More on SNPs and PCa

Terrence P McGarty
White Paper No 119
November, 2014

There has been a continuing focus on SNPs and Prostate Cancer. This document reviews some of this recent works and adds to what we have done previously.

Notice

This document represents the personal opinion of the author and is not meant to be in any way the offering of medical advice or otherwise. It represents solely an analysis by the author of certain data which is generally available. The author furthermore makes no representations that the data available in the referenced papers is free from error. The Author also does not represent in any manner or fashion that the documents and information contained herein can be used other than for expressing the opinions of the Author. Any use made and actions resulting directly or otherwise from any of the documents, information, analyses, or data or otherwise is the sole responsibility of the user and The Author expressly takes no liability for any direct or indirect losses, harm, damage or otherwise resulting from the use or reliance upon any of the Author's opinions as herein expressed. There is no representation by The Author, express or otherwise, that the materials contained herein are investment advice, business advice, legal advice, medical advice or in any way should be relied upon by anyone for any purpose. The Author does not provide any financial, investment, medical, legal or similar advice in this document or in its publications on any related Internet sites.

Contents

1 INTRODUCTION

The problem often found in examining PCa patients with Gleason 7 lesions is to assess whether they are aggressive or indolent. There has been an explosion of putative markets ranging from SNPs, Genes, promoters, miRNAs, methylations and any combination thereto. We generally understand on a cell by cell basis what is occurring in many malignancies and the logic to such changes. But examining cells in a broad spectrum basis, say from any part of the body, to ascertain what a specific tumor portends is highly problematic. The reason is that

Some researchers have argued that SNPs are highly useful. For example Yonggang et al argue:

A major advantage of using SNP data over microarray data to study genetic predisposition is that, unlike microarray data, a person's SNP pattern is unlikely to change over time. Loosely stated, the SNP pattern collected from a person with a disease is likely to be the same pattern that would have been collected from that person at birth or early in life. Thus, we can use SNP data from patients at any stage of their life and at any stage of their disease progression.

However there is no fully accepted basis of that assertion. For example if the lesion is initiated by some methylation resulting from some excess inflammation, and the methylation induces some resulting transcription blockage, then the SNP is irrelevant unless it can be expressly shown to be causative. The mechanisms for such are problematic. Admittedly a SNP in a promoter region may demonstrate blockage of the promoter but most likely must do so on both chromosomes.

A second major advantage of using SNP data is that the data can be collected from any tissue in the body. With microarray data, the mRNA samples for cancer patients are taken from tumor tissue (e.g., from the colon), and the mRNA samples for healthy donors are taken from healthy tissue of the same type (e.g., colon again). SNP data, on the other hand, is not taken directly from tumor samples, but from any tissue in the body. The benefit of this is that, in addition to being faster to obtain, SNP data is also easier to obtain since less invasive procedures can be used. On the other hand, when using SNP data, we do not expect to have predictors of as high accuracy as we get with microarray data.

Again one must examine this assertion in detail. Does every cell reflect all others in the body? In PCa for example we know that as the tumor progresses the cells express differently, for example look at PTEN, and also in the case of a PCa stem cell we again may have a substantially different expression.

Thus each time we see a result promoting SNPs we must be somewhat cautious.

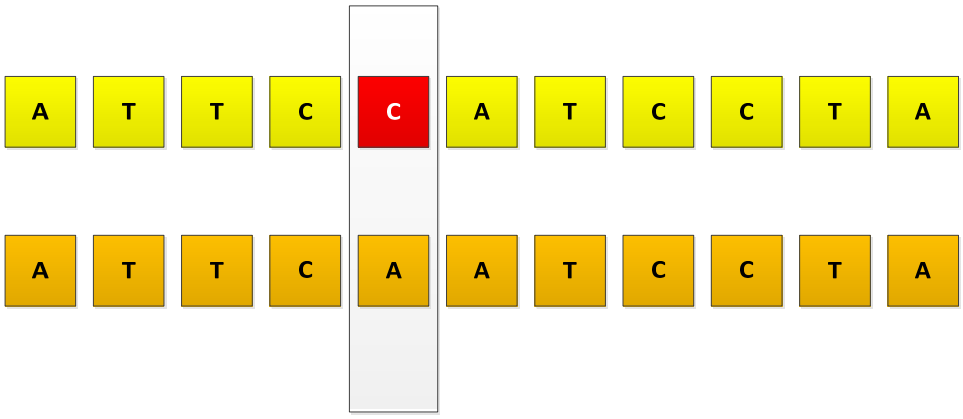
For example in a paper by Lin et al we have an interesting and supportive developed in a causative manner. They state:

