-reported PC diagnosis was confirmed in 800 of the 801 medical records received (99.9%). Therefore, all self-reported PROGRESS PC cases are included in this study. Furthermore, as the affection status in the PROGRESS pedigrees is constantly updated, 47 men have been diagnosed with PC since the ICPCG analysis was performed in 2003. Finally, the nine Mayo Clinic and PROGRESS families included in the ICPCG study were at that time only analyzed with 6 or 7 genome-wide microsatellite markers. In this study, additional fine-mapping markers have been genotyped in each family, and as such, greater power to discriminate between haplotypes exists for each family. As a result, one PROGRESS family included in the original ICPCG study was not included in this study after genotyping of additional markers dropped the family LOD score to <0.58. Not surprisingly, when the nine families overlapping

Fig. 3 Known candidate genes located in shared consensus interval at 22q12.3. Positions of known genes (*upper panel*) are taken from "RefSeq Genes" at the UCSC genome-browser, the March 2006 human reference sequence (NCBI Build 36.1). The lines (*lower panel*) represent the position and number of pedigrees sharing a consensus interval on the physical map. The *red line* represents all 42 combined families from the Mayo Clinic and PROGRESS. The *blue line* represents the 14 pedigrees that have ≥5 affected men in the same dataset. Data from previous studies based on 14 families from the University of Utah (*yellow line*) (Camp et al. 2006), and 54 families from ICPCG (*green line*) (Camp et al. 2007) are included (the *green bar* represent the consensus three-recombination interval). The *black line* represents the 25 families from the Mayo Clinic, PROGRESS and the University of Utah that have been fine mapped and have ≥5 affected men, combined with the 29 remaining families from the ICPCG study
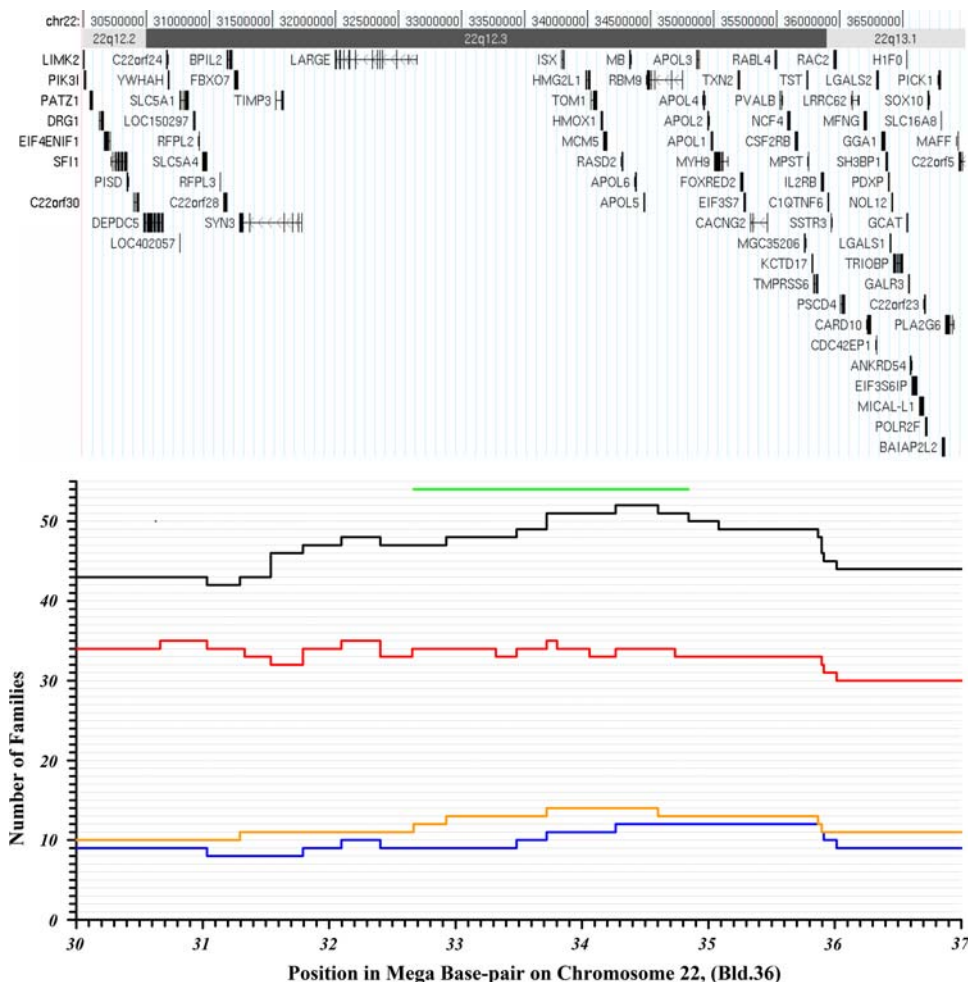
both this and the ICPCG study are compared, the addition of new markers shifts the location of several recombination breakpoints. Specifically, in five of the nine families, the centromeric and/or the telomeric recombination sites are 0.03–35 cM shorter (0.03, 0.5, 2.47, 6.60, 7.25, 35 cM).

