Asian Journal of Andrology (2013) 15, 309–313
© 2013 AJA, SIMM & SJTU. All rights reserved 1008-682X/13 $32.00
www.nature.com/aja

REVIEW

# Genetics and genomics of prostate cancer

Michael Dean[1] and Hong Lou[2]

**Prostate cancer (PCa) is one of the most common malignancies in the world with over 890 000 cases and over 258 000 deaths worldwide each year. Nearly all mortalities from PCa are due to metastatic disease, typically through tumors that evolve to be hormone-refractory or castrate-resistant. Despite intensive epidemiological study, there are few known environmental risk factors, and age and family history are the major determinants. However, there is extreme heterogeneity in PCa incidence worldwide, suggesting that major determining factors have not been described. Genome-wide association studies have been performed and a considerable number of significant, but low-risk loci have been identified. In addition, several groups have analyzed PCa by determination of genomic copy number, fusion gene generation and targeted resequencing of candidate genes, as well as exome and whole genome sequencing. These initial studies have examined both primary and metastatic tumors as well as murine xenografts and identified somatic alterations in *TP53* and other potential driver genes, and the disturbance of androgen response and cell cycle pathways. It is hoped that continued characterization of risk factors as well as gene mutation and misregulation in tumors will aid in understanding, diagnosing and better treating PCa.**
*Asian Journal of Andrology* (2013) **15,** 309–313; doi:10.1038/aja.2013.29; published online 8 April 2013

**Keywords:** androgen receptor; cancer progression; chromatin remodeling; metastasis; prostate cancer (PCa); risk factor; somatic mutation; tumor heterogeneity

## PROSTATE CANCER (PCA)—GLOBAL IMPACT AND RISK FACTORS

PCa is the second most common cancer of men with 899 000 new cases (14% of all cancers) reported each year.[1,2] However, the rate of PCa incidence varies by over 25-fold globally with Australia and New Zealand, Europe and Northern America having the highest incidence rates. Incidence is also high in populations of African descent, the Caribbean, South America and sub-Saharan Africa, and very low in South-Central Asia (4.1 per 100 000).

In 2008 (the last year for which global statistics are available), there was an estimated 258 000 deaths from PCa making it the sixth most common cause of death from cancer in men (**Figure 1**). The percentage of cancer mortalities in men due to PCa ranges from 0.6%–4.8% in Asia, to 9.7%–16% in North, Central and South America Europe and Australia (**Supplementary Table 1**). It is generally agreed that the widespread use of the prostate-specific antigen test in developed countries leads to a higher diagnosis and PCa represents 20%–30% of all male cancer diagnoses in these countries (**Supplementary Table 1**). However, in Asia, PCa represents only 1%–10% of male cancer cases, but the incidence is rising dramatically.[3]

PCa is a disease of aging with a majority of both cases and deaths occurring in men in their 70s (http://seer.cancer.gov/statfacts/html/prost.html).[4] The extreme differences in incidence and mortality in Asia as compared to the rest of the world have prompted searches for dietary or environmental factors. As for breast cancer, Japanese have extremely low rates that rise if they are first-generation immigrants to the United States and yet further for subsequent generations, suggesting that there are indeed very strong environmental factors.[5,6] Dietary fat and particular animal fats have been implicated with odds ratios as high as 3.6 for animal fat consumption in the most positive studies. Similar modest associations with body mass index and PCa have been found. Neither smoking or alcohol use show strong associations nor have studies of nutrient levels and supplementation yielded highly positive correlations. Androgens are clearly required as eunuchs and people with androgen deficiency have very low risk, but attempts to correlate testosterone and other androgen levels have not shown dramatic associations.

In conclusion, PCa is one of the few common cancers without a strong known lifestyle or infectious agent as a risk factor. The parallels to breast cancer are striking (i) gender-specific; (ii) hormonally driven; (iii) large geographic difference in incidence; and (iv) evidence for major lifestyle factors from immigration studies, but lack of clear association to specific dietary components.

