# A genome-wide search for loci interacting with known prostate cancer risk-associated genetic variants

Sha Tao[1], Zhong Wang[2,3], Junjie Feng[2,3], Fang-Chi Hsu[2,4], Guangfu Jin[2,3], Seong-Tae Kim[2,3], Zheng Zhang[2,3], Henrik Gronberg[5], Lilly S.Zheng[2,3], William B.Isaacs[6], Jianfeng Xu[1,2,3,7] and Jielin Sun[2,3,*]

[1]Center for Genetic Epidemiology and Prevention, Van Andel Research Institute, Grand Rapids, MI, USA, [2]Center for Cancer Genomics, Medical Center Boulevard, [3]Center for Genomics and Personalized Medicine Research, [4]Department of Biostatistical Sciences, Wake Forest University School of Medicine, Winston-Salem, NC 27157, USA, [5]Department of Medical Epidemiology and Biostatistics, Karolinska Institute, Stockholm 17177, Sweden, [6]Department of Urology, Johns Hopkins Medical Institutions, Baltimore, MD 21201, USA and [7]Department of Urology, Wake Forest University School of Medicine, Winston-Salem, NC 27157, USA

[*]To whom correspondence should be addressed. Tel: +1 (336) 713 7500; Fax: +1 (336) 713-7566;
Email: jisun@wfubmc.edu

**Genome-wide association studies (GWAS) have identified ∼30 single-nucleotide polymorphisms (SNPs) consistently associated with prostate cancer (PCa) risk. To test the hypothesis that other sequence variants in the genome may interact with those 32 known PCa risk-associated SNPs identified from GWAS to affect PCa risk, we performed a systematic evaluation among three existing PCa GWAS populations: CAncer of the Prostate in Sweden population, a Johns Hopkins Hospital population, and the Cancer Genetic Markers of Susceptibility population, with a total sample size of 4723 PCa cases and 4792 control subjects. Meta-analysis of the interaction term between each of those 32 SNPs and SNPs in the genome was performed in three PCa GWAS populations. The most significant interaction detected was between rs12418451 in *MYEOV* and rs784411 in *CEP152*, with a $P_{interaction}$ of $1.15 \times 10^{-7}$ in the meta-analysis. In addition, we emphasized two pairs of interactions with potential biological implication, including an interaction between rs7127900 near insulin-like growth factor-2 (*IGF2*)/*IGF2AS* and rs12628051 in *TNRC6B*, with a $P_{interaction}$ of $3.39 \times 10^{-6}$ and an interaction between rs7679763 near *TET2* and rs290258 in *SYK*, with a $P_{interaction}$ of $1.49 \times 10^{-6}$. Those results show statistical evidence for novel loci interacting with known risk-associated SNPs to modify PCa risk. The interacting loci identified provide hints on the underlying molecular mechanism of the associations with PCa risk for the known risk-associated SNPs. Additional studies are warranted to further confirm the interaction effects detected in this study.**

## Introduction

Prostate cancer (PCa) is the most common non-skin cancer affecting men in western countries. Inherited genetic variants play an important role in contributing to familial aggregation of PCa. Since 2007, genome-wide association studies (GWAS) successfully identified at least 33 PCa risk-associated single-nucleotide polymorphisms (SNPs) (1–15).

Although those risk-associated SNPs are well replicated in multiple studies (16–20), very few studies assess the potential epistasis or gene–gene interaction between those SNPs and the rest of SNPs that reside in the genome. In fact, epistatic effect is the norm rather than

**Abbreviations:** CAPS, CAncer of the Prostate in Sweden; CGEMS, Cancer Genetic Markers of Susceptibility; CI, confidence interval; GWAS, genome-wide association studies; IGF, insulin-like growth factor; JHH, Johns Hopkins Hospital; OR, odds ratio; PCa, prostate cancer; RNAi, RNA interference; SNP, single-nucleotide polymorphism.

exception for complex diseases, such as PCa. Inference from tumor-igenesis and results from genetic modeling studies suggest that multiple susceptibility genes, either additively or multiplicatively, determine individual risk to PCa. The importance of epistasis is also supported by empirical evidence from model organisms and human studies (21–23).

The evidence of epistatic effect from the empirical data suggests that gene–gene interactions need to be examined in GWAS when searching for PCa risk variants. Actually, assessment of gene–gene interaction may reveal additional PCa risk variants, especially in the situation where multiple risk-associated variants have been identified. It is computationally possible to use a logistic regression model to search the genome and to identify additional variants that interact with these known risk variants to modify the risk of developing PCa.

Recently, Ciampa *et al.* (24) reported a two-stage GWAS of epistasis between 13 known PCa risk-associated SNPs and SNPs across the genome in the National Cancer Institute Cancer Genetic Markers of Susceptibility (CGEMS) Stage I population with 523 841 SNPs and Stage II population with 27 383 SNPs, which were selected based on the main effects in Stage I. No SNP–SNP interaction was identified that reached a genome-wide significance level in Stage I or Stage II data, and a list of top interactions were suggested and warranted replication in other studies. The lack of replication data in Ciampa's study emphasized the importance of evaluating gene–gene interaction in multiple GWAS populations. More importantly, combining individual level data of multiple GWAS can improve the power to identify SNPs that interact with the known risk-associated SNPs to impact PCa risk. To this end, we performed a combined genome-wide search for SNPs that interact with 32 PCa risk-associated variants identified from GWAS in three case–control populations of European descents, including 1583 PCa cases and 519 control subjects from the CAncer Prostate in Sweden (CAPS), 1964 PCa cases and 3172 control subjects from a Johns Hopkins Hospital (JHH) PCa and iControl database and 1176 PCa cases and 1101 control subjects in the National Cancer Institute CGEMS study. We also evaluated the list of SNP–SNP interactions suggested by Ciampa's study in the two independent GWAS populations (CAPS and JHH).

## Materials and methods

### Study populations

The first GWAS population included 1583 PCa patients and 519 control subjects that matched the age distribution of case subjects from CAPS, a population-based PCa case–control study from Sweden (CAPS) (6). Briefly, the CAPS population was recruited from four regional cancer registries in Sweden and diagnosed between July 2001 and October 2003. The clinical characteristics of these patients are presented in Supplementary Table 1, available at *Carcinogenesis* Online.

The second population was from a JHH PCa GWAS, which included 1964 PCa cases and 3172 control subjects. The cases are Caucasian PCa patients who underwent radical prostatectomy for the treatment of PCa at JHH from 1 January 1999 through 31 December 2008 (25). The clinical characteristics of these patients are presented in Supplementary Table 2, available at *Carcinogenesis* Online. The control subjects for this population were an independent group of Caucasian individuals from the Illumina iControlDB (iControls) dataset (https://www.illumina.com/science/icontrodb.ilmn).

The third population was obtained from Stage I of the National Cancer Institute CGEMS study. It included 1176 PCa cases and 1101 control subjects, selected from the Prostate, Lung, Colon and Ovarian Cancer Screening Trial (6,9). The genotype and phenotype data of the study are publicly available and our use of the data was approved by CGEMS.

### Genotype data, imputation and quality control

GWAS of the CAPS population was performed using Affymetrix 5.0 chip. GWAS of the JHH case population was performed using the Illumina 610K chip (24). GWAS of the iControls population (25) was performed using

Illumina Hap300 and Hap550 chips. GWAS of the CGEMS population was performed using HumanHap300 and HumanHap240 assays from Illumina Corp.