There are several possible explanations for the difficulty in identifying a common consensus shared region in all pedigrees. First, with such a high prevalence of disease and a strong environmental component to PC etiology, phenocopies (non-gene carriers with disease) are likely to be common (Jemal et al. 2007). The presence of a phenocopy within a family could provide erroneous recombinational boundaries. Second, as prostate cancer is a later-onset disease, individuals in the parental generation of the probands are usually deceased and individuals in the offspring generations are usually too young to manifest the phenotype, thus making many pedigrees only marginally informative for linkage analysis. Many of the families that have a positive, but low LOD score (e.g., 0.58), may represent a false-positive signal for linkage to this chromosome. These families may also provide erroneous recombination boundaries. Finally, although less likely, chromosome 22q may harbor

two or more susceptibility loci for PC. One way to reduce the impact for the first two of these problems, however, is to stratify HPC families based on the number of affected men per family. Interestingly, the shared consensus "plateau" on the centromeric side of the ICPCG LOD-1 susceptibility region gradually disappears when families with increasing numbers of affected men are analyzed. Our data suggest that the shared interval on the centromeric side is mainly a contribution from families with fewer affected men. In contrast, in the 14 families having ≥5 affected men, a 2.53-Mb consensus interval defined by three recombination events is identified, and this interval overlaps with the region previously identified in the ICPCG and University of Utah studies. Of note, the families used in the University of Utah study are large, composed mainly of extended families with ≥5 affected men. Additionally, the significant 22q12.3 LOD score in the ICPCG study was identified in the subset of families having ≥5 affected men. Unfortunately, because of software limitations that do not adequately handle LD at high marker density, we were not able to look for the presence of a (founder) haplotype across all of our families.

We observed intriguing results when combining fine-mapping data sets from the Mayo Clinic, PROGRESS, and the University of Utah. For the analysis of families with ≥5 affected men, a total of 25 families are identified: eight Mayo Clinic, six PROGRESS, and eleven from the University of Utah. The Utah pedigrees contribute one more critical recombination breakpoint on the centromeric and three on the telomeric side (Camp et al. 2006), which is shared by 23 of the families (92%). Thus, the ≥5 affected minimal consensus interval defined by one recombination event on each side is reduced to only 336 Kb (34,265,420–34,601,446 bp), and the three recombination interval is reduced to 2.18 Mb (33,719,908–35,895,844 bp).

The ≥5 affected shared consensus interval discussed above can be further refined by incorporating data from the remaining research groups participating in the ICPCG study (Camp et al. 2007). Even though the remaining ICPCG families have only a few markers genotyped, and possibly include families with only four affected men, we added data from them to the 25 Mayo Clinic, PROGRESS and University of Utah families that have fine-mapping data and ≥5 affected men per family. After removing Mayo Clinic, PROGRESS and University of Utah families from the ICPCG dataset, a total of 29 new families are available for analysis. We found that 52 of the 54 families (96%) share a consensus interval, which overlaps with the ≥5 affected shared consensus interval described above. In the 29 ICPCG families, no extra recombination breakpoints are detected that would reduce the minimal consensus interval (Camp et al. 2007). However, two recombination breakpoints are found in the ICPCG pedigrees that shrink the three-recombination interval on the telomeric side to a 1.36-Mb interval between 33.72 Mb and 35.08 Mb. In the ICPCG study, the previous limit on the telomeric side of the three-recombination interval was defined by a PROGRESS family with only four affected men that is not included in this combined analysis. As such, the edge of the interval shifts to 35.08 Mb instead of 34.84 Mb. This increases the interval by 240 Kb and includes five new genes (APOL1, APOL2, APOL3, APOL4 and MYH9) to the list of 11 previously noted genes (ISX, HMG2L1, TOM1, HMOX1, MCM5, RASD2, MB, LOC284912, APOL6, APOL5 and RBM9). The overall region is nevertheless reduced by 38% (2.2 Mb vs. 1.36 Mb) when compared to that reported by the ICPCG study (Camp et al. 2007).