However, in many disorders including prostate cancer, the balance between stimulators and inhibitors is tilted to favor stimulators, resulting in an ‘‘angiogenic switch’’. The so- called ‘‘angiogenic switch’’ may result from changes in the expression levels of genes in the angiogenesis pathway. Single nucleotide polymorphisms (SNP) in angiogenesis genes may alter gene expression and influence the process of angiogenesis in prostate cancer and inhibited tumor growth in animal models. Indeed, several SNPs in angiogenesis genes that affect gene expression have been identified. These variants may potentially contribute to inter-individual variation in the risk and progression of prostate tumors. Furthermore, angiogenesis is shown to be associated with the Gleason score, tumor stage, progression, metastasis and survival among prostate cancer patients. Although the number of studies for evaluating the role of SNPs in angiogenesis genes is limited, several of the studies support the association between angiogenesis and prostate cancer aggressiveness. So far, results from several candidate gene and genome-wide association (GWA) studies suggest that SNPs in the angiogenesis pathway may be important in prostate cancer progression and aggressiveness.

Here the authors have not only a correlative connection but a causative one, perhaps.

2 SNPS AGAIN

SNPs are single nucleotide changes in a chromosome. There are millions and the clinical significance is at best problematic. The impact of a SNP on the

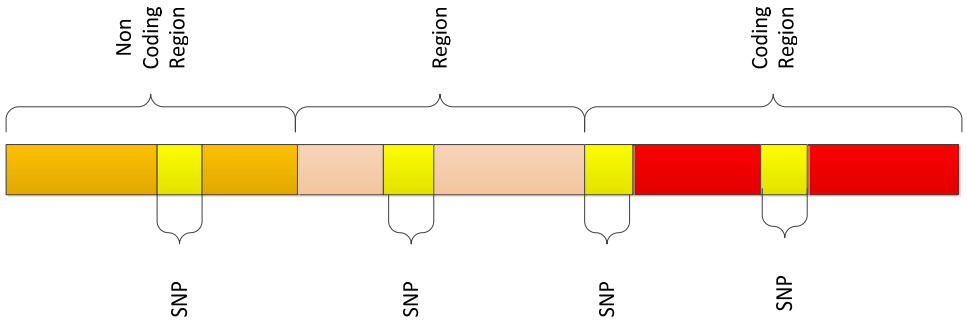


A SNP is a single nucleotide change which occurs in 1% or more of the population.

As Yonggang et al state:

A significant contribution to the genetic variation among individuals is the cumulative effect of a number of discrete, single-base changes in the human genome that are relatively easy to detect. These single positions of variation in DNA are called single nucleotide polymorphisms, or SNPs. While it is presently infeasible to obtain the sequence of all the DNA of a patient, it is feasible to quickly measure that patient's SNP pattern – the particular DNA bases present at a large number of these SNP positions.

The statement of SNPs being substantial elements of genetic variation is not all that obvious. We do observe that some clusters of SNP individuals have a higher propensity for a disease but that may be correlative rather than causative.



SNPs can appear anywhere in a chromosome. As shown above they can be in coding regions, non-coding regions and across promoter regions. What are the effects of these changes? That has been a driving question and it is the issue that must be addressed before correlative effects are used.

3 RECENT REPORTS

In a paper by Yonggang et al they report:

Gleason score (GS) 7 prostate cancer is a heterogeneous disease with different clinical behavior. We sought to identify genetic biomarkers that may predict the aggressiveness of GS 7 diseases. We genotyped 72 prostate cancer susceptibility SNPs identified in genome-wide association studies in 1,827 white men with histologically confirmed prostate adenocarcinoma. SNPs associated with disease aggressiveness were identified by comparing high-aggressive (GS ≥8) and low-aggressive (GS ≤6) cases. The significant SNPs were then tested to see whether they could further stratify GS 7 prostate cancer.