For each GWAS population, we imputed all the known SNPs that are catalogued in HapMap Phase II (www.hapmap.org) using the IMPUTE computer program (26) with a posterior probability of 0.9 as a threshold to call genotypes. Individuals with a call rate <0.95 were removed from GWAS analysis. The following quality control criteria were used to filter SNPs: Minor Allele Frequency < 0.01, Hardy-Weinberg Equilibrium < 0.001 and call rate <0.95.

*PCa-known risk SNPs identified from GWAS*

The 33 PCa-known risk-associated SNPs were discovered by GWAS and the following fine-mapping studies, with $P$-values $\leq 10 \times 10^{-7}$ (1–15). The detailed information for the 33 risk SNPs are presented in Table I. The SNP rs16901979 was not evaluated in the following interaction analysis due to the unavailability of imputation of this SNP since it was not catalogued in the HapMap database.

*Statistical analysis*

Multiplicative interactions between each one of the 32 known PCa risk variants and each SNP in the genome were systematically tested by including both SNPs and an interaction term (product of two SNPs), as implemented in the computer program PLINK (27). Ancestral proportions obtained based on EIGENSOFT software (28) were included as covariates to minimize the impact of potential population stratification in the JHH population. An additive genetic model was used, where the genotypes were coded as 0, 1 and 2 and each SNP was treated as a continuous variable. The interaction term was tested using a Wald test with degree of freedom of 1. A meta-analysis of the interaction term for the three study populations was performed using the method developed by Manning *et al.* (29). Briefly, the meta-odds ratio ($OR_M$) of the interaction term across the three populations was estimated using an inverse variance weighted meta-analysis, where $\ln\left(OR_M\right) = \sum_{i=1}^{3}\left(w_i \ln\left(OR_i\right) \middle/ \sum_{i=1}^{3} w_i\right)$, $w_i = 1/\mathrm{var}\left(\ln(OR_i)\right)$ and $se\left(\ln\left(OR_M\right)\right) = \left[\sum_{i=1}^{3}(1/\mathrm{var}(\ln(OR_i)))\right]^{-1/2}$ (29).

**Results**

After imputation and applying quality control criteria, 1 314 700, 1 646 196 and 1 757 946 SNPs remained for CAPS, JHH and CGEMS studies, respectively. A total of 1 117 531 common SNPs for those three populations were used in the interaction analysis.

We examined the inflation factor and the quantile–quantile plots for interaction tests in the combined analysis of three populations. No systematic bias was observed as the inflation factors for the 32 GWAS scans for SNP–SNP interactions ranged from 0.98 to 1.03 (Supplementary Table 1 is available at *Carcinogenesis* Online).

The results for the top-ranked SNPs that interacted with each of the 32 known PCa risk SNPs ($P_{\mathrm{interaction}} < 1.0 \times 10^{-5}$ in the meta-analysis) were presented in Supplementary Table 2, available at *Carcinogenesis* Online. For SNPs in linkage disequilibrium (as defined by $r^2 > 0.5$), only the one with the smallest $P$-value based on meta-analysis was included in the Supplementary Table 2, available at *Carcinogenesis* Online. We then further examined the interaction effects for the top-ranked SNPs ($P_{\mathrm{interaction}} < 1 \times 10^{-5}$) in each of the three populations. SNPs that significantly interacted with the 32 SNPs in all three populations at a nominal $P_{\mathrm{interaction}}$ of 0.05 were presented in Table II. No SNP–SNP interaction reached a genome-wide significant level of $1.5 \times 10^{-9}$ [$0.05/(1 \times 10^{-6} \times 32)$]. The most significant interaction was observed between rs12418451 in the *MYEOV* gene region and rs784411 in the intron of *CEP152*, with a $P_{\mathrm{interaction}}$ of $1.15 \times 10^{-7}$ [$OR_{\mathrm{interaction}} = 1.42$; 95% confidence interval (CI): 1.25–1.61] in the meta-analysis. This interaction pair was significant in all three populations and the effects of the interaction were in the same direction [$P_{\mathrm{interaction}} = 0.008$, $OR_{\mathrm{interaction}} = 1.55$ (95% CI: 1.12–2.16) for CAPS; $P_{\mathrm{interaction}} = 0.005$, $OR_{\mathrm{interaction}} = 1.34$ (95% CI: 1.14–1.58) for JHH and $P_{\mathrm{interaction}} = 0.001$, $OR_{\mathrm{interaction}} = 1.53$ (95% CI = 1.18–1.99) in CGEMS, respectively] (Table II).

**Table I.** Reported SNPs associated with PCa

| Chr | SNPs | Cytogenetic bands | Position | Known genes | m/M allele | Risk allele |
|---|---|---|---|---|---|---|
| 2 | rs1465618 | 2p21 | 43,407,453 | *THADA* | A/G | A |
| 2 | rs721048 | 2p15 | 62,985,235 | *EHBP1* | A/G | A |
| 2 | rs12621278 | 2q31.1 | 173,019,799 | *ITGA6* | G/A | A |
| 3 | rs2660753 | 3p12 | 87,193,364 | — | T/C | T |
| 3 | rs10934853 | 3q21.3 | 129,521,063 | *EEFSEC* | A/C | A |
| 4 | rs17021918 | 4q22.3 | 95,781,900 | *PDLIM5* | T/C | C |
| 4 | rs7679673 | 4q24 | 106,280,983 | *TET2* | A/C | C |
| 6 | rs9364554 | 6q25 | 160,753,654 | *SLC22A3* | T/C | T |
| 7 | rs10486567 | 7p15 | 27,943,088 | *JAZF1* | A/G | G |
| 7 | rs6465657 | 7q21 | 97,654,263 | *LMTK2* | T/C | C |
| 8 | rs2928679 | 8p21.2 | 23,494,920 | *SLC25A37* | A/G | A |
| 8 | rs1512268 | 8p21.2 | 23,582,408 | *NKX3.1* | T/C | T |
| 8 | rs10086908 | 8q24 (5) | 128,081,119 | — | C/T | T |
| 8 | rs16901979 | 8q24 (2) | 128,194,098 | — | A/C | A |
| 8 | rs16902094 | 8q24.21 | 128,389,528 | — | N/A | G |
| 8 | rs620861 | 8q24 (4) | 128,404,855 | — | A/G | G |
| 8 | rs6983267 | 8q24 (3) | 128,482,487 | — | G/T | G |
| 8 | rs1447295 | 8q24 (1) | 128,554,220 | — | A/C | A |
| 9 | rs1571801 | 9q33 | 123,467,194 | *DAB2IC* | G/A | A |
| 10 | rs10993994 | 10q11 | 51,219,502 | *MSMB* | T/C | T |
| 10 | rs4962416 | 10q26 | 126,686,862 | *CTBP2* | C/T | C |
| 11 | rs7127900 | 11p15.5 | 2,190,150 | *IGF2, IGF2AS, INS, TH* | G/A | A |
| 11 | rs12418451 | 11q13 (2) | 68,691,995 | — | A/G | A |
| 11 | rs10896449 | 11q13 (1) | 68,751,243 | *MYEOV* | A/G | G |
| 17 | rs11649743 | 17q12 (2) | 33,149,092 | *HNF1B* | A/G | G |
| 17 | rs4430796 | 17q12 (1) | 33,172,153 | *HNF1B* | A/G | A |
| 17 | rs1859962 | 17q24.3 | 66,620,348 | — | G/T | G |
| 19 | rs8102476 | 19q13.2 | 43,427,453 | *PPP1R14A* | T/C | C |
| 19 | rs887391 | 19q13 | 46,677,464 | — | C/T | T |
| 19 | rs2735839 | 19q13 | 56,056,435 | *KLK3* | A/G | G |
| 22 | rs9623117 | New 22q13 | 38,782,065 | *TNRC6B* | C/T | C |
| 22 | rs5759167 | New 22q13.2 | 41,830,156 | *TTLL1, BIK, MCAT, PACSIN2* | T/G | G |
| 23 | rs5945619 | Xp11 | 51,258,412 | *NUDT10, NUDT11, LOC340602* | C/T | C |

Base pair position is based on NCBI build 36; Chr, chromosome; m/M, minor allele/major allele.