## FAMILY HISTORY AND FAMILIAL PCA

While family history is a risk factor, only 15% of men with the disease have a first-degree relative with PCa compared to 8% of the general population.[7,8] A meta-analysis of multiple studies suggested that the risk is higher for men with a brother affected than a father, and that the familial risk is higher for early onset disease.[9] In Sweden, a nation-wide registry of cases was used to document 3- to 10-fold higher risks in men with a family history and genetic factors account for 12% of the disease.[10]

[1]Cancer and Inflammation Program, National Cancer Institute, National Institutes of Health, Frederick, MD 21702, USA and [2]Basic Science Program, SAIC-Frederick, National Cancer Institute, National Institutes of Health, Frederick, MD 21702, USA
Correspondence: Dr M Dean (deanm@mail.nih.gov)
Received: 16 January 2013; Revised: 25 February 2013; Accepted: 26 February 2013; Published online: 8 April 2013

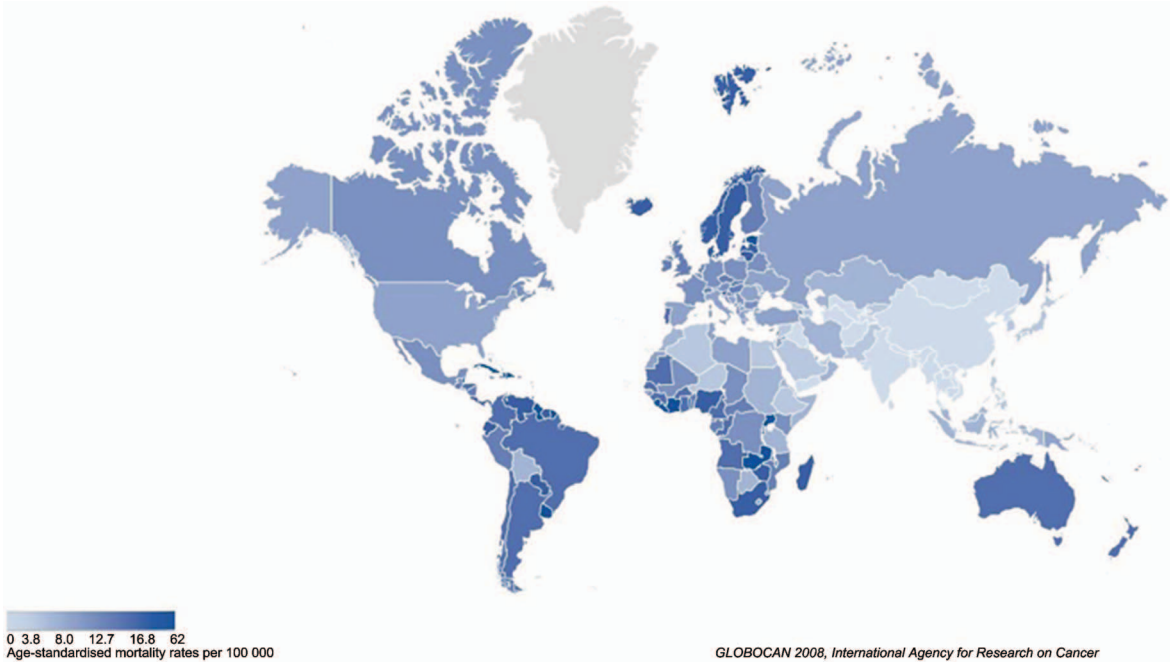


**Figure 1** PCa mortality. The map shows the incidence of PCa in the countries of the world. The data are from GLOBACAN2008. PCa, prostate cancer.

Linkage studies have been performed on large collections of PCa pedigrees and several candidate familial loci have been identified. The most convincing locus is the *HOXB13* gene, found in a region with a log-of-odds score of 5.5. A G84E mutation was identified in four of the linked pedigrees and this same mutation and others in the *HOXB13* gene were identified in additional families.[11–13] The G84E mutation is significantly associated with disease in men with a family history and/or early disease onset. The *HOXB13* gene binds to the androgen receptor and plays a role in prostate development.

Other loci implicated in familial studies include HPC1 (chromosome 1q25) with the *RNASEL* gene mutated in some families; PCAP (1q42-43); HPCX (Xq27-28); CAPB (1q36), with the *EPHB2* gene implicated;[14] and HPC20 (20q13). At least a dozen other loci have been implicated but no specific genes were identified and confirmed in these regions.[15]

Several studies have demonstrated that in Ashkenazi Jewish populations, the common *BRCA1* and *BRCA2* mutations identified in breast and ovarian cancer families confer risk for PCa. The odds ratios range from 2.1 to 4.8 and generally reach statistical significance for *BRCA2* but not *BRCA1*.[16–18] Additional studies in multiple populations support a role for *BRCA2* and some studies for *BRCA1* in early-onset, high-Gleason score disease and/or death by PCa.[18]

## GENOME-WIDE ASSOCIATION STUDIES (GWAS)

With the advent of microarrays capable of interrogating one million or more single nucleotide polymorphisms (SNPs), GWAS have been performed on most major cancer types including multiple studies on PCa.[12,19–24] Over 33 loci have been identified that qualify as being genome-wide significant and replicated in more than one study, or internally in multiple validation sets (**Supplementary Table 2**). The majority of these loci are in intergenic regions, many near known genes, and a few have known biological relevance to PCa. The odds ratios of these associations are modest (1.02–1.66) and so no one loci contributes highly to PCa risk.