Three SNPs—rs2735839, rs10486567, and rs103294—were associated with biopsy-proven high-aggressive (GS ≥8) prostate cancer (P < 0.05).

Furthermore, the frequency of the variant allele (A) at rs2735839 was significantly higher in patients with biopsy-proven GS 4+3 disease than in those with GS 3 + 4 disease (P = 0.003). In multivariate logistic regression analysis, patients carrying the A allele at rs2735839 exhibited a 1.85-fold (95% confidence interval, 1.31–2.61) increased risk of being GS 4 + 3 compared with those with GS 3 + 4.

The rs2735839 is located 600 base pair downstream of the KLK3 gene (encoding PSA) on 19q13.33 and has been shown to modulate PSA level, providing strong biologic plausibility for its association with prostate cancer aggressiveness. We confirmed the association of the rs2735839 with high-aggressive prostate cancer (GS ≥8).

The question is how does it modulate the activity and if it does then why does a malignancy occur all too often late in life if that SNP has been sitting there for so long. They continue:

Moreover, we reported for the first time that rs2735839 can stratify GS 7 patients, which would be clinically important for more accurately assessing the clinical behavior of the intermediate-grade prostate cancer and for tailoring personalized treatment and post-treatment management.

In effect the above mentioned SNP, which modulates KLK3 or the PSA gene somehow, can be used as a monitor itself. One could then argue that changes in PSA are reflective of changes in the SNP modulation effect and this have a further basis for continuing PSA measurements.

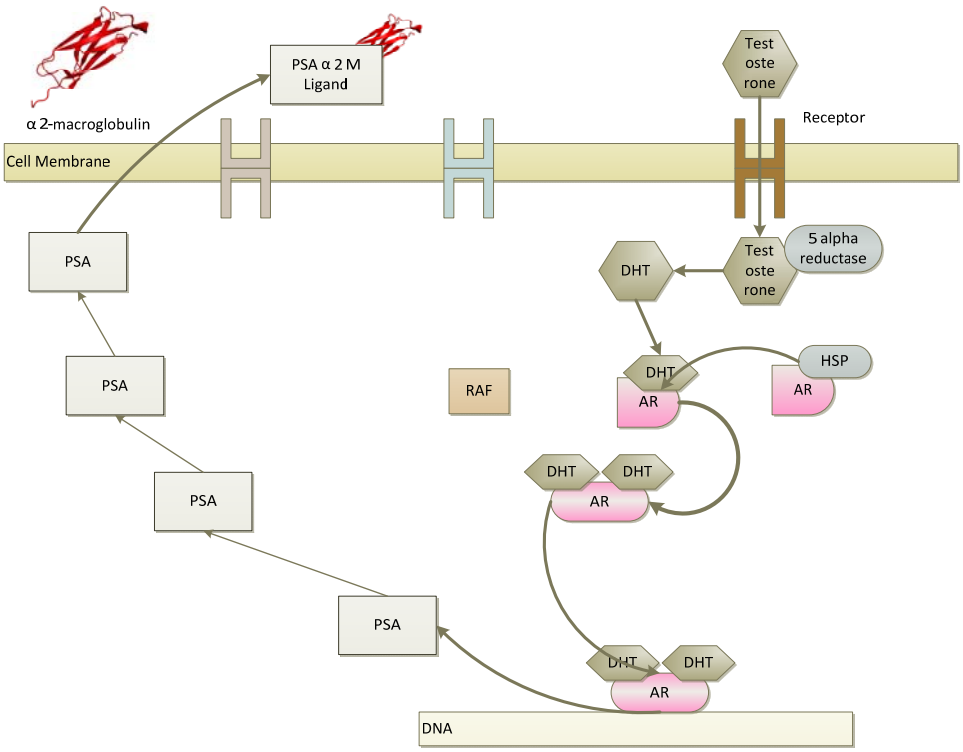
From NCBI we have the following summary discussing KLK3[1]:

Kallikreins are a subgroup of serine proteases having diverse physiological functions. Growing evidence suggests that many kallikreins are implicated in carcinogenesis and some have potential as novel cancer and other disease biomarkers. This gene is one of the fifteen kallikrein

[1] http://www.ncbi.nlm.nih.gov/gene/354

subfamily members located in a cluster on chromosome 19. Its protein product is a protease present in seminal plasma. It is thought to function normally in the liquefaction of seminal coagulum, presumably by hydrolysis of the high molecular mass seminal vesicle protein. Serum level of this protein, called PSA in the clinical setting, is useful in the diagnosis and monitoring of prostatic carcinoma. Alternate splicing of this gene generates several transcript variants encoding different isoforms.

The PSA process is shown below:



The above demonstrates the normal process for PSA production.

From Waltering[2]:

Kallikrein related peptidase 3 (KLK3), better known as prostate specific antigen (PSA), is located in chromosome 19q13.41. KLK3 encodes a single chain glycoprotein with a molecular mass of 33 kDa and functions as a serine protease. It belongs to the family of the fifteen kallikrein members located in a cluster in the same chromosomal region. All kallikrein genes encode five exons of similar size and have high sequence homology with other family members. Many of these peptidases also have several alternative splice variants and are known to be regulated by androgens. KLK3 was cloned in 1987. KLK3 expression has been shown to be elevated in BPH and in highly differentiated PCs, but it is decreased during PC progression.

[2] Waltering, K., Androgen Receptor Signaling Pathway in Prostate Cancer, PhD Thesis, Univ Tampere, Sept 2010.

The use of KLK3 as a PC biomarker (the so-called PSA test) began in the mid-1980s. In a recent European study, which included more than 160,000 men aged 55 to 69; it was found that PSA based screening reduced PC mortality by 20%. However, there was a high risk of overdiagnosis. Androgen regulation of KLK3 includes both the proximal promoter and the enhancer ARE located 4 kb upstream from the TSS. Recruitment of AR and its co-regulators create a chromosomal loop from the enhancer to the core promoter. Kallikrein family members have also been suggested to play a putative role in PC progression. For example, KLK3 has been suggested to directly degrade extracellular matrix glycoproteins and facilitate cell migration.

From a Eureka report on this work they state[3]:

Researchers at The University of Texas MD Anderson Cancer Center have identified a biomarker living next door to the KLK3 gene that can predict which GS7 prostate cancer patients will have a more aggressive form of cancer.

The results reported in the journal of Clinical Cancer Research, a publication of the American Association of Cancer Research, indicate the KLK3 gene – a gene on chromosome 19 responsible for encoding the prostate-specific antigen (PSA) – is not only associated with prostate cancer aggression, but a single nucleotide polymorphism (SNP) on it is more apparent in cancer patients with GS7.

Researchers have linked Gleason score, an important predictor of prostate cancer outcomes, to several clinical end points, including clinical stage, cancer aggression and survival. There has been much research associated with prostate cancer outcomes as well as GS7 prostate cancers, which is an intermediate grade of cancer accounting for 30 to 40 percent of all prostate cancers.

"This is the first report that I am aware of that indicates a genetic variant can stratify GS7 prostate cancer patients," said Jian Gu, Ph.D., associate professor at MD Anderson, and a key investigator on the study. "This is important because this group with heterogeneous prognosis is difficult to predict and there are no reliable biomarkers to stratify this group."

In this study, researchers investigated inherited genetic variants to see if there would be any promising biomarkers for prostate cancer patients. The investigators studied the genetic makeup of 72 SNPs identified from the genome-wide association studies (GWAS) in 1,827 prostate cancer patients. They analyzed associations of these SNPs with disease aggression, comparing them in clinically defined high and low aggressive cases. They found a SNP on the KLK3 gene that can predict an aggressive form of GS7 disease.