**Table II.** Results for top SNPs that interact with the known PCa risk SNPs ($P_{\text{interaction}} < 1.0 \times 10^{-5}$ in the meta-analysis and $P_{\text{interaction}} < 0.05$ in each of the three populations)

| Chr | SNP | Gene | Chr | SNP | Position | Minor allele | Gene | Location | Relative position | P | OR (95% CI) | P | OR (95% CI) | P | OR (95% CI) | P | OR (95% CI) |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 2 | rs1465618 | THADA | 4 | rs11735008 | 393,303 | G | ABCA11P | Intergenic | 15,921 | $6.65 \times 10^{-6}$ | 0.76 (0.68–0.86) | $5.04 \times 10^{-3}$ | 0.69 (0.54–0.90) | $3.11 \times 10^{-2}$ | 0.83 (0.71–0.98) | $1.35 \times 10^{-3}$ | 0.69 (0.55–0.86) |
|  |  |  | 13 | rs9567349 | 43,535,405 | G | NCRNA00284 | Intergenic | 32,806 | $3.93 \times 10^{-6}$ | 0.61 (0.49–0.75) | $4.57 \times 10^{-4}$ | 0.41 (0.25–0.67) | $3.24 \times 10^{-2}$ | 0.72 (0.54–0.97) | $3.97 \times 10^{-3}$ | 0.57 (0.39–0.84) |
| 3 | rs10934853 | EEFSEC | 9 | rs7847271 | 116,870,633 | A | TNC | Intron |  | $3.85 \times 10^{-6}$ | 0.67 (0.56–0.79) | $7.73 \times 10^{-3}$ | 0.60 (0.41–0.87) | $2.56 \times 10^{-3}$ | 0.68 (0.54–0.88) | $1.87 \times 10^{-2}$ | 0.69 (0.50–0.94) |
|  |  |  | 18 | rs998124 | 40,979,660 | G | MIR4319 | Intergenic | 175,531 | $5.21 \times 10^{-6}$ | 1.33 (1.18–1.51) | $3.56 \times 10^{-2}$ | 1.39 (1.02–1.88) | $3.64 \times 10^{-3}$ | 1.28 (1.08–1.51) | $3.56 \times 10^{-2}$ | 1.42 (1.12–1.80) |
| 4 | rs17021918 | PDLIM5 | 3 | rs9757252 | 86,977,168 | T | VGLL3 | Intergenic | 92,645 | $4.73 \times 10^{-6}$ | 1.25 (1.13–1.37) | $8.45 \times 10^{-3}$ | 1.35 (1.08–1.69) | $2.23 \times 10^{-3}$ | 1.22 (1.07–1.38) | $2.16 \times 10^{-2}$ | 1.24 (1.03–1.50) |
| 4 | rs7679673 | TET2 | 9 | rs290258 | 92,595,560 | G | SYK | Intergenic | −8,273 | $1.49 \times 10^{-6}$ | 0.75 (0.67–0.84) | $2.11 \times 10^{-3}$ | 0.66 (0.51–0.86) | $3.01 \times 10^{-3}$ | 0.78 (0.67–0.92) | $1.39 \times 10^{-2}$ | 0.75 (0.59–0.94) |
|  |  |  | 22 | rs5751168 | 21,175,240 | T | ZNF280B | Intron |  | $4.11 \times 10^{-6}$ | 1.44 (1.23–1.67) | $4.75 \times 10^{-5}$ | 2.19 (1.50–3.19) | $3.38E \times 10^{-2}$ | 1.25 (1.02–1.53) | $9.09 \times 10^{-3}$ | 1.48 (1.10–1.99) |
| 7 | rs10486567 | JAZF1 | 3 | rs1795355 | 41,574,530 | T | ULK4 | Intron |  | $9.46 \times 10^{-6}$ | 0.79 (0.71–0.88) | $3.37 \times 10^{-2}$ | 0.77 (0.60–0.98) | $9.11 \times 10^{-3}$ | 0.83 (0.72–0.95) | $2.14 \times 10^{-3}$ | 0.73 (0.60–0.89) |
|  |  |  | 3 | rs11720607 | 174,325,971 | G | SPATA16 | Intron |  | $4.87 \times 10^{-6}$ | 0.73 (0.63–0.83) | $2.45 \times 10^{-3}$ | 0.62 (0.45–0.84) | $2.34 \times 10^{-3}$ | 0.75 (0.62–0.90) | $5.43 \times 10^{-2}$ | 0.77 (0.58–1.00) |
| 7 | rs6465657 | LMTK2 | 16 | rs8057939 | 47,951,777 | C | C16orf78 | Intergenic | −13,532 | $4.71 \times 10^{-6}$ | 1.37 (1.20–1.57) | $3.31 \times 10^{-2}$ | 1.43 (1.03–1.98) | $1.03 \times 10^{-3}$ | 1.36 (1.13–1.63) | $1.70 \times 10^{-2}$ | 1.36 (1.06–1.75) |
| 8 | rs10086908 |  | 6 | rs10456809 | 17,921,804 | T | KIF13A | Intron |  | $4.83 \times 10^{-6}$ | 1.25 (1.14–1.38) | $1.30 \times 10^{-2}$ | 1.31 (1.06–1.62) | $2.28 \times 10^{-3}$ | 1.23 (1.08–1.40) | $1.69 \times 10^{-2}$ | 1.26 (1.04–1.52) |
| 8 | rs1447295 | NA | 7 | rs7789197 | 40,931,652 | A | INHBA | Intergenic | 763,474 | $3.36 \times 10^{-6}$ | 0.66 (0.56–0.79) | $2.57 \times 10^{-3}$ | 0.55 (0.38–0.81) | $7.27 \times 10^{-3}$ | 0.72 (0.56–0.91) | $1.06 \times 10^{-2}$ | 0.66 (0.48–0.91) |
|  |  |  | 9 | rs12682851 | 8,002,418 | G | C9orf123 | Intergenic | 212,619 | $1.53 \times 10^{-6}$ | 0.72 (0.63–0.82) | $9.66 \times 10^{-3}$ | 0.67 (0.50–0.91) | $2.15 \times 10^{-3}$ | 0.75 (0.62–0.90) | $6.56 \times 10^{-3}$ | 0.70 (0.54–0.90) |
|  |  |  | 10 | rs10885582 | 116,317,540 | T | ABLIM1 | Intron |  | $3.70 \times 10^{-6}$ | 0.73 (0.63–0.83) | $9.33 \times 10^{-5}$ | 0.54 (0.39–0.73) | $1.27 \times 10^{-2}$ | 0.79 (0.66–0.95) | $3.46 \times 10^{-2}$ | 0.75 (0.57–0.98) |
| 8 | rs6983267 | NA | 6 | rs1011119 | 19,972,144 | G | ID4 | Intergenic | 23,250 | $7.20 \times 10^{-6}$ | 0.81 (0.74–0.89) | $1.45 \times 10^{-2}$ | 0.76 (0.61–0.95) | $2.85 \times 10^{-2}$ | 0.83 (0.74–0.94) | $1.58 \times 10^{-2}$ | 0.80 (0.67–0.96) |
| 9 | rs1571801 | DAB2IC | 8 | rs2219968 | 79,119,213 | A | PKIA | Intergenic | 471,678 | $6.07 \times 10^{-7}$ | 1.30 (1.17–1.43) | $1.18 \times 10^{-2}$ | 1.35 (1.07–1.70) | $1.64 \times 10^{-4}$ | 1.30 (1.14–1.50) | $3.17 \times 10^{-2}$ | 1.24 (1.02–1.52) |
|  |  |  | 8 | rs13264970 | 83,236,384 | C | SNX16 | Intergenic | 319,308 | $3.53 \times 10^{-6}$ | 0.77 (0.69–0.86) | $1.50 \times 10^{-2}$ | 0.74 (0.59–0.94) | $5.02 \times 10^{-3}$ | 0.80 (0.68–0.93) | $3.92 \times 10^{-3}$ | 0.72 (0.58–0.90) |