The strongest associations are to a region on chromosome 8q24 about 500 kb 5′ to the *MYC* oncogene. Multiple SNPs in this region are associated, with association being different in different populations and multiple variants in this region are associated with other cancers.[19,20,25–27] Despite some effort to directly link these variants to *MYC* gene expression levels, no clear correlation could be found, but this region is known to interact at the chromatin level with the *MYC* locus.[26,27]

One of the strongest SNP associations is to the rs10993994 SNP on chromosome 10, in front of the *MSMB* gene.[21,23,28,29] *MSMB* encodes beta-microseminoprotein, an immunoglobulin binding factor family protein produced by the epithelial cells of the prostate gland and secreted into the seminal plasma. The MSMB protein has inhibin-like activity and may play a role as an autocrine paracrine factor in uterine, breast and other female reproductive tissues. The expression of the MSMB protein is decreased in PCa, and antibodies to the protein are used as biomarkers of PCa recurrence and/or progression. Due to the obvious connection of this protein to PCa and to incident metastatic disease,[30] considerable effort has been made in validating and understanding this result.

The rs10993994 SNP affects a predicted CREB transcription factor binding site[31] and this activity was validated by functional studies.[32] This association is reflected in the finding of reduced levels of expression of MSMB protein in urine and prostate tissue in multiple populations.[33–35] The SNP also shows association to PCa in Chinese populations and to MSMB levels in serum. Fine-mapping and resequencing confirmed that this SNP is the most significantly associated and that there are no common mutations/variants in the *MSMB* coding region.[36] However, the *MSMB* gene is also adjacent to the *NCOA4* gene, a known coactivator of androgen receptors and Lou *et al.*[37] documented that there are hybrid transcripts expressed from the *MSMB* promoter that contain the entire coding region of *NCOA4*. These hybrid transcripts can produce a stable protein and are regulated by CREB binding to the rs10993994 site. However,

Fitzgerald *et al.*[38] could find no association between SNPs in this region and overall *NCOA4* expression and Rinckleb *et al.*[39] presented data that there is an association of rs10993994 risk homozygosity and DNA repair capacity. Therefore, this PCa GWAS locus is very complex and the association signal may be a combination of very subtle effects on more than one gene/pathway.

Another loci with some functional validation is the rs12653946 polymorphism near the *IRX4* (Iroquois homeobox 4) gene at chromosome 5p15. *IRX4* is expressed in the prostate and heart, and there is a significant association between the genotype of rs12653946 and *IRX4* expression in normal prostate tissues. Knockdown of the *IRX4* gene enhances the growth of PCa cell lines and the protein interacts with the vitamin D receptor.[40]

PCa has been studied by GWAS more thoroughly than any other cancer except perhaps breast. Multiple studies in many countries and in all major ethnic groups have been carried out and a group of associated loci with high confidence are now known. However, translating this knowledge into further understanding of PCa development and progression or in use as effective biomarkers will require considerably more functional and clinical studies.

## SOMATIC ALTERATIONS IN TUMORS

In order to understand the somatic alterations that take place in the genomes of PCa tumors, several groups have begun sequencing exomes and genomes, carrying out methylation studies and performing RNA sequencing studies. Prior to undertaking this, there had been studies on mutations in prostate tumors of known oncogenes and tumor suppressors, revealing common mutations in genes such as *TP53*, *PIK3CA* and *PTEN*. The identification of high expression of *ERG*, an ETS-family transcription factor led to the identification of DNA deletions that result in the activation by gene fusion of *ERG* in many prostate tumors.[41,42] While the most common lesion is a nearly 3-MB deletion of chromosome 21 resulting in the fusion of *TMPRSS2* and *ERG*, other deletions and translocations have been described resulting in fusions of *ERG* or other ETS-family genes (*ETV1*, *ETV4* and *ETV5*) resulting in overexpression of a hybrid ETS-family protein (reviewed in Clark and Cooper[43]). These lesions can be found in up to half of all PCas in people of European descent; however, there has been little advance in using these lesions as biomarkers for PCa therapy or recurrence, and the function of the ERG protein in the tumor is poorly understood (reviewed in Rosen *et al.*[44]).