"Treatment options for the GS7 disease are controversial because the burden of combined treatment modalities may outweigh the potential benefit in some patients," said Xifeng Wu, M.D., Ph.D., professor and chair of Epidemiology, and lead investigator on the study. "It is critical that we develop personalized treatments based on additional biomarkers to stratify GS7 prostate

[3] http://www.eurekalert.org/pub_releases/2014-10/uotm-rdg100214.php also see http://medicalxpress.com/news/2014-10-gene-aggressive-prostate-cancer-diagnosis.html

cancers. Additional biomarkers may help us achieve personalized clinical management of low and intermediate risk prostate cancer patients."

Wu also said her team are expanding the study and taking a pathway-based approach to systemically investigate genetic variants in microRNA regulatory pathways as biomarkers for the prognosis of prostate cancer patients. "We are also working on circulating biomarkers. Eventually, we will incorporate all biomarkers, epidemiological and clinical variants into nomograms to best predict the prognosis of prostate cancer patients at diagnosis."

Now many others have studies SNPs and PCa[4]. In a recent paper by Mikropoulos et al on Medscape the authors provide an excellent up to date summary[5]:

Several SNPs associated with PrCa risk in the 8q24 locus were among the earliest identified. The 8q24 region is a gene-poor region located upstream of the MYC proto-oncogene and this suggested an association with its expression, which was later proven to occur in a tissue-specific manner. Another important SNP is rs10993994 in the region containing the MSMB gene on chromosome 10. This risk allele associates with reduced MSMB protein expression. MSMB expression is high in normal and benign prostate tissue and low in PrCa. MSMB regulates cell growth and when lost, tumor cells grow in an uncontrolled manner. The odds ratio (OR) for this SNP's association to PrCa was established as 1.61. This is a potential biomarker since urine MSMB assays have been developed and their role in screening is being evaluated.

The Myc region is always a sensitive region. Myc controls cell proliferation and as such needs close control. They continue:

SNP rs2735839 was identified between the KLK2 and KLK3 genes on chromosome 19 where there is a kallikrein gene cluster. Kallikreins are serum proteases and the most well-known member of this group is the prostate-specific antigen (PSA), which is widely used for screening and monitoring PrCa. SNPs were also identified in the intronic region areas of the LMTK2 gene, which codes for cdk5, the SLC22A3 gene, which codes for an organic cation transporter and NUDT10, which regulates DNA phosphorylation.

Again the proximity to PSA gene expression is noted. This has been the case for many previous works not just the one we have focused on herein.

In proximity to the TERT gene (encoding TERT) on 5p15, a further susceptibility SNP was identified (rs2242652). Telomerase is important in counterbalancing telomere-dependent replicative aging. SNPs in this region have been associated with numerous cancers, such as basal cell carcinoma, lung cancer, bladder cancer, glioma and testicular cancer. This SNP showed an association with high PSA levels, as well as increased risk of developing PrCa. Fine-

[4] We had written extensively on this in July 2013.
http://www.telmarc.com/Documents/White%20Papers/99%20SNPs.pdf

[5] http://www.medscape.com/viewarticle/830689

mapping analysis identified a total of four loci independently associated with PrCa risk in the TERT region, one of which was associated with changes in gene expression.

rs2121875 is a SNP located at 5p12 within the FGF10 locus associated with an increased risk of PrCa. FGF10 is often overexpressed in breast carcinomas, and encodes a FGF essential for a range of developmental processes, which also has an important role in the growth of normal prostatic epithelial cells.

In 2013, we reported on 23 new susceptibility alleles associated with PrCa, 16 of which were also associated with aggressive disease.. A SNP located at 1q32 (rs4245739) in proximity to the MDM4 gene is of potential clinical significance. MDM4 inhibits cell cycle arrest and apoptosis, via p53 downregulation.[30] Another SNP (rs11568818) with a potential prognostic value is situated at 11q22 within a region containing the gene MMP7. MMP7 encodes for a matrix metalloproteinase, which is pivotal for tumor metastasis and overexpression of MMP7 is a potential biomarker for PrCa aggressiveness and risk of metastatic disease. Finally, SNP (rs7141529) at 14q24 is an intronic SNP within the RAD51B gene, which is an important DNA repair gene involved in homologous recombination, also associated with PrCa risk.