|  |  |  | 10 | rs1547851 | 92,364,806 | T | HTR7 | Intergenic | 125,750 | $7.45 \times 10^{-6}$ | 1.59 (1.30–1.95) | $2.72 \times 10^{-3}$ | 1.98 (1.27–3.09) | $7.35 \times 10^{-3}$ | 1.49 (1.11–1.99) | $2.42 \times 10^{-2}$ | 1.52 (1.06–2.20) |
| 10 | rs4962416 | CTBP2 | 5 | rs10940579 | 57,166,575 | C | ACTBL2 | Intergenic | 352,182 | $3.81 \times 10^{-6}$ | 1.32 (1.18–1.49) | $4.85 \times 10^{-3}$ | 1.36 (1.00–1.84) | $1.77 \times 10^{-3}$ | 1.29 (1.10–1.51) | $4.74 \times 10^{-3}$ | 1.39 (1.10–1.74) |
| 11 | rs10896449 | MYEOV | 12 | rs17354197 | 88,169,501 | T | DUSP6 | Intergenic | 96,467 | $8.82 \times 10^{-6}$ | 1.41 (1.21–1.64) | $3.97 \times 10^{-2}$ | 1.49 (1.02–2.18) | $2.35 \times 10^{-3}$ | 1.37 (1.12–1.67) | $1.10 \times 10^{-2}$ | 1.45 (1.09–1.92) |
| 11 | rs12418451 | MYEOV | 3 | rs10513723 | 176,062,702 | A | NAALADL2 | Intron |  | $5.61 \times 10^{-6}$ | 1.41 (1.21–1.63) | $7.22 \times 10^{-3}$ | 1.58 (1.13–2.21) | $1.61 \times 10^{-2}$ | 1.27 (1.05–1.54) | $1.46 \times 10^{-3}$ | 1.64 (1.21–2.22) |
|  |  |  | 8 | rs7829048 | 4,689,690 | C | CSMD1 | Intron |  | $9.76 \times 10^{-6}$ | 0.74 (0.65–0.85) | $1.22 \times 10^{-2}$ | 0.68 (0.50–0.92) | $4.52 \times 10^{-2}$ | 0.84 (0.71–1.00) | $1.81 \times 10^{-4}$ | 0.60 (0.46–0.78) |
|  |  |  | 15 | rs784411 | 46,827,089 | C | CEP152 | Intron |  | $1.15 \times 10^{-7}$ | 1.42 (1.25–1.61) | $8.83 \times 10^{-3}$ | 1.55 (1.12–2.16) | $5.28 \times 10^{-4}$ | 1.34 (1.14–1.58) | $1.32 \times 10^{-3}$ | 1.53 (1.18–1.99) |
| 11 | rs7127900 | IGF2, IGF2AS | 8 | rs13258681 | 124,783,903 | C | ANXA13 | Intron |  | $3.65 \times 10^{-6}$ | 1.32 (1.17–1.48) | $4.58 \times 10^{-2}$ | 1.32 (1.01–1.73) | $5.43 \times 10^{-4}$ | 1.33 (1.13–1.56) | $1.90 \times 10^{-2}$ | 1.31 (1.05–1.64) |
|  |  | INS, TH | 22 | rs12628051 | 38,984,222 | C | TNRC6B | Intron |  | $3.39 \times 10^{-6}$ | 1.30 (1.17–1.46) | $1.82 \times 10^{-3}$ | 1.50 (1.16–1.93) | $6.14 \times 10^{-3}$ | 1.24 (1.06–1.44) | $1.44 \times 10^{-2}$ | 1.32 (1.06–1.65) |
| 17 | rs4430796 | HNF1B | 1 | rs731174 | 37,969,428 | C | EPHA10 | Intron |  | $4.55 \times 10^{-6}$ | 1.27 (1.15–1.40) | $5.03 \times 10^{-2}$ | 1.24 (1.00–1.54) | $3.20 \times 10^{-2}$ | 1.18 (1.01–1.38) | $1.13 \times 10^{-4}$ | 1.41 (1.19–1.68) |
|  |  |  | 2 | rs12694942 | 158,518,681 | T | UPP2 | Intergenic | 41,256 | $7.84 \times 10^{-6}$ | 0.80 (0.73–0.88) | $5.22 \times 10^{-3}$ | 0.75 (0.61–0.92) | $1.87 \times 10^{-2}$ | 0.79 (0.68–0.92) | $6.36 \times 10^{-3}$ | 0.85 (0.72–1.01) |
|  |  |  | 9 | rs10812303 | 25,712,117 | T | TUSC1 | Intergenic | 43,261 | $5.59 \times 10^{-6}$ | 1.34 (1.18–1.52) | $6.63 \times 10^{-3}$ | 1.46 (1.11–1.91) | $7.57 \times 10^{-3}$ | 1.27 (1.07–1.52) | $9.01 \times 10^{-3}$ | 1.37 (1.08–1.73) |
| 17 | rs1859962 | NA | 2 | rs16867225 | 180,749,531 | A | CWC22 | Intergenic | 169,506 | $3.12 \times 10^{-6}$ | 0.64 (0.53–0.77) | $1.33 \times 10^{-3}$ | 0.48 (0.30–0.75) | $4.80 \times 10^{-3}$ | 0.70 (0.54–0.90) | $1.84 \times 10^{-2}$ | 0.65 (0.45–0.93) |
|  |  |  | 7 | rs10277209 | 108,790,810 | C | C7orf66 | Intergenic | 478,937 | $3.81 \times 10^{-6}$ | 1.36 (1.19–1.55) | $4.08 \times 10^{-2}$ | 1.39 (1.01–1.89) | $4.04 \times 10^{-3}$ | 1.29 (1.08–1.53) | $1.52 \times 10^{-2}$ | 1.52 (1.17–1.97) |
| 19 | rs2735839 | KLK3 | 20 | rs6089829 | 61,139,481 | A | LOC63930 | Intron |  | $3.21 \times 10^{-6}$ | 0.74 (0.65–0.84) | $2.53 \times 10^{-2}$ | 0.71 (0.52–0.96) | $1.26 \times 10^{-2}$ | 0.75 (0.64–0.90) | $1.10 \times 10^{-2}$ | 0.72 (0.56–0.93) |
| 19 | rs8102476 | PPP1R14A | 1 | rs1866967 | 29,958,249 | G | PTPRU | Intergenic | 432,337 | $5.16 \times 10^{-6}$ | 0.82 (0.75–0.89) | $2.70 \times 10^{-2}$ | 0.79 (0.64–0.97) | $5.10 \times 10^{-3}$ | 0.85 (0.75–0.95) | $2.97 \times 10^{-3}$ | 0.78 (0.66–0.92) |
| 19 | rs887391 | NA | 4 | rs735172 | 5,809,770 | C | EVC | Intron |  | $2.03 \times 10^{-6}$ | 1.31 (1.17–1.46) | $6.58 \times 10^{-3}$ | 1.43 (1.10–1.85) | $2.37 \times 10^{-3}$ | 1.27 (1.09–1.47) | $1.02 \times 10^{-2}$ | 1.32 (1.07–1.63) |
|  |  |  | 5 | rs4463179 | 13,558,432 | A | DNAH5 | Intergenic | 185,005 | $2.22 \times 10^{-6}$ | 0.64 (0.53–0.77) | $3.00 \times 10^{-2}$ | 0.64 (0.42–0.96) | $8.35 \times 10^{-5}$ | 0.59 (0.46–0.77) | $8.28 \times 10^{-2}$ | 0.73 (0.51–1.04) |
| 22 | rs9623117 | TNRC6B | 4 | rs1713511 | 43,472,127 | A | KCTD8 | Intergenic | 398,550 | $7.87 \times 10^{-6}$ | 1.31 (1.16–1.47) | $4.94 \times 10^{-3}$ | 1.54 (1.14–2.09) | $6.71 \times 10^{-3}$ | 1.23 (1.06–1.44) | $1.06 \times 10^{-2}$ | 1.36 (1.07–1.72) |