Three studies [45–47] have been published to date with genome/exome sequence of more than 10 prostate tumors. One challenge with PCa is the high proportion of normal cells and heterogeneity in the primary tumor. To address this, some groups have either passaged xenografts in mice or sequenced metastatic tumors. Kumar *et al.*[45] performed exome capture and sequenced xenografts from 16 lethal metastatic tumors and three primary tumors. They did not have corresponding normal DNA, but used filters to identify the most likely somatic variants and identified recurrent mutations in the *TP53*, *DLK2*, *GPC6*, *SDF4* and 19 other genes. Three of these tumors exhibited a dramatic increase in the number of mutations (2500–4000) as compared to the other 20 samples that had $362 \pm 147$ mutations. The molecular basis for this hyper-mutability phenotype was not determined, but has been seen in other tumors.

Barbieri *et al.*[46] performed exome capture and sequenced 112 treatment naive prostate adenocarcinomas and matched normal samples from the United States and Australia. They identified mutations in the known genes, *PIK3CA*, *TP53* and *PTEN* as well as *FOXA1*, *MED12*, *THSD7B*, *SCN11A* and *ZNF595*, and *CDKN1B* was mutated in three and deleted in 16 tumors, indicating that this gene is another common driver. The mutations in the *FOXA1* gene were missense variants clustered in the Forkhead domain. *FOXA1* regulates androgen receptor-driven transcription[48] providing a potential mechanism for these mutations to influence PCa development. *MED12* encodes a subunit of the cyclin-dependent kinase 8 (CDK8) and *MED12* mutations are found in up to 70% of uterine leiomyomas.[49] The mutations found in PCa are clustered in the active site.[46] The *SPOP* gene had been previously shown to be mutated in PCa[50] and was sequenced in an additional 300 primary tumors and metastases from the United States and Europe and recurrent heterozygous *SPOP* substitutions were identified in 6%–13% of primary tumors and in 15% of subjects with metastatic disease.[46] These *SPOP* mutations affect conserved residues in the putative substrate-binding cleft and these altered residues (Tyr87, Trp131 and Phe133) have key roles in substrate interaction,[51] and knockdown of *SPOP* increases invasiveness of PCa cell lines.[46]

Grasso *et al.*[47] sequenced the exomes of 50 lethal, metastatic castrate-resistant PCas and 11 treatment-naive, high-grade localized tumors. They identified mutations in the known genes *TP53*, *AR*, *ZFHX3*, *RB1*, *PTEN* and *APC* and described three additional significantly mutated genes: *MLL2*, *OR5L1* and *CDK12*. Through integration of the mutation and copy number data, they implicated the WNT and PTEN pathways as being altered in a large number of tumors. Grasso *et al.*[47] also described a frequent copy number loss on chromosome 5q21 at the location of *CHD1*, a gene encoding a chromatin-remodeling enzyme, and documented loss or mutation in multiple chromatin-remodeling genes, as has been found for multiple tumor types.[47,52–54] They documented that multiple of these chromatin-modifying genes directly interact with the androgen receptor and along with additional regulators, such as FOXA1, can partially explain the androgen-resistant nature of the tumors they sequenced.

**Supplementary Table 3** summarizes the results from all three of these studies. Only two genes were found mutated in all three studies, *TP53* and *MLL2*, with *PTEN* significantly mutated in two studies. The *SPOP*, *FOXA1*, *CDKN1B* and *MED12* genes described above are mutated in two of the three studies. However, there is clearly a great diversity of genes with many genes significantly mutated in one study and not detected in the other three. It is true that the designs of the studies were quite different with xenograft, primary tumors and metastatic tumors being the principle source of DNA for the studies, respectively. Thus part of the difference may be due to the starting material selected. Sensitivity of detection can also be an issue, especially in the Barbieri *et al.*[46] study using primary tumors as the contamination of normal cells and the heterogeneity of the primary tumor may mask the ability to detect some mutations.

Clearly the mutational landscape of PCa is extremely complex, and the several studies[47,55] that have also integrated expression and copy number analysis into the picture introduce other complications. The finding of multiple chromatin-remodeling genes mutated and connecting these mutations and that of other genes, such as *FOXA1*, in hormone signaling is an important advance, as nearly all lethal PCas are castrate-resistant. However, relatively few genes have emerged from these studies that can be targeted by currently available directed compounds, and the availability of metastatic tumor tissue is very limited in PCa, creating a problem for an individualized medicine strategy. While circulating tumor cells could provide a solution to this problem, it remains to be seen if these cells represent the actual cells with metastatic capacity.