From this report we also present below the detailed tabular results on a wide variety of SNPs.

SNP	***Susceptibility loci***	***Nearby gene***	***Gene function***
rs1218582	1q21	*KCNN3*	Calmodulin binding
rs4245739	1q32	*MDM4* and *PIK3C2B*	Negative TP53 regulator and therefore inhibits cell cycle arrest and apoptosis and positive regulation of cell proliferation
rs10187424	2p11	*GGCX/VAMP8*	SNARE interactions in vesicular transport
rs721048	2p15	Intronic in *EHBP1*	Eps15 homology domain binding protein
rs1465618	2p21	Intronic in *THADA*	Complex locus
rs13385191	2p24	*C2orf43*	Catalytic activity
rs11902236	2p25	*TAF1B*:*GRHL1*	TBP-associated factor
rs12621278	2q31	Intronic in *ITGA6*	Integrins-cell adhesion cell surface-mediated signaling
rs2292884	2q37	*MLPH*	Exophilin subfamily of Rab effector proteins
rs3771570	2q37	*FARP2*	Rac protein signal transduction
rs2055109	3p11	–	
rs2660753	3p12	–	
rs7611694	3q13	*SIDT1*	Unknown
rs10934853	3q21	Intronic in *EEFSEC*	GTP binding, GTPase activity, nucleotide binding, translation elongation factor activity
rs6763931	3q23	Intronic in *ZBTB38*	Transcriptional activator that binds methylated DNA
rs10936632	3q26	*CLDN11/SKIL*	CNS myelin
rs1894292	4q13	*AFM* and *RASSF6*	Structurally-related serum transport proteins

rs17021918	4q22	Intronic in *PDLIM5*	Cytoskeleton organization, cell lineage specification and organ development oncogenesis
rs12500426	4q22	–	
rs7679673	4q24	*TET2*	Metal ion binding, oxidoreductase activity
rs2121875	5p12	*FGF10*	Important role in the growth of normal prostatic epithelial cells
rs2242652	5p15	*TERT*	Telomerase is important in counterbalancing telomere-dependent replicative aging†
rs12653946	5p15	*IRX4*	Regulation of transcription, DNA dependent
rs6869841	5q35	*FAM44B (BOD1)*	Encoding biorientation of chromosomes in cell division 1
rs130067	6p21	Missense coding in *CCHCR1*	Protein binding
rs1983891	6p21	*FOXP4*	FOX transcription factor family
rs3096702	6p21	*NOTCH4*	Notch signaling network
rs2273669	6p21	*ARMC2* and *SESN1*	*ARMC2*
rs339331	6q22	*RFX6*	RFX family of transcription factors
rs9364554	6q25	Intronic in *SLC22A3*	Cation transporter gene
rs1933488	6q25	*RSG17*	
rs10486567	7p15	Intronic in *JAZF1*	Transcriptional repressor
rs12155172	7p21	*SP8*	Transcription factor
rs6465657	7q21	Intronic in *LMTK2*	Tyrosine kinase
rs2928679	8p21	*SLC25A37*	Mitochondrial carrier proteins
rs1512268	8p21	*NKX3.1*	Homeodomain-containing transcription factor NKX3–1†
rs11135910	8p21	*EBF2*	Regulation of transcription
rs10993994	8q24	*c-MYC* oncogene	Transcription factor activity controlling cell cycle progression, apoptosis and cellular transformation†
rs1447295	8q24	–	
rs6983267	8q24	–	
rs16901979	8q24	–	
rs10086908	8q24	–	
rs12543663	8q24	–	
rs620861	8q24	–	
rs1571801	9q33	Intronic in *DAB2IP*	GAP tumor suppressor
rs817826	9q31	*RAD23B-KLF4*	Nucleotide excision repair
rs1571801	9q33	*DAB2IP*	Ras GAP tumor suppressor
rs10993994	10q13	*MSMB* gene	MSMB regulates cell growth†
rs3850699	10q24	*TRIM8*	Ligase activity