SNP1 indicate the 32 known PCa-risk SNPs. SNP 2 indicates the interacting SNPs; Chr, chromosome; Position is based on NVBI Build 36: relative position is the distance of SNP2 relative to the nearest gene if SNP2 is located in the intergenic region; P and OR are for the multiplicative interaction term. P and OR for the meta-analysis are calculated based on a Cochran–Mantel–Haenszel test.

Among the other 34 pairs of interactions that were significant at a $P_{\text{interaction}}$ cutoff of $1 \times 10^{-5}$ in the meta-analysis, two pairs were noteworthy to be emphasized when considering possible biological function. The first pair involved an interaction between rs7127900 at insulin-like growth factor (*IGF2*)/*IGF2AS* region and rs12628051 in the intron of *TNRC6B*, with a $P_{\text{interaction}}$ of $3.39 \times 10^{-6}$ ($\text{OR}_{\text{interaction}}$ = 1.30; 95% CI = 1.17–1.46) (Table II). The interaction was significant in all three populations and the effects of the interaction were in the same direction ($P_{\text{interaction}}$ = 0.002, $\text{OR}_{\text{interaction}}$ = 1.50, 95% CI = 1.16–1.93 in CAPS; $P_{\text{interaction}}$ = 0.006, $\text{OR}_{\text{interaction}}$ = 1.24, 95% CI = 1.06–1.44 in JHH and $P_{\text{interaction}}$ = 0.014, $\text{OR}_{\text{interaction}}$ = 1.32, 95% CI = 1.06–1.65 in CGEMS). The second pair of interaction was between rs7679763 in *TET2* gene region and rs290258 in the promoter region of SYK, with a $P_{\text{interaction}}$ of $1.49 \times 10^{-6}$ (OR = 0.75; 95% CI = 0.67–0.84) (Table II). Similarly, the interaction effect was consistently observed in all three populations with the same direction of interaction effect ($P_{\text{interaction}}$ = 0.002, $\text{OR}_{\text{interaction}}$ = 0.66, 95% CI = 0.51–0.86 in CAPS; $P_{\text{interaction}}$ = 0.003, $\text{OR}_{\text{interaction}}$ = 0.78, 95% CI = 0.67–0.92 in JHH and $P_{\text{interaction}}$ = 0.014, $\text{OR}_{\text{interaction}}$ = 0.75, 95% CI = 0.59–0.94 in CGEMS).

We then carefully examined the significant pairs of SNP–SNP interactions reported by Ciampa *et al.* (24) in CAPS and JHH population. Among the 25 pairs reported in the previous study, 16 pairs were also evaluated in our data. Three pairs of SNP–SNP interaction reached nominal $P_{\text{interaction}}$ of 0.05 in CAPS population (Table III). The most significant interaction replicated in CAPS was between rs6983267 and rs4953347 ($P_{\text{interaction}}$ = 0.001, $\text{OR}_{\text{interaction}}$ = 1.42). However, this interaction was not significant in the JHH population ($P$ = 0.69). The other two pairs of SNPs were replicated in CAPS at a nominal $P_{\text{interaction}}$ of 0.05, including the interaction between rs2735839 and rs12196677 ($P_{\text{interaction}}$ = 0.017, same direction of interaction effect) and the interaction between rs10934853 and rs10458466 ($P_{\text{interaction}}$ = 0.02 but with opposite direction of interaction effect). The interaction effect of the two pairs of SNP–SNP interactions that were significant in JHH population at a $P_{\text{interaction}}$ cutoff of 0.05 were in the opposite direction compared with the previous study (24) (Table III).

## Discussion

To our knowledge, our study represents one of the first comprehensive gene–gene interaction scans in three PCa GWAS populations. Specifically, we performed a genome-wide gene–gene interaction scan for each of the 32 known PCa risk-associated variants identified from GWAS in three case–control populations of European descents, which includes a total of 4723 PCa cases and 4792 controls. In the meta-analysis, we found 35 pairs of SNP–SNP interactions that were significantly associated with PCa risk ($P_{\text{interaction}} < 1 \times 10^{-5}$). In addition, the interactions for those 35 pairs were significant in all three populations (all $P_{\text{interaction}} < 0.05$). Among those 35 pairs of statistically significant interactions, we emphasized three pairs of interactions with potential biological implication, including an interaction between rs12418451 in *MYEOV* and rs16961635 in *CEP152*, with a $P_{\text{interaction}}$ of $1.15 \times 10^{-7}$ (OR = 1.42, 95% CI = 1.25–1.61), an interaction between rs7127900 at *IGF2/IGF2AS* region and rs12628051 in the intron of *TNRC6B*, with a $P_{\text{interaction}}$ of $3.39 \times 10^{-6}$ (OR = 1.30, 95% CI = 1.17–1.46) and an interaction between rs7679763 in *TET2* gene region and rs290258 in the promoter region of SYK, with a $P_{\text{interaction}}$ of $1.49 \times 10^{-6}$ (OR = 0.75, 95% CI = 0.67–0.84).