The current studies were all performed on European populations, and will need to be repeated in people of African descent (where

incidence and mortality is considerably higher) and in Asian populations. Available data support that there are considerable molecular differences in prostate tumors between different ethnicities. Ren *et al.*[55] have shown that only about 20% of Chinese tumors have a *TMPRSS2-ERG* fusion, and *PTEN* deletions are also less common in Chinese tumors.[56] *ERG* fusions are also less frequent in African-American tumors (24%–31%) compared to Caucasian Americans (42%–50%) and ERG staining was also significantly different (29% and 63%, respectively)[57] as are Japanese tumors (16%).[58]

## CONCLUSIONS AND FUTURE PERSPECTIVES

The high incidence of PCa in much of the world and the very poor prognosis for metastatic disease has placed a high priority on understanding the causes and risk factors, developing effective prevention and detection strategies as well as effective treatments. While immigration studies and the rapidly rising incidence in parts of Asia support a key role of lifestyle factors, identifying those factors has been challenging. Some studies support an influence of 'Western diet' and both high fat content and animal protein have been implicated, but either proving this or changing diet sufficiently is unlikely to occur in the near term.

The serum prostate-specific antigen test is the most common diagnostic tool to identify early lesions and is widely used in many countries. However, the test results in up to 80% false-positive results, leading to a high number of prostate biopsies and additional tests that are unneeded. More concerning is the estimate of 17%–50% overdiagnosis (diagnosis of men who will remain asymptomatic during their lifetime). Nearly 90% of these men receive either radiation therapy, surgery or androgen deprivation therapy, all of which have moderate to severe side effects. Based on these concerns, the United State Preventive Services Task Force recently concluded that the benefits of prostate-specific antigen screening outweigh the harms (http://www.uspreventiveservicestaskforce.org/prostatecancerscreening.htm). Therefore, there is a tremendous need for other biomarkers of prostate tumors likely to develop into lethal disease.

The apparently large influence of genetic factors has influenced the undertaking of a large number of GWAS in multiple populations. This has led to the discovery of a large number of loci highly significantly associated. The most directly relevant is the rs10993994 SNP in the promoter region of MSMB that regulates promoter activity, expression and serum and urine levels of this important marker. However, clinical application of this result is uncertain. And although the constellation of markers can be used to estimate the risk of a given man, few of these markers are informative for outcome.

The best hope for improving our knowledge of this disease is comprehensive molecular characterization of the tumor. This has begun to be carried out and the first studies have uncovered new genes mutated, amplified, deleted and/or activated in PCa. The studies of castrate-resistant tumors have shed new light on genes implicated in androgen-responsiveness such as MLL2 and other chromatin-remodeling proteins and the FOXA1 transcription factor. However, the sequencing studies of primary tumors, and those of metastatic or xenografts are very different, suggesting that there is considerable heterogeneity and that a sample of the metastatic tumor is needed to identify the major cancer driver genes. The difficulty in accessing metastatic material in patients makes the application of personalized treatment difficult and the lack of targeted agents for the genes altered in PCa needs to be addressed.

The next generation of molecular studies of PCa need to address, not only exome and genome sequencing, but copy number analysis, epigenome analysis and expression and chromatin studies. As in all solid tumors, generating, annotating and combining these results in a meaningful fashion requires large numbers of samples, properly collected and annotated, but this approach will hopefully be informative for reducing the burden of the disease.

## COMPETING FINANCIAL INTERESTS

All authors declare that there are no competing financial interests.

## ACKNOWLEDGMENTS

This work was supported by the Intramural Research Program of the NIH, the National Cancer Institute Center for Cancer Research.

Supplementary Information accompanies the paper on *Asian Journal of Andrology*'s website (http://www.nature.com/aja).

