Published in final edited form as:
Cancer Res. 2011 May 15; 71(10): 3442–3446. doi:10.1158/0008-5472.CAN-11-0314.

National Cancer Institute Prostate Cancer Genetics Workshop

William J. Catalona[1], **Joan E. Bailey-Wilson**[2], **Nicola J. Camp**[3], **Stephen J. Chanock**[4], **Kathleen A. Cooney**[5], **Douglas F. Easton**[6], **Rosalind A. Eeles**[7], **Liesel M. FitzGerald**[8], **Matthew L. Freedman**[9], **Julius Gudmundsson**[10], **Rick A. Kittles**[11], **Elliott H. Margulies**[12], **Barry B. McGuire**[1], **Elaine A. Ostrander**[13], **Timothy R. Rebbeck**[14], **Janet L. Stanford**[15], **Stephen N. Thibodeau**[16], **John S. Witte**[17], and **William B. Isaacs**[18]

[1]Department of Urology, Northwestern University Feinberg School of Medicine, Chicago, Illinois, USA [2]National Human Genome Research Institute, National Institutes of Health, Baltimore, Maryland, USA [3]Division of Genetic Epidemiology, University of Utah, Salt Lake City, Utah, USA [4]Division of Cancer Epidemiology and Genetics, National Cancer Institute, NIH, DHHS, Rockville, Maryland, USA and Core Genotyping Facility, SAIC-Frederick, Inc., National Cancer Institute, NIH, DHHS, Gaithersburg, Maryland, USA [5]Department of Internal Medicine, University of Michigan Medical School, Ann Arbor, Michigan, USA [6]Department of Public Health and Primary Care, University of Cambridge, Strangeways Research Laboratory, Cambridge, UK [7]The Institute of Cancer Research, Surrey, UK, and The Royal Marsden NHS Foundation Trust, Sutton, UK [8]Division of Public Health Sciences, Fred Hutchinson Cancer Research Center, Seattle, Washington, USA [9]Department of Medical Oncology, Dana-Farber Cancer Institute, Boston, Massachusetts, USA and The Broad Institute, Cambridge, Massachusetts, USA [10]deCODE genetics, IS-101 Reykjavik, Iceland [11]Department of Medicine, University of Illinois, Chicago, Illinois, USA [12]Genome Technology Branch, National Human Genome Research Institute, National Institutes of Health, Bethesda, Maryland, USA [13]Cancer Genetics Branch, National Human Genome Research Institute, Bethesda, Maryland, USA [14]Center for Clinical Epidemiology and Biostatistics, University of Pennsylvania, Philadelphia, Pennsylvania [15]Division of Public Health Sciences, Fred Hutchinson Cancer Research Center, Seattle, Washington, and Department of Epidemiology, School of Public Health, University of Washington, Seattle, Washington, USA [16]Department of Laboratory Medicine and Pathology, Mayo Clinic, Rochester, Minnesota, USA [17]Departments of Epidemiology and Biostatistics, and Urology, Institute for Human Genetics, University of California, San Francisco, San Francisco, California, USA [18]Department of Urology and Oncology, Johns Hopkins Medical Institutions, Baltimore, Maryland, USA

Abstract

Compelling evidence supports a genetic component to prostate cancer (PC) susceptibility and aggressiveness. Recent genome-wide association studies (GWAS) have identified >30 single nucleotide polymorphisms associated with PC susceptibility. It remains unclear, however, whether such genetic variants are associated with disease aggressiveness—one of the most important questions in PC research today. To help clarify this and substantially expand research in the genetic determinants of PC aggressiveness, the first National Cancer Institute Prostate Cancer Genetics Workshop assembled researchers to develop plans for a large new research consortium and patient cohort. The workshop reviewed the prior work in this area and addressed the practical issues in planning future studies. With new DNA sequencing technology, the potential application of sequencing information to patient care is emerging. The workshop, therefore, included state-of-

Corresponding Author: William J. Catalona, MD, 675 North St. Clair Street, Suite 20-150, Chicago, IL 60611. Phone: 312-695-4471; Fax: 312-695-1482; wcatalona@nmff.org.

the-art presentations by experts on new genotyping technologies, including sequencing and associated bioinformatics issues, which are just beginning to be applied to cancer genetics.

Introduction

William Catalona (Northwestern University, Chicago, Illinois) opened the workshop with a discussion of the public health problem caused by the heterogeneity of prostate cancer (PC) aggressiveness, i.e., the inability to accurately identify those men destined to suffer and die from PC. Elucidation of genetic markers associated with aggressive disease is one of the most important endeavors in PC research today, as better markers to identify aggressive PC could substantially improve patient care.