The discovery of approximately three dozen PCa risk variants using single SNP analysis suggests that it is possible to detect individual risk variants. However, when the underlying genetic model involves interaction of multiple genes, a single gene approach is less effective and may not be able to explain the complex etiology of the disease. Therefore, evaluation of the joint effect (epistasis) of multiple genetic variants is critical to understand the underlying causes of complex diseases (30), especially in the situation where several individual risk variants have been identified. The next question is to explore whether other SNPs interact with those SNPs to modify risk to PCa. The identified loci that interact with the known PCa risk-associated SNPs may help to elucidate the underlying molecular mechanisms of the associations of those risk SNPs.

The most significant interaction was seen between the PCa risk-associated SNPs rs12418451 and rs784411. The SNP rs12418451 is located at the 11q13.2 that is ~77 kb upstream of *TPCN2*, a putative cation-selective ion channel gene and ~126 kb upstream of *MYEOV*, an oncogene that has been implicated in multiple cancers (31–35). The SNP rs784411 resides in the intron of CEP152, a centrosomal protein that was recently shown to function as a regulator of genomic integrity (36) and cellular response to DNA damage (37). Given the limited information, we speculate that observed interaction may reflect the close collaboration of MYEOV (or TPCN2, even though it is less likely) and CEP152 in the same or different oncogenic pathways that drive the tumorigenesis of prostatic epithelial cells.

Among the two SNPs that were shown to consistently interact with the PCa risk-associated SNP rs7127900 at 11p15.5, one SNP (rs12628051) is located within TNRC6B, which encodes an RNA

**Table III.** Results for the top pairs of SNP–SNP interactions reported by the previous study (24) in JHH and CAPS population

| Gene | Conditioning SNP | Interacting SNP | Chr | Gene | CAPS | | | JHH | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | | | | | OR[a] | 95% CI | $P$[a] | OR[a] | 95% CI | $P$[a] |
| *DAB2IP* | rs1571801 | rs4242 | 22 | | 1.14 | 0.63–2.05 | 0.662 | 0.80 | 0.7–0.93 | **0.003** |
| *EEFSEC* | rs10934853 | rs12489404 | 3 | *GRM7* | 0.99 | 0.76–1.29 | 0.941 | 1.18 | 1.03–1.34 | **0.016** |
| | rs10934853 | rs10458466 | 1 | | 0.77 | 0.62–0.96 | **0.020** | 0.93 | 0.81–1.06 | 0.259 |
| *EHBP1* | rs721048 | rs10514124 | 5 | | 0.98 | 0.66–1.46 | 0.921 | 0.91 | 0.78–1.05 | 0.194 |
| | rs721048 | rs704638 | 4 | | 1.12 | 0.83–1.52 | 0.462 | 1.10 | 0.94–1.29 | 0.240 |
| | rs721048 | rs7604809 | 2 | *GTF3C3* | 0.91 | 0.60–1.38 | 0.656 | 0.92 | 0.75–1.14 | 0.455 |
| *JAZF1* | rs10486567 | rs2899748 | 15 | *GLCE* | 0.99 | 0.73–1.34 | 0.948 | 0.96 | 0.83–1.11 | 0.573 |
| *KLK2-KLK3* | rs2735839 | rs12196677 | 6 | *PNPLA1* | 3.90 | 1.27–11.95 | **0.017** | 1.06 | 0.83–1.36 | 0.616 |
| *MSMB* | rs10993994 | rs12605415 | 18 | | 1.07 | 0.87–1.32 | 0.525 | 1.08 | 0.96–1.22 | 0.179 |
| | rs10993994 | rs11083271 | 18 | | 1.12 | 0.89–1.41 | 0.329 | 1.04 | 0.93–1.18 | 0.486 |
| | rs10993994 | rs9880831 | 3 | *LOC391524* | 1.00 | 1.00–1.00 | 0.975 | 0.95 | 0.84–1.07 | 0.372 |
| | rs10993994 | rs7921651 | 10 | *FAM107B* | 1.18 | 0.93–1.49 | 0.164 | 1.07 | 0.93–1.22 | 0.365 |
| *MYEOV* | rs10896449 | rs1240224 | 12 | | 1.08 | 0.86–1.36 | 0.506 | 1.01 | 0.89–1.15 | 0.851 |
| *8q24* | rs16901979 | rs991000 | 2 | | 1.14 | 0.64–2.05 | 0.661 | 1.00 | 0.70–1.43 | 0.997 |
| | rs16901979 | rs943323 | 14 | *KIAA1409* | 0.87 | 0.50–1.52 | 0.623 | 0.92 | 0.65–1.30 | 0.628 |
| *8q24* | rs6983267 | rs4953347 | 2 | *EPAS1* | 1.42 | 1.15–1.76 | **0.001** | 0.98 | 0.87–1.10 | 0.688 |

Bold indicates P- values with a nominal significant level of 0.05. Chr, chromosome.
[a]*P* values, OR and 95% CI are calculated for the multiplicative interaction term.

interference (RNAi) machinery component protein crucial for the microRNA/small interfering RNA-dependent translational repression or degradation of target messenger RNAs. It is worthy to mention that this gene also contains a GWAS-identified PCa risk-associated SNP (rs9623117). Several mechanisms may potentially explain for these interactions. Firstly, we noticed that at ~70 kb telomeric to rs7127900 reside the PCa-implicated *IGF2* gene and its antisense transcript-encoding *IGF2AS*. IGF2 encodes a member of the insulin family of polypeptide growth factors that promotes cell proliferation during fetal development but becomes less active in healthy adults due to genomic imprinting. Dysregulated overexpression of IGF2 caused by loss of imprinting has been associated with a variety of human cancers, including PCa (38–41). IGF2AS encodes a predictably non-coding RNA that is antisense to IGF2 and thus may potentially regulate IGF2 expression through RNAi in a similar manner as some other natural antisense transcripts. Thus, one plausible scenario is that TNRC6B may affect the RNAi-mediated transcriptional regulation of IGF2AS on IGF2, which may underlie the observed interaction between genetic variants within these two loci. Secondly, there are two microRNA genes located at 11p15.5, miR-4686 (~40 kb from the PCa-risk SNP rs7127900) and miR-483 (~80 kb from rs7127900). Although the role of miR-4686 remains to be determined, miR-483 has been demonstrated to act as an oncogene to suppress proapoptotic BBC3 (PUMA) or tumor suppressive DPC4 (Smad4) in a variety of human cancers (42,43). Thus, an alternative mechanism for the observed interaction between the 11p15.5 locus and the TNRC6B locus is that genetic variants in TNRC6B may affect the miR-483 (or miR-4686)-mediated RNAi toward its/their target tumor suppressor genes.

Another pair of interacting SNPs was found between rs7679673 (~6 kb upstream of *TET2*) and rs290258 (~8 kb upstream of *SYK*). TET2 encodes an enzyme hydroxylating methylcytosine and is implicated in epigenetic programming that involves DNA methylation and demethylation (reviewed in ref. 44). The critical role of TET2 in cancer is suggested by the observation that loss-of-function mutations of TET2 are frequently identified in various hematological malignancies (45,46). As a non-receptor tyrosine protein kinase that mediates cellular proliferation and differentiation, SYK is believed to function as a potential tumor suppressive gene (reviewed in ref. 47). It is noteworthy that hypermethylation of SYK gene promoter has been frequently found in and widely associated with lung, gastric and breast cancer (48,49). Thus, although it remains to be determined whether SYK promoter in prostatic tumors also undergoes silencing via DNA methylation, the observed interaction between TET2 and SYK suggests that it is a plausible hypothesis.