1 Jemal A, Bray F, Center MM, Ferlay J, Ward E *et al*. Global cancer statistics. *CA Cancer J Clin* 2011; **61**: 69–90.
2 Ferlay J, Shin HR, Bray F, Forman D, Mathers C *et al*. Estimates of worldwide burden of cancer in 2008: GLOBOCAN 2008. *Int J Cancer* 2010; **127**: 2893–917.
3 Sim HG, Cheng CW. Changing demography of prostate cancer in Asia. *Eur J Cancer* 2005; **41**: 834–45.
4 Hayat MJ, Howlader N, Reichman ME, Edwards BK. Cancer statistics, trends, and multiple primary cancer analyses from the Surveillance, Epidemiology, and End Results (SEER) Program. *Oncologist* 2007; **12**: 20–37.
5 Haenszel W, Kurihara M. Studies of Japanese migrants. I. Mortality from cancer and other diseases among Japanese in the United States. *J Natl Cancer Inst* 1968; **40**: 43–68.
6 Shimizu H, Ross RK, Bernstein L, Yatani R, Henderson BE *et al*. Cancers of the prostate and breast among Japanese and white immigrants in Los Angeles County. *Br J Cancer* 1991; **63**: 963–6.
7 Steinberg GD, Carter BS, Beaty TH, Childs B, Walsh PC. Family history and the risk of prostate cancer. *Prostate* 1990; **17**: 337–47.
8 Stanford JL, Ostrander EA. Familial prostate cancer. *Epidemiol Rev* 2001; **23**: 19–23.
9 Zeegers MP, Jellema A, Ostrer H. Empiric risk of prostate carcinoma for relatives of patients with prostate carcinoma: a meta-analysis. *Cancer* 2003; **97**: 1894–903.
10 Damber L, Gronberg H, Damber JE. Familial prostate cancer and possible associated malignancies: nation-wide register cohort study in Sweden. *Int J Cancer* 1998; **78**: 293–7.
11 Ewing CM, Ray AM, Lange EM, Zuhlke KA, Robbins CM *et al*. Germline mutations in HOXB13 and prostate-cancer risk. *N Engl J Med* 2012; **366**: 141–9.
12 Gudmundsson J, Sulem P, Gudbjartsson DF, Blondal T, Gylfason A *et al*. Genome-wide association and replication studies identify four variants associated with prostate cancer susceptibility. *Nat Genet* 2009; **41**: 1122–6.
13 Breyer JP, Avritt TG, McReynolds KM, Dupont WD, Smith JR. Confirmation of the HOXB13 G84E germline mutation in familial prostate cancer. *Cancer Epidemiol Biomarkers Prev* 2012; **21**: 1348–53.
14 Huusko P, Ponciano-Jackson D, Wolf M, Kiefer JA, Azorsa DO *et al*. Nonsense-mediated decay microarray analysis identifies mutations of EPHB2 in human prostate cancer. *Nat Genet* 2004; **36**: 979–83.
15 Easton DF, Schaid DJ, Whittemore AS, Isaacs WJ. Where are the prostate cancer genes?—A summary of eight genome wide searches. *Prostate* 2003; **57**: 261–9.
16 Giusti RM, Rutter JL, Duray PH, Freedman LS, Konichezky M *et al*. A twofold increase in BRCA mutation related prostate cancer among Ashkenazi Israelis is not associated with distinctive histopathology. *J Med Genet* 2003; **40**: 787–92.
17 Kirchhoff T, Kauff ND, Mitra N, Nafa K, Huang H *et al*. BRCA mutations and risk of prostate cancer in Ashkenazi Jews. *Clin Cancer Res* 2004; **10**: 2918–21.
18 Gallagher DJ, Gaudet MM, Pal P, Kirchhoff T, Balistreri L *et al*. Germline BRCA mutations denote a clinicopathologic subset of prostate cancer. *Clin Cancer Res* 2010; **16**: 2115–21.
19 Yeager M, Orr N, Hayes RB, Jacobs KB, Kraft P *et al*. Genome-wide association study of prostate cancer identifies a second risk locus at 8q24. *Nat Genet* 2007; **39**: 645–9.
20 Yeager M, Chatterjee N, Ciampa J, Jacobs KB, Gonzalez-Bosquet J *et al*. Identification of a new prostate cancer susceptibility locus on chromosome 8q24. *Nat Genet* 2009; **41**: 1055–7.
21 Thomas G, Jacobs KB, Yeager M, Kraft P, Wacholder S *et al*. Multiple loci identified in a genome-wide association study of prostate cancer. *Nat Genet* 2008; **40**: 310–5.
22 Schumacher FR, Berndt SI, Siddiq A, Jacobs KB, Wang Z *et al*. Genome-wide association study identifies new prostate cancer susceptibility loci. *Hum Mol Genet* 2011; **20**: 3867–75.
23 Eeles RA, Kote-Jarai Z, Giles GG, Olama AA, Guy M *et al*. Multiple newly identified loci associated with prostate cancer susceptibility. *Nat Genet* 2008; **40**: 316–21.
24 Eeles RA, Kote-Jarai Z, Al Olama AA, Giles GG, Guy M *et al*. Identification of seven new prostate cancer susceptibility loci through a genome-wide association study. *Nat Genet* 2009; **41**: 1116–21.