Although a number of studies (Camp et al. 2005; Chang et al. 2005, 2006; Cunningham et al. 2003; Janer et al. 2003; Lange et al. 2003; Stanford et al. 2006; Xu et al. 2005) provide convincing evidence for a PC genetic susceptibility locus on chromosome 22 at q12.3, a candidate susceptibility gene has not yet been identified. This current study represents the most detailed fine-mapping analysis of

this region reported to date. The most conservative region is defined by an 8.74-Mb overlapping consensus region between 26.00 and 34.74 Mb, which encompasses all previously published consensus regions. Based on these data, along with other published results, a minimal three-recombinant consensus interval of approximately 1.36 Mb between 33.72 and 35.08 Mb is suggested. Given the inherent difficulty in performing linkage studies for PC (high phenocopy rate, genetic, allelic and phenotypic heterogeneity, incomplete and age-dependent penetrance, etc.), further reduction of the minimal consensus intervals may prove to be difficult. The results presented here, then, are likely to provide the most comprehensive framework achievable for candidate gene testing. While the indicated interval may be overly generous due to the conservative nature of our analysis, it is unlikely that the chromosome 22 susceptibility gene lies outside of it. Ongoing studies, then, are aimed at evaluating genes in this region for variants associated with prostate cancer risk.

Materials and methods

Selection of families

For the Mayo Clinic families, details of the survey, telephone follow-up, and family recruitment can be found elsewhere (Berry et al. 2000; Cunningham et al. 2003; Schaid et al. 1998). Briefly, HPC families were selected through a proband, who received treatment for PC at the Mayo Clinic, with the requirement of at least three men with prostate cancer in the family of whom two or more were still alive for recruitment. The current study included 173 families with a total of 482 sampled affected men. Study materials and protocols were approved by the Mayo Clinic Human Subjects Internal Review Board.

The 254 PROGRESS HPC families have been ascertained by the Fred Hutchinson Cancer Research Center from throughout North America by advertising and public media (Janer et al. 2003; McIndoe et al. 1997). The families had to fulfill one of following criteria in order to participate: (1) have three or more first-degree relatives with PC; (2) have three generations (maternal or paternal) with PC; or (3) have two first-degree relatives with PC diagnosed before age 65 or who were African–American. All prostate cancer survivors and selected unaffected men and women were invited to join PROGRESS. On average, eight members of each family completed a study questionnaire on medical and family cancer history and provided a blood sample. The affected men were also asked to sign a consent form for release of medical records related to the prostate cancer diagnosis and treatment. Collection details and family characteristics have been described previously in papers

summarizing our initial genome-wide scans (Janer et al. 2003; McIndoe et al. 1997; Stanford et al. 2006). Since then, the family data have been reviewed and medical and family cancer history updated. In the 254 families utilized for this analysis, 47 new cases have developed PC, including nine individuals previously coded as unknown affection status. Currently, there are 858 sampled affected men in the 254 families. Study forms and protocols were approved by the Institutional Review Boards of the Fred Hutchinson Cancer Research Center and the National Human Genome Research Institute.

SNP selection for linkage analysis and haplotyping for Mayo Clinic families

Chromosome 22 at 22q12 had previously been mapped in the Mayo pedigrees using 40 single nucleotide polymorphisms (SNPs) available from the Early Access Affymetrix Mapping 10 K SNP array (Schaid et al. 2004). A maximum dominant HLOD of 1.97 was obtained on chromosome 22q with the LOD-1 support interval surrounding this peak ranging from 25.7 to 37.4 Mb. To refine this initial linkage signal and perform association studies specific to those genes identified within this region, additional SNPs were genotyped by The Center for Inherited Disease Research (CIDR) using the Illumina GoldenGate platform. Utilizing a variety of public databases and bioinformatics tools, a physical and transcript map was constructed encompassing the LOD-1 support interval described above (25.7–37.4 Mb). Chromosomal start and stop positions were obtained for all genes ($n = 216$) within this region of interest, 10 Kb was added to the 5′ and 1 Kb to the 3′ end of each gene, and then adjacent or overlapping genes were combined into a single segment, or "contig" (continuous regions of overlapping genes).