Two SNPs (rs731174 and rs10812303) were found to interact with the GWAS-identified PCa risk-associated SNP rs4430796, residing within *HNF1B*, a homeodomain-containing transcription factor whose expression alteration has been widely implicated in various human cancers, including PCa. The SNP rs731174 is located within the intron of *EPHA10*, a member of the EPH subfamily of receptor tyrosine kinases. This family of receptor tyrosine kinases play an important role in cell–cell communication regulating cell attachment, shape and mobility in epithelial cells and are believed to be implicated in carcinogenesis (reviewed in ref. 50). It is possible that HNF1Ba and EPHA10 collaborate in the signaling network that is crucial for the well-being of prostatic cells, whereas the genetic variants located within these two genes may synergistically contribute to the oncogenesis of PCa. The other SNP rs10812303 is ~40 kb upstream of *TUSC1*, an intronless gene that has been suggested to serve as a tumor suppressor in lung tumorigenesis (51). Thus, the interaction between genetic variants in TUSC1 and HNF1B may also suggest a plausible collaboration of these two genes.

Besides our novel findings, we also replicated the most notable finding reported by the study of Ciampa *et al.* (24) in our CAPS population. The interaction involves 8q24 region 3 (rs6983267) and *EPAS1* (rs4953347). We also observed the interaction between the *KLK2–KLK3* (rs2735839) and *PNPLA1* in the CAPS population. However, none of those two pairs of interaction were significant in the JHH dataset. Therefore, further statistical evidence supports from additional replication studies are needed to reach a more robust conclusion.

Our results need to be interpreted with caution. The most significant SNP–SNP interaction pair detected in the meta-analysis had a $P_{\text{interaction}}$ of $1.15 \times 10^{-7}$. It did not reach a genome-wide significance level of $1.5 \times 10^{-9}$ as we performed multiple tests in this study (1 million × 32 = 32 million tests). One possible reason for not achieving genome-wide significance may be due to limited statistical power to detect small to modest interaction effects. In the meta-analysis of 4723 PCa cases and 4759 controls, we had 80% power to detect relatively large interaction effects (OR > 1.7), using the stringent Bonferonni-corrected $P$-value cutoff of $1.5 \times 10^{-9}$ to claim a genome-wide significant level. Therefore, additional samples were needed to detect modest interaction effects at a genome-wide significant level, which was the case in our study (OR approximately ranged from 1.2 to 1.5). However, the interaction effect detected in our study was consistently implicated in all three populations and with the same direction of interaction effects. In addition, similar pattern of integrations were observed (quantitative interactions) among the three populations. This may represent statistically meaningful SNP–SNP interactions with a modest magnitude of interaction effect. However, the pairs of SNP–SNP interactions identified by our study still warrant follow-up in other populations to further exclude the possibilities of false positive findings. More importantly, functional studies are needed to better understand the underlying molecular mechanisms of the interactions implicated.

The limitations of our study include the use of a pairwise gene–gene interaction approach and limiting the search of interactions on SNPs that confer main effects only. Firstly, the search for gene–gene interactions based on a pairwise approach leads to reduced or no power when high-order interactions are present. However, the search for high-order interactions in a GWAS scale still represents a computational challenge and warrants novel statistical approaches to handle this difficult task. Secondly, we did not evaluate gene–gene interactions among SNPs that didn't confer a main effect. We can't exclude the possibilities that the epistasis between other pairs of SNPs without strong main effect may also affect risk of PCa. However, the exhaustive search for gene–gene interaction is beyond the scope of the current study as we focused on identifying genes that interact with known risk-associated SNPs. We would like to continue our research on this field in future studies. In addition, the current analysis of the imputed data was based on the most likely genotype with the posterior probability of ≥0.90. Alternatively, using the imputed dosage data would be more accurate and may improve the power of this study. Future epistasis analysis based on the imputed dosage data would be explored once the statistical and computation methods are available.

In summary, our systematic evaluation of gene–gene interactions in three GWAS populations suggested a list of loci interacting with known PCa risk-associated SNPs that may warrant follow-up in other study populations. Three pairs of interactions are worthwhile to be emphasized, including an interaction between rs12418451 in the *MYEOV* gene region and rs784411 in the intron of *CEP152*, an interaction between rs7127900 in the *IGF2/IGF2AS* gene region and rs12628051 in the intron of *TNRC6B* and an interaction between rs7679673 in the *TET2* gene region and rs290258 in the intron of *SYK*. Those results showed statistical evidence for genes interacting with known risk-associated SNPs on PCa risk. The interacting loci identified also provide more hints on the underlying molecular mechanism of the associations with PCa risk for the known risk-associated SNPs. Additional studies are warranted to further confirm the gene–gene interaction effects detected in this study.

## Supplementary material

Supplementary Tables 1 and 2 can be found at http://carcin.oxfordjournals.org/.