Catalona presented the history of the Genetics Working Group (GWG) and outlined the goals of the workshop. The National Cancer Institute's (NCI) Specialized Program of Research Excellence (SPORE) GWG was formed in 2007, with the aim to perform a case-case association study of aggressive vs. non-aggressive PC, genotyping the 30+ validated PC susceptibility single nucleotide polymorphisms (SNPs) in >20,000 patients from the PC SPORE sites. The SNPs associated with PC aggressiveness would then be tested for association with other disease characteristics and treatments. The NCI funded the GWG's proposal to hold a workshop to: 1) plan the proposal for funding the consortium's case-case study, 2) "fine tune" the strategies for conducting future studies, and 3) plan an expanded collaboration beyond the SPORE sites.

Matthew Freedman (Dana Farber Cancer Institute, Boston, Massachusetts) commented that the combined SPOREs alone bring a tremendous potential to bear on PC biology. The characteristics distinguishing the SPORE populations from the other existing cohorts are the deep clinical annotation; their size; the potential to study multiple ancestral groups; the potential to study pharmacogenetics; and, because of the availability of thousands of tumor samples from the same patient population, studies of combined germ-line and somatic abnormalities could be enhanced.

The GWG will leverage its combined resources to conduct research that otherwise would not be possible without this synergy. A large multi-institutional virtual biobank of clinically- and genetically-annotated samples would be extremely useful to a wide array of researchers for a number of activities including: 1) discovery of rare, highly penetrant susceptibility variants, 2) validation of discoveries in biomarker development, and 3) identification of patients who are carriers of known genetic variants and, thus, potential candidates for clinical trials of targeted therapy.

Genetic determinants of lethal disease are likely to have small effect sizes, and larger studies may be necessary to identify them (1). Until research is directed at this area, identifying genetic variants that contribute to PC aggressiveness is unlikely. Many speakers echoed the need for establishing a broader-based consortium to provide the large sample sizes required for studies on PC aggressiveness.

Practical Issues

The workshop addressed the short- and long-term objectives of the GWG and also possible future projects. The consensus was that it would be wise first to perform a preliminary "meta-analysis" of aggressive vs. non-aggressive PC from the existing studies with genotype and clinical data available for the 30+ SNPs. It was also agreed that additional promising candidate SNPs, identified from ongoing genetic studies, should be analyzed, thereby constantly expanding the dataset to validate newly discovered variants.

The effect sizes for genetic variants associated with PC risk are small, and it is currently unknown whether the effect size for variants associated with aggressive disease will be similar. From the PC GWAS scans already available, it might be possible to estimate whether effect sizes are smaller or larger for aggressiveness than for risk; if effect sizes are predicted to be larger, a smaller sample size might yield sufficient statistical power. However, if smaller, many speakers echoed the need for establishing a broader-based consortium to provide the large sample sizes necessary to identify statistically significant associations. To generate reliable results, tens of thousands of samples are needed. In a complex disease such as PC, stratification for aggressiveness considerably reduces the number of usable samples. As the variants associated with the less common aggressive phenotype may in turn be less common, very large numbers are needed to have a higher power to discover variants with a minor allele frequency (MAF) of <2.5%.

There was also agreement that while the initial study to launch the consortium should include the 30+ SNPs as the model, ultimately the consortium should perform larger case-case-control analyses of promising candidate SNPs identified from other genetic studies, and possibly even undertake a GWAS, constantly expanding the datasets to validate newly discovered variants.

Coordinating and harmonizing data were also noted as being critical issues. While decentralized genotyping can be problematic and inefficient, other consortia have successfully used decentralized genotyping for small numbers of variants. Phenotype presents perhaps the greatest challenge for standardization and harmonization of data. Defining the aggressive phenotype has been problematic. Comparing patients from the tails of the aggressiveness distribution, without unduly compromising sample size and power, enhances the ability to detect the genetic associations; however, merely dividing patients into those one would predict would do well or poorly is not sufficient.

The proposed aim of the GWG was to perform a case-case association study; however, several speakers suggested that comparing genotypes of cases to controls should be included. If the focus is just on cases and significant associations are found, one would not know whether the association represented increased risk in the more aggressive cases or decreased risk in the less aggressive ones. The use of publicly available control data was discussed in this context.

Functional Studies and Other Opportunities

Matthew Freedman pointed out that there are other important questions to ask on a genome-wide scale with either germ-line or somatic tissues, using a variety of platforms. He discussed the need to decipher the relationships between the germ-line and somatic genomes and how noncoding regions interact with other genes. He pointed out that gene expression itself is a heritable trait.

Genetics of PC Aggressiveness

(i) Genetics of PC Progression

William Isaacs (Johns Hopkins