SNP selection within this region relied on tagSNPs selected on the basis of linkage disequilibrium (LD). A list of all candidate SNPs covering this region, including measures of assay fitness (such as design scores, error codes and degree of duplication in the genome), was obtained from Illumina. Publicly available genotype data from the HapMap Consortium (version 2, October 2005) and Perlegen Sciences were used to calculate pairwise LD for each pair of SNPs and the algorithm implemented in ldSelect (Carlson et al. 2004) was used to select tagSNPs. A TagSNP in a given bin was defined as a SNP that exceeded an r^2 threshold ($r^2 \geq 0.80$) with all other SNPs in the same bin. Additional selection criteria were then employed to select a single SNP within the set of tagSNPs to represent each bin. To choose between multiple tagSNPs within an LD bin, hierarchical selection criteria were implemented and relied on the design metrics provided by Illumina, minor allele frequency (MAF, ≥ 0.05), and type of SNP (c-snp vs. non-coding).

Only contigs/gene regions with at least 70% of the sequence covered by LD bins were selected for further analysis, resulting in the exclusion of 33 from the original 216 genes. The average LD bin coverage for the remaining contigs was 87%. Utilizing this process, 738 tagSNPs were selected to cover 183 genes within our region of interest on chromosome 22q. Of the 738 selected SNPs, genotype data were obtained for 681; 670 were of sufficient quality for further analysis. Eleven SNPs were excluded for the following reasons: (1) Minor allele frequency (MAF) <0.01 ($n = 7$); (2) marker call rate <90% ($n = 7$); or (3) P value <0.001 for the test for Hardy–Weinberg equilibrium ($n = 1$). Note that some excluded SNPs fell into more than one category. From the combined set of the 670 Illumina tagSNPs and the 40 Early Access Affymetrix Mapping 10 K SNPs, a reduced set of 174 independent markers was selected for the linkage analysis. To minimize the effect of LD on the linkage analysis, SNPs were chosen by selecting the most informative SNP from each LD bin using an $r^2 < 0.10$.

Microsatellite and SNP selection for linkage analysis and haplotyping of PROGRESS families

Previously, six markers (D22S420, ATTT019, D22S689, D22S685, D22S683 and D22S445) on chromosome 22 that were selected from Human Screening Sets 6 and 8 (Research Genetics) had been genotyped in all 254 PROGRESS families (Janer et al. 2003). An additional set of 34 microsatellites was then typed in the subset of PROGRESS families selected for this study. The name, primer-sequence and physical position for each marker are found in Supplemental Table S-1. Markers with known genetic position in the revised deCODE map (Nievergelt et al. 2004) were selected to cover the region with at least one marker per cM. More markers were subsequently added during the process of haplotyping in order to delineate recombination breakpoints in different families. In a region between D22S281 and D22S277 (32.6–34.6 Mb), some families had unclear phase for critical recombination breakpoints, even after all known microsatellites had been exploited. Therefore, using HaploView software and the data from the HapMap Caucasian population (IHC 2003), an additional 14 SNPs not in LD with each other ($r^2 \geq 0.80$, MAF >0.4) were genotyped (Supplemental Table S-1). Thus, a total of 48 new markers were added to the region beyond the initial six genome scan markers.

Genetic and physical maps

The genetic position for all markers was determined from the revised deCODE map (Nievergelt et al. 2004). To

place SNPs and microsatellites that were not part of the deCODE map, the UCSC Genome Browser build 36 (March 2006 human reference sequence, NCBI Build 36.1) provided the framework to perform linear interpolation. One marker, ATTT019, was not found on either map, and therefore, was located by blasting its primer sequences.

Genotyping Mayo Clinic and PROGRESS families

SNPs genotyped on the Mayo Clinic pedigrees were typed using standardized protocols in accordance with the manufacturer's recommendations for either the "Early Access Affymetrix Mapping 10 K SNP array" (analysis performed at Mayo) or the "Illumina GoldenGate platform" (analysis performed at CIDR). Genotypes were called using software provided by either Affymetrix for the 10 K SNP array or Illumina for the GoldenGate assays.