## Funding

## Acknowledgements

## References

1. Amundadottir,L.T. *et al*. (2006) A common variant associated with prostate cancer in European and African populations. *Nat. Genet.*, **38**, 652–658.
2. Gudmundsson,J. *et al*. (2007) Genome-wide association study identifies a second prostate cancer susceptibility variant at 8q24. *Nat. Genet.*, **39**, 631–637.
3. Yeager,M. *et al*. (2007) Genome-wide association study of prostate cancer identifies a second risk locus at 8q24. *Nat. Genet.*, **39**, 645–649.
4. Gudmundsson,J. *et al*. (2007) Two variants on chromosome 17 confer prostate cancer risk, and the one in TCF2 protects against type 2 diabetes. *Nat. Genet.*, **39**, 977–983.
5. Duggan,D. *et al*. (2007) Two genome-wide association studies of aggressive prostate cancer implicate putative prostate tumor suppressor gene DAB2IP. *J. Natl Cancer Inst.*, **99**, 1836–1844.
6. Thomas,G. *et al*. (2008) Multiple loci identified in a genome-wide association study of prostate cancer. *Nat. Genet.*, **40**, 310–315.
7. Gudmundsson,J. *et al*. (2008) Common sequence variants on 2p15 and Xp11.22 confer susceptibility to prostate cancer. *Nat. Genet.*, **40**, 281–283.
8. Eeles,R.A. *et al*. (2008) Multiple newly identified loci associated with prostate cancer susceptibility. *Nat. Genet.*, **40**, 316–321.
9. Yeager,M. *et al*. (2009) Identification of a new prostate cancer susceptibility locus on chromosome 8q24. *Nat. Genet.*, **41**, 1055–1057.
10. Gudmundsson,J. *et al*. (2009) Genome-wide association and replication studies identify four variants associated with prostate cancer susceptibility. *Nat. Genet.*, **41**, 1122–1126.
11. Eeles,R.A. *et al*. (2009) Identification of seven new prostate cancer susceptibility loci through a genome-wide association study. *Nat. Genet.*, **41**, 1116–1121.
12. Sun,J. *et al*. (2008) Association between sequence variants at 17q12 and 17q24.3 and prostate cancer risk in European and African Americans. *Prostate*, **68**, 691–697.
13. Hsu,F.C. *et al*. (2009) A novel prostate cancer susceptibility locus at 19q13. *Cancer Res.*, **69**, 2720–2723.
14. Sun,J. *et al*. (2009) Sequence variants at 22q13 are associated with prostate cancer risk. *Cancer Res.*, **69**, 10–15.
15. Zheng,S.L. *et al*. (2009) Two independent prostate cancer risk-associated loci at 11q13. *Cancer Epidemiol. Biomarkers Prev.*, **18**, 1815–1820.
16. Kote-Jarai,Z. *et al*. (2008) Multiple novel prostate cancer predisposition loci confirmed by an international study: the PRACTICAL Consortium. *Cancer Epidemiol. Biomarkers Prev.*, **17**, 2052–2061.
17. Chang,B.L. *et al*. (2011) Validation of genome-wide prostate cancer associations in men of African descent. *Cancer Epidemiol. Biomarkers Prev.*, **20**, 23–32.
18. Zheng,S.L. *et al*. (2010) Association of 17 prostate cancer susceptibility loci with prostate cancer risk in Chinese men. *Prostate*, **70**, 425–432.
19. Yamada,H. *et al*. (2009) Replication of prostate cancer risk loci in a Japanese case-control association study. *J. Natl Cancer Inst.*, **101**, 1330–1336.
20. Lindstrom,S. *et al*. (2011) Characterizing associations and SNP-environment interactions for GWAS-identified prostate cancer risk markers—results from BPC3. *PLoS One*, **6**, e17142.
21. Mackay,T.F. (2001) Quantitative trait loci in Drosophila. *Nat. Rev. Genet.*, **2**, 11–20.
22. Williams,S.M. *et al*. (2004) The use of animal models in the study of complex disease: all else is never equal or why do so many human studies fail to replicate animal findings? *Bioessays*, **26**, 170–179.
23. Routman,E.J.C. *et al*. (1997) Gene effects on a quantitative trait: two-locus epistatic effects measured at microsatellite markers and at estimated QTL. *Evolution*, **51**, 1654–1662.
24. Ciampa,J. *et al*. (2011) Large-scale exploration of gene-gene interactions in prostate cancer using a multistage genome-wide association study. *Cancer Res.*, **71**, 3287–3295.
25. Xu,J. *et al*. (2010) Inherited genetic variant predisposes to aggressive but not indolent prostate cancer. *Proc. Natl Acad. Sci. USA*, **107**, 2136–2140.
26. Marchini,J. *et al*. (2007) A new multipoint method for genome-wide association studies by imputation of genotypes. *Nat. Genet.*, **39**, 906–913.
27. Purcell,S. *et al*. (2007) PLINK: a tool set for whole-genome association and population-based linkage analyses. *Am. J. Hum. Genet.*, **81**, 559–575.
28. Price,A.L. *et al*. (2006) Principal components analysis corrects for stratification in genome-wide association studies. *Nat. Genet.*, **38**, 904–909.
29. Manning,A.K. *et al*. (2011) Meta-analysis of gene-environment interaction: joint estimation of SNP and SNP x environment regression coefficients. *Genet. Epidemiol.*, **35**, 11–18.
30. Cordell,H.J. (2009) Detecting gene-gene interactions that underlie human diseases. *Nat. Rev. Genet.*, **10**, 392–404.
31. Janssen,J.W. *et al*. (2000) Concurrent activation of a novel putative transforming gene, myeov, and cyclin D1 in a subset of multiple myeloma cell lines with t(11;14)(q13;q32). *Blood*, **95**, 2691–2698.
32. Janssen,J.W. *et al*. (2002) MYEOV, a gene at 11q13, is coamplified with CCND1, but epigenetically inactivated in a subset of esophageal squamous cell carcinomas. *J. Hum. Genet.*, **47**, 460–464.
33. Janssen,J.W. *et al*. (2002) MYEOV: a candidate gene for DNA amplification events occurring centromeric to CCND1 in breast cancer. *Int. J. Cancer*, **102**, 608–614.
34. Specht,K. *et al*. (2004) Different mechanisms of cyclin D1 overexpression in multiple myeloma revealed by fluorescence *in situ* hybridization and quantitative analysis of mRNA levels. *Blood*, **104**, 1120–1126.
35. Moss,A.C. *et al*. (2006) ETV4 and Myeov knockdown impairs colon cancer cell line proliferation and invasion. *Biochem. Biophys. Res. Commun.*, **345**, 216–221.
36. Hatch,E.M. *et al*. (2010) Cep152 interacts with Plk4 and is required for centriole duplication. *J. Cell Biol.*, **191**, 721–729.
37. Kalay,E. *et al*. (2011) CEP152 is a genome maintenance protein disrupted in Seckel syndrome. *Nat. Genet.*, **43**, 23–26.
38. Cui,H. *et al*. (2003) Loss of IGF2 imprinting: a potential marker of colorectal cancer risk. *Science*, **299**, 1753–1755.
39. Zhao,R. *et al*. (2009) Loss of imprinting of the insulin-like growth factor II (IGF2) gene in esophageal normal and adenocarcinoma tissues. *Carcinogenesis*, **30**, 2117–2122.
40. Vorwerk,P. *et al*. (2003) Loss of imprinting of IGF-II gene in children with acute lymphoblastic leukemia. *Leuk. Res.*, **27**, 807–812.
41. Jarrard,D.F. *et al*. (1995) Regional loss of imprinting of the insulin-like growth factor II gene occurs in human prostate tissues. *Clin. Cancer Res.*, **1**, 8.
42. Veronese,A. *et al*. (2010) Oncogenic role of miR-483-3p at the IGF2/483 locus. *Cancer Res.*, **70**, 3140–3149.
43. Hao,J. *et al*. (2011) MicroRNA 483-3p suppresses the expression of DPC4/Smad4 in pancreatic cancer. *FEBS Lett.*, **585**, 207–213.
44. Mohr,F. *et al*. (2011) TET genes: new players in DNA demethylation and important determinants for stemness. *Exp. Hematol.*, **39**, 272–281.
45. Lorsbach,R.B. *et al*. (2003) TET1, a member of a novel protein family, is fused to MLL in acute myeloid leukemia containing the t(10;11)(q22;q23). *Leukemia*, **17**, 637–641.
46. Abdel-Wahab,O. *et al*. (2009) Genetic characterization of TET1, TET2, and TET3 alterations in myeloid malignancies. *Blood*, **114**, 144–147.
47. Coopman,P.J. *et al*. (2006) The Syk tyrosine kinase: a new negative regulator in tumor growth and progression. *Cancer Lett.*, **241**, 159–173.
48. Ma,L. *et al*. (2010) The relationship between methylation of the Syk gene in the promoter region and the genesis of lung cancer. *Clin. Lab.*, **56**, 407–416.
49. Yuan,Y. *et al*. (2001) Hypermethylation leads to silencing of the SYK gene in human breast cancer. *Cancer Res.*, **61**, 5558–5561.
50. Wang,B. (2011) Cancer cells exploit the Eph-ephrin system to promote invasion and metastasis: tales of unwitting partners. *Sci. Signal.*, **175**, pe28–pe30.
51. Shan,Z. *et al*. (2004) Identifying novel homozygous deletions by microsatellite analysis and characterization of tumor suppressor candidate 1 gene, TUSC1, on chromosome 9p in human lung cancer. *Oncogene*, **23**, 6612–6620.