25 Gudmundsson J, Sulem P, Gudbjartsson DF, Masson G, Agnarsson BA *et al*. A study based on whole-genome sequencing yields a rare variant at 8q24 associated with prostate cancer. *Nat Genet* 2012; **44**: 1326–9.
26 Ahmadiyeh N, Pomerantz MM, Grisanzio C, Herman P, Jia L *et al*. 8q24 prostate, breast, and colon cancer risk loci show tissue-specific long-range interaction with MYC. *Proc Natl Acad Sci USA* 2010; **107**: 9742–6.
27 Sotelo J, Esposito D, Duhagon MA, Banfield K, Mehalko J *et al*. Long-range enhancers on 8q24 regulate c-Myc. *Proc Natl Acad Sci USA* 2010; **107**: 3001–5.
28 Guy M, Kote-Jarai Z, Giles GG, Al Olama AA, Jugurnauth SK *et al*. Identification of new genetic risk factors for prostate cancer. *Asian J Androl* 2009; **11**: 49–55.
29 Chang BL, Cramer SD, Wiklund F, Isaacs SD, Stevens VL *et al*. Fine mapping association study and functional analysis implicate a SNP in MSMB at 10q11 as a causal variant for prostate cancer risk. *Hum Mol Genet* 2009; **18**: 1368–75.
30 Ahn J, Kibel AS, Park JY, Rebbeck TR, Rennert H *et al*. Prostate cancer predisposition loci and risk of metastatic disease and prostate cancer recurrence. *Clin Cancer Res* 2011; **17**: 1075–81.
31 Buckland PR, Hoogendoorn B, Coleman SL, Guy CA, Smith SK *et al*. Strong bias in the location of functional promoter polymorphisms. *Hum Mutat* 2005; **26**: 214–23.
32 Lou H, Yeager M, Li H, Bosquet JG, Hayes RB *et al*. Fine mapping and functional analysis of a common variant in MSMB on chromosome 10q11.2 associated with prostate cancer susceptibility. *Proc Natl Acad Sci USA* 2009; **106**: 7933–8.
33 Waters KM, Stram DO, le Marchand L, Klein RJ, Valtonen-Andre C *et al*. A common prostate cancer risk variant 5′ of microseminoprotein-beta (MSMB) is a strong predictor of circulating beta-microseminoprotein (MSP) levels in multiple populations. *Cancer Epidemiol Biomarkers Prev* 2010; **19**: 2639–46.
34 Whitaker HC, Kote-Jarai Z, Ross-Adams H, Warren AY, Burge J *et al*. The rs10993994 risk allele for prostate cancer results in clinically relevant changes in microseminoprotein-beta expression in tissue and urine. *PLoS One* 2010; **5**: e13363.
35 Xu B, Wang J, Tong N, Mi Y, Min Z *et al*. A functional polymorphism in MSMB gene promoter is associated with prostate cancer risk and serum MSMB expression. *Prostate* 2010; **70**: 1146–52.
36 Yeager M, Deng Z, Boland J, Matthews C, Bacior J *et al*. Comprehensive resequence analysis of a 97 kb region of chromosome 10q11.2 containing the MSMB gene associated with prostate cancer. *Hum Genet* 2009; **126**: 743–50.
37 Lou H, Li H, Yeager M, Im K, Gold B *et al*. Promoter variants in the MSMB gene associated with prostate cancer regulate MSMB/NCOA4 fusion transcripts. *Hum Genet* 2012; **131**: 1453–66.
38 Fitzgerald LM, Zhang X, Kolb S, Kwon EM, Liew YC *et al*. Investigation of the relationship between prostate cancer and MSMB and NCOA4 genetic variants and protein expression. *Hum Mutat* 2013; **34**: 149–56.
39 Rinckleb AE, Surowy HM, Luedeke M, Varga D, Schrader M *et al*. The prostate cancer risk locus at 10q11 is associated with DNA repair capacity. *DNA Repair (Amst)* 2012; **11**: 693–701.
40 Nguyen HH, Takata R, Akamatsu S, Shigemizu D, Tsunoda T *et al*. IRX4 at 5p15 suppresses prostate cancer growth through the interaction with vitamin D receptor, conferring prostate cancer susceptibility. *Hum Mol Genet* 2012; **21**: 2076–85.