Primer-sequences for microsatellites used in the PROGRESS pedigrees were provided by GDB (Genomic Data Base, http://www.gdb.org), and ordered from Invitrogen (Invitrogen, Carlsbad, CA) (Supplemental Table S-1). The amplification reaction (10 μl) contained 20 ng of amplified-template DNA and utilized standard amplification conditions. Reactions were analyzed on a GeneAmp PCR System 9700 thermocycler (Applied Biosystems, Foster City, CA). The size of the microsatellite fragments was detected using an ABI 3730 xl Sequencer (Applied Biosystems) and the resulting data analyzed using the GENEMAPPER software (v4.0) (Applied Biosystems). Erroneous genotypes were detected using the "error" command in the Merlin software, and by visual examination of the predicted haplotypes. If untyped affected individuals could be extrapolated from typed relatives, they were included in the analysis as well. This allowed inclusion of six affected men from five PROGRESS families.

For the SNP analysis (PROGRESS families), primer sequences were designed using Primer3 (Rozen and Skaletsky 2000), and genotyped by direct sequencing using 10 ng genomic DNA as template with standard PCR conditions for the BigDye Terminator Cycle Sequencing Kit (v3.1) (Applied Biosystems). Products were separated on the ABI 3730 xl Sequencer (Applied Biosystems) and analyzed using the Mutation Surveyor software, (http://www.softgenetics.com/mutationSurveyor.html). All genotypes were obtained using both forward and reverse sequencing data.

Linkage analysis

Analysis of the data derived from the PROGRESS families utilized a model-based multipoint linkage approach performed using GENEHUNTER (v2.1_r5 beta) (Kruglyak

et al. 1996; Markianos et al. 2001). Because of the large number of SNPs utilized for the analysis, the Mayo Clinic families were analyzed using MERLIN (v1.0.1) (Abecasis et al. 2002; Cook Jr 2002). A simplification of the "Smith model" was used (Smith et al. 1996), such that the dominant mode and rare disease-allele frequency ($q = 0.003$) remained the same, but only two liability classes were assigned and an "affected only" analysis was done. Class I included all affected men with 100% penetrance and no phenocopies. Class 2 included all other individuals and assumed 50% penetrance for all genotypes, which means that their phenotype did not contribute to the LOD scores, only their genetic marker data. By using 100% penetrance and no phenocopies, we forced recombinations to be observed. However, because phenocopies might be present, which could infer false recombinations, we did not rely on single flanking recombinants, but rather considered two or three recombinants on either side. The frequencies of all marker alleles were estimated from all subjects, ignoring genetic relationships. The increment value was set to calculate the LOD score at every cM for PROGRESS families and to calculate the LOD score at five equally spaced locations between each pair of markers for the Mayo Clinic families.

Haplotyping

For both the Mayo and PROGRESS families, the most likely haplotypes given the observed genotype data were reconstructed using the MERLIN "best" command (v1.0.1) (Abecasis et al. 2002; Cook Jr 2002). Least number of crossing-over events was assumed. The outermost marker adjacent to any recombination breakpoint was included in the shared haplotype. If two haplotypes were found to segregate among all affected men in the family, a "combined" shared haplotype was constructed consisting of the two overlapping chromosomal intervals. When a breakpoint was located in a region of several markers that had unclear phase information, the whole region with unclear phase was included in the haplotype. For the Mayo pedigrees, recombinants identified in the most likely haplotypes reported by MERLIN were further verified by the presence of a sharp drop in the individual pedigree LOD trace of >0.5 LOD units. If a MERLIN-inferred recombinant was not accompanied by a corresponding drop in the individual LOD trace, the pedigree was considered nonrecombinant. Recombinant boundaries were defined as the physical position of the SNPs residing immediately outside the shared haplotype segment. Note that this type of analysis focused on within-family haplotypes to focus on recombination breakpoints, but not between family haplotypes, because LD could bias inference of haplotypes across families.