rs4962416	10q26	Intronic in *CTBP2*	Wnt signaling pathway and Notch signaling pathway
rs2252004	10q26	–	
rs7127900	11p15	–	
rs1938781	11q12	*FAM111A*	Proteolysis
rs7931342	11q13	–	
rs11568818	11q22	*MMP7*	Matrix metalloproteinase associated with metastatic potential
rs902774	12q13	*KRT8*	Cellular structural integrity
rs10875943	12q13	*TUBA1C/PRPH*	Protein binding, GTP binding, GTPase activity, nucleotide binding and structural molecule activity
rs1270884	12q24	*TBX5*	Transcription factors involved in the regulation of developmental processes
rs9600079	13q22	–	
rs8008270	14q22	*FERMT2*	Actin cytoskeleton organization, cell adhesion, regulation of cell shape
rs7141529	14q24	*RAD51*	DNA repair
rs684232	17p13	*VPS53* and *FAM57A*	Protein transport
rs7210100	17q21	*ZNF652*	Transcription regulation
rs11650494	17q21	*HOXB13*	Encoding transcription factor homeobox B13
rs4430796	17q12	Intronic in *HNF1B*	Homeodomain-containing superfamily of transcription factors†
rs11649743	17q12	–	
rs1859962	17q24	–	
rs7241993	18q23	*SALL3*	Regulation of transcription†
rs2735839	19q13	*KLK2* and *KLK3* regions	Serine proteases†
rs8102476	19q13	–	
rs11672691	19q13	–	
rs103294	19q13	*LILRA3*	Immunoreceptors expressed predominantly on monocytes and B cells
rs11672691	19q13	–	
rs6062509	20q13	*ZGPAT*	Transmembrane adaptor phosphoprotein
rs2427345	20q13	*GATAS* and *CABLES2*	Cyclin-dependent protein kinase regulator activity
rs2405942	Xp22	*SHROOM2*	Amiloride-sensitive sodium channel activity beta-catenin binding
rs5945619	Xp11	*NUDT11*	Diphosphoinositol-polyphosphate diphosphatase activity, hydrolase activity and metal ion binding†
rs591943	Xq12	Androgen receptor	Androgen receptor regulation

4 OBSERVATIONS

This paper that we have been discussing presents a SNP analysis which has some logical nexus to PSA and pathways often found aberrant in PCa. We are left asking a few questions:

1. What is the controlling mechanism between the SNP and the PSA production?

There seems at best closeness to PSA and an argument that the proximity is reflective of the aggressiveness of the malignancy. There must be a clearer understanding of the entire process before arguing as is done above.

2. Why if the SNP is in the gene does it not cause a PCa effect earlier? What then is the precipitating sequence of events?

This is the key question. If these SNPs are pandemic in all cells then why is PCa specific and why does it take so long? What is truly occurring here?

3. What are the pathway effects?

There appears to be a great deal of inferential data but no clear definitive linkages. The problem with SNPs all too often is the correlative and non-causative relationships.

4. Can one examine a means to block the deleterious effects of this modulation and if so what are they?

This is the therapeutic question. Again one needs the details and not just single nucleotide suggestions.

5. How cell specific is this SNP and as we have seen, many SNPs have broader imputed effects.

We have examined many of the ROC curves and they are interesting but not conclusive. One may not want to bet one's patient's lives on these specific markers

5 REFERENCES

1. Lin , H. et al, SNP-SNP Interaction Network in Angiogenesis Genes Associated with Prostate Cancer Aggressiveness, PLOS ONE, www.plosone.org , April 2013, Volume 8, Issue 4
2. Waddell, M., et al, Predicting Cancer Susceptibility from Single-Nucleotide Polymorphism Data: A Case Study in Multiple Myeloma, BIOKDD '05, August 2005.
3. Waltering, K., Androgen Receptor Signaling Pathway in Prostate Cancer, PhD Thesis, Univ Tampere, Sept 2010.
4. Yonggang, H., et al, The Prostate Cancer Susceptibility Variant rs2735839 Near KLK3 Gene Is Associated with Aggressive Prostate Cancer and Can Stratify Gleason Score 7 Patients, http://clincancerres.aacrjournals.org/content/20/19/5133.abstract