41 Petrovics G, Liu A, Shaheduzzaman S, Furusato B, Sun C *et al*. Frequent overexpression of ETS-related gene-1 (ERG1) in prostate cancer transcriptome. *Oncogene* 2005; **24**: 3847–52.
42 Tomlins SA, Rhodes DR, Perner S, Dhanasekaran SM, Mehra R *et al*. Recurrent fusion of TMPRSS2 and ETS transcription factor genes in prostate cancer. *Science* 2005; **310**: 644–8.
43 Clark JP, Cooper CS. ETS gene fusions in prostate cancer. *Nat Rev Urol* 2009; **6**: 429–39.
44 Rosen P, Sesterhenn IA, Brassell SA, McLeod DG, Srivastava S *et al*. Clinical potential of the ERG oncoprotein in prostate cancer. *Nat Rev Urol* 2012; **9**: 131–7.
45 Kumar A, White TA, MacKenzie AP, Clegg N, Lee C *et al*. Exome sequencing identifies a spectrum of mutation frequencies in advanced and lethal prostate cancers. *Proc Natl Acad Sci USA* 2011; **108**: 17087–92.
46 Barbieri CE, Baca SC, Lawrence MS, Demichelis F, Blattner M *et al*. Exome sequencing identifies recurrent SPOP, FOXA1 and MED12 mutations in prostate cancer. *Nat Genet* 2012; **44**: 685–9.
47 Grasso CS, Wu YM, Robinson DR, Cao X, Dhanasekaran SM *et al*. The mutational landscape of lethal castration-resistant prostate cancer. *Nature* 2012; **487**: 239–43.
48 Gao N, Zhang J, Rao MA, Case TC, Mirosevich J *et al*. The role of hepatocyte nuclear factor-3 alpha (Forkhead Box A1) and androgen receptor in transcriptional regulation of prostatic genes. *Mol Endocrinol* 2003; **17**: 1484–507.
49 Makinen N, Mehine M, Tolvanen J, Kaasinen E, Li Y *et al*. MED12, the mediator complex subunit 12 gene, is mutated at high frequency in uterine leiomyomas. *Science* 2011; **334**: 252–5.
50 Berger MF, Lawrence MS, Demichelis F, Drier Y, Cibulskis K *et al*. The genomic complexity of primary human prostate cancer. *Nature* 2011; **470**: 214–20.
51 Zhuang M, Calabrese MF, Liu J, Waddell MB, Nourse A *et al*. Structures of SPOP-substrate complexes: insights into molecular architectures of BTB-Cul3 ubiquitin ligases. *Mol Cell* 2009; **36**: 39–50.
52 Li M, Zhao H, Zhang X, Wood LD, Anders RA *et al*. Inactivating mutations of the chromatin remodeling gene ARID2 in hepatocellular carcinoma. *Nat Genet* 2011; **43**: 828–9.
53 Gui Y, Guo G, Huang Y, Hu X, Tang A *et al*. Frequent mutations of chromatin remodeling genes in transitional cell carcinoma of the bladder. *Nat Genet* 2011; **43**: 875–8.
54 Jones S, Li M, Parsons DW, Zhang X, Wesseling J *et al*. Somatic mutations in the chromatin remodeling gene ARID1A occur in several tumor types. *Hum Mutat* 2012; **33**: 100–3.
55 Ren S, Peng Z, Mao JH, Yu Y, Yin C *et al*. RNA-seq analysis of prostate cancer in the Chinese population identifies recurrent gene fusions, cancer-associated long noncoding RNAs and aberrant alternative splicings. *Cell Res* 2012; **22**: 806–21.
56 Mao X, Yu Y, Boyd LK, Ren G, Lin D *et al*. Distinct genomic alterations in prostate cancers in Chinese and Western populations suggest alternative pathways of prostate carcinogenesis. *Cancer Res* 2010; **70**: 5207–12.
57 Rosen P, Pfister D, Young D, Petrovics G, Chen Y *et al*. Differences in frequency of ERG oncoprotein expression between index tumors of Caucasian and African American patients with prostate cancer. *Urology* 2012; **80**: 749–53.
58 Magi-Galluzzi C, Tsusuki T, Elson P, Simmerman K, LaFargue C *et al*. TMPRSS2-ERG gene fusion prevalence and class are significantly different in prostate cancer of Caucasian, African-American and Japanese patients. *Prostate* 2011; **71**: 489–97.