References

Abecasis GR, Cherny SS, Cookson WO, Cardon LR (2002) Merlin–rapid analysis of dense genetic maps using sparse gene flow trees. Nat Genet 30:97–101

Berry R, Schroeder JJ, French AJ, McDonnell SK, Peterson BJ, Cunningham JM, Thibodeau SN, Schaid DJ (2000) Evidence for a prostate cancer-susceptibility locus on chromosome 20. Am J Hum Genet 67:82–91

Bratt O (2002) Hereditary prostate cancer: clinical aspects. J Urol 168:906–913

Camp NJ, Farnham JM, Cannon Albright LA (2005) Genomic search for prostate cancer predisposition loci in Utah pedigrees. Prostate 65:365–374

Camp NJ, Farnham JM, Cannon-Albright LA (2006) Localization of a prostate cancer predisposition gene to an 880-kb region on chromosome 22q12.3 in Utah high-risk pedigrees. Cancer Res 66:10205–10212

Camp NJ, Cannon-Albright LA, Farnham JM, Baffoe-Bonnie AB, George A, Powell I, Bailey-Wilson JE, Carpten JD, Giles GG, Hopper JL, Severi G, English DR, Foulkes WD, Maehle L, Moller P, Eeles R, Easton D, Badzioch MD, Whittemore AS, Oakley-Girvan I, Hsieh CL, Dimitrov L, Xu J, Stanford JL, Johanneson B, Deutsch K, McIntosh L, Ostrander EA, Wiley KE, Isaacs SD, Walsh PC, Thibodeau SN, McDonnell SK, Hebbring S, Schaid DJ, Lange EM, Cooney KA, Tammela TL, Schleutker J, Paiss T, Maier C, Gronberg H, Wiklund F, Emanuelsson M, Isaacs WB (2007) Compelling evidence for a prostate cancer gene at 22q12.3 by the International Consortium for Prostate Cancer Genetics. Hum Mol Genet 16:1271–1278

Carlson CS, Eberle MA, Rieder MJ, Yi Q, Kruglyak L, Nickerson DA (2004) Selecting a maximally informative set of single-nucleotide polymorphisms for association analyses using linkage disequilibrium. Am J Hum Genet 74:106–120

Carter BS, Beaty TH, Steinberg GD, Childs B, Walsh PC (1992) Mendelian inheritance of familial prostate cancer. Proc Natl Acad Sci USA 89:3367–3371

Chang BL, Isaacs SD, Wiley KE, Gillanders EM, Zheng SL, Meyers DA, Walsh PC, Trent JM, Xu J, Isaacs WB (2005) Genome-wide screen for prostate cancer susceptibility genes in men with clinically significant disease. Prostate 64:356–361

Chang BL, Lange EM, Dimitrov L, Valis CJ, Gillanders EM, Lange LA, Wiley KE, Isaacs SD, Wiklund F, Baffoe-Bonnie A, Langefeld CD, Zheng SL, Matikainen MP, Ikonen T, Fredriksson H, Tammela T, Walsh PC, Bailey-Wilson JE, Schleutker J, Gronberg H, Cooney KA, Isaacs WB, Suh E, Trent JM, Xu J (2006) Two-locus genome-wide linkage scan for prostate cancer susceptibility genes with an interaction effect. Hum Genet 118:716–724

Chang KW, Lee TC, Yeh WI, Chung MY, Liu CJ, Chi LY, Lin SC (2004) Polymorphism in heme oxygenase-1 (HO-1) promoter is related to the risk of oral squamous cell carcinoma occurring on male areca chewers. Br J Ca 91:1551–1554

Conlon EM, Goode EL, Gibbs M, Stanford JL, Badzioch M, Janer M, Kolb S, Hood L, Ostrander EA, Jarvik GP, Wijsman EM (2003) Oligogenic segregation analysis of hereditary prostate cancer

pedigrees: evidence for multiple loci affecting age at onset. Int J Cancer 105:630–635

Cook Jr EH (2002) Merlin: faster linkage analysis with improved genotyping error detection. Pharmacogenomics J 2:139–140

Cui J, Staples MP, Hopper JL, English DR, McCredie MR, Giles GG (2001) Segregation analyses of 1,476 population-based Australian families affected by prostate cancer. Am J Hum Genet 68:1207–1218

Cunningham JM, McDonnell SK, Marks A, Hebbring S, Anderson SA, Peterson BJ, Slager S, French A, Blute ML, Schaid DJ, Thibodeau SN (2003) Genome linkage screen for prostate cancer susceptibility loci: results from the Mayo Clinic Familial Prostate Cancer Study. Prostate 57:335–3346

Cybulski C, Wokolorczyk D, Huzarski T, Byrski T, Gronwald J, Gorski B, Debniak T, Masojc B, Jakubowska A, Gliniewicz B, Sikorski A, Stawicka M, Godlewski D, Kwias Z, Antczak A, Krajka K, Lauer W, Sosnowski M, Sikorska-Radek P, Bar K, Klijer R, Zdrojowy R, Malkiewicz B, Borkowski A, Borkowski T, Szwiec M, Narod SA, Lubinski J (2006) A large germline deletion in the Chek2 kinase gene is associated with an increased risk of prostate cancer. J Med Genet 43:863–866

Dong X, Wang L, Taniguchi K, Wang X, Cunningham JM, McDonnell SK, Qian C, Marks AF, Slager SL, Peterson BJ, Smith DI, Cheville JC, Blute ML, Jacobsen SJ, Schaid DJ, Tindall DJ, Thibodeau SN, Liu W (2003) Mutations in CHEK2 associated with prostate cancer risk. Am J Hum Genet 72:270–280

Dorkeld F, Bernheim A, Dessen P, Huret JL (1999) A database on cytogenetics in haematology and oncology. Nucleic Acids Res 27:353–354

Easton DF, Schaid DJ, Whittemore AS, Isaacs WJ (2003) Where are the prostate cancer genes?–A summary of eight genome wide searches. Prostate 57:261–269

Goldgar DE, Easton DF, Cannon-Albright LA, Skolnick MH (1994) Systematic population-based assessment of cancer risk in first-degree relatives of cancer probands. J Natl Cancer Inst 86:1600–1608

Gong G, Oakley-Girvan I, Wu AH, Kolonel LN, John EM, West DW, Felberg A, Gallagher RP, Whittemore AS (2002) Segregation analysis of prostate cancer in 1,719 white, African–American and Asian–American families in the United States and Canada. Cancer Causes Control 13:471–482

Gronberg H, Damber L, Damber JE, Iselius L (1997) Segregation analysis of prostate cancer in Sweden: support for dominant inheritance. Am J Epidemiol 146:552–557

IHC (2003) The international HapMap project. Nature 426:789–796

Janer M, Friedrichsen DM, Stanford JL, Badzioch MD, Kolb S, Deutsch K, Peters MA, Goode EL, Welti R, DeFrance HB, Iwasaki L, Li S, Hood L, Ostrander EA, Jarvik GP (2003) Genomic scan of 254 hereditary prostate cancer families. Prostate 57:309–319

Jemal A, Siegel R, Ward E, Murray T, Xu J, Thun MJ (2007) Cancer statistics, 2007. CA Cancer J Clin 57:43–66

Keetch DW, Rice JP, Suarez BK, Catalona WJ (1995) Familial aspects of prostate cancer: a case control study. J Urol 154:2100–2102

Kruglyak L, Daly MJ, Reeve-Daly MP, Lander ES (1996) Parametric and nonparametric linkage analysis: a unified multipoint approach. Am J Hum Genet 58:1347–1363

Lange EM, Gillanders EM, Davis CC, Brown WM, Campbell JK, Jones M, Gildea D, Riedesel E, Albertus J, Freas-Lutz D, Markey C, Giri V, Dimmer JB, Montie JE, Trent JM, Cooney KA (2003) Genome-wide scan for prostate cancer susceptibility genes using families from the University of Michigan prostate cancer genetics project finds evidence for linkage on chromosome 17 near BRCA1. Prostate 57:326–334

Markianos K, Daly MJ, Kruglyak L (2001) Efficient multipoint linkage analysis through reduction of inheritance space. Am J Hum Genet 68:963–977

McIndoe RA, Stanford JL, Gibbs M, Jarvik GP, Brandzel S, Neal CL, Li S, Gammack JT, Gay AA, Goode EL, Hood L, Ostrander EA (1997) Linkage analysis of 49 high-risk families does not support a common familial prostate cancer-susceptibility gene at 1q24–25. Am J Hum Genet 61:347–353

Nievergelt CM, Smith DW, Kohlenberg JB, Schork NJ (2004) Large-scale integration of human genetic and physical maps. Genome Res 14:1199–1205

Ostrander EA, Stanford JL (2000) Genetics of prostate cancer: too many loci, too few genes. Am J Hum Genet 67:1367–1375

Ostrander EA, Markianos K, Stanford JL (2004) Finding prostate cancer susceptibility genes. Annu Rev Genomics Hum Genet 5:151–175

Pakkanen S, Baffoe-Bonnie AB, Matikainen MP, Koivisto PA, Tammela TL, Deshmukh S, Ou L, Bailey-Wilson JE, Schleutker J (2007) Segregation analysis of 1,546 prostate cancer families in Finland shows recessive inheritance. Hum Genet 121:257–267

Rozen S, Skaletsky H (2000) Primer3 on the WWW for general users and for biologist programmers. Methods Mol Biol 132:365–386

Schaid DJ (2004) The complex genetic epidemiology of prostate cancer. Hum Mol Genet 13(Spec No 1):R103–R121

Schaid DJ, McDonnell SK, Blute ML, Thibodeau SN (1998) Evidence for autosomal dominant inheritance of prostate cancer. Am J Hum Genet 62:1425–1438

Schaid DJ, Guenther JC, Christensen GB, Hebbring S, Rosenow C, Hilker CA, McDonnell SK, Cunningham JM, Slager SL, Blute ML, Thibodeau SN (2004) Comparison of microsatellites versus single-nucleotide polymorphisms in a genome linkage screen for prostate cancer-susceptibility Loci. Am J Hum Genet 75:948–965

Seppala EH, Ikonen T, Mononen N, Autio V, Rokman A, Matikainen MP, Tammela TL, Schleutker J (2003) CHEK2 variants associate with hereditary prostate cancer. Br J Cancer 89:1966–1970

Smith JR, Freije D, Carpten JD, Gronberg H, Xu J, Isaacs SD, Brownstein MJ, Bova GS, Guo H, Bujnovszky P, Nusskern DR, Damber JE, Bergh A, Emanuelsson M, Kallioniemi OP, Walker-Daniels J, Bailey-Wilson JE, Beaty TH, Meyers DA, Walsh PC, Collins FS, Trent JM, Isaacs WB (1996) Major susceptibility locus for prostate cancer on chromosome 1 suggested by a genome-wide search. Science 274:1371–1374

Stanford JL, McDonnell SK, Friedrichsen DM, Carlson EE, Kolb S, Deutsch K, Janer M, Hood L, Ostrander EA, Schaid DJ (2006) Prostate cancer and genetic susceptibility: a genome scan incorporating disease aggressiveness. Prostate 66:317–325

Tapper WJ, Morton NE, Dunham I, Ke X, Collins A (2001) A sequence-based integrated map of chromosome 22. Genome Res 11:1290–1295

Valeri A, Briollais L, Azzouzi R, Fournier G, Mangin P, Berthon P, Cussenot O, Demenais F (2003) Segregation analysis of prostate cancer in France: evidence for autosomal dominant inheritance and residual brother-brother dependence. Ann Hum Genet 67:125–137

Verhage BA, Baffoe-Bonnie AB, Baglietto L, Smith DS, Bailey-Wilson JE, Beaty TH, Catalona WJ, Kiemeney LA (2001) Autosomal dominant inheritance of prostate cancer: a confirmatory study. Urology 57:97–101

Xu J, Dimitrov L, Chang BL, Adams TS, Turner AR, Meyers DA, Eeles RA, Easton DF, Foulkes WD, Simard J, Giles GG, Hopper JL, Mahle L, Moller P, Bishop T, Evans C, Edwards S, Meitz J, Bullock S, Hope Q, Hsieh CL, Halpern J, Balise RN, Oakley-Girvan I, Whittemore AS, Ewing CM, Gielzak M, Isaacs SD, Walsh PC, Wiley KE, Isaacs WB, Thibodeau SN, McDonnell SK, Cunningham JM, Zarfas KE, Hebbring S, Schaid DJ, Friedrichsen DM, Deutsch K, Kolb S, Badzioch M, Jarvik GP, Janer M, Hood L, Ostrander EA, Stanford JL, Lange EM, Beebe-Dimmer JL, Mohai CE, Cooney KA, Ikonen T, Baffoe-Bonnie A, Fredriksson H, Matikainen MP, Tammela T, Bailey-Wilson J, Schleutker J, Maier C, Herkommer K, Hoegel JJ, Vogel W, Paiss T, Wiklund F, Emanuelsson M, Stenman E, Jonsson BA, Gronberg H, Camp NJ, Farnham J, Cannon-Albright LA, Seminara D (2005) A combined genomewide linkage scan of 1,233 families for prostate cancer-susceptibility genes conducted by the international consortium for prostate cancer genetics. Am J Hum Genet 77:219–229

Zhang Y, Kiel DP, Ellison RC, Schatzkin A, Dorgan JF, Kreger BE, Cupples LA, Felson DT (2002) Bone mass and the risk of prostate cancer: the Framingham Study. Am J Med 113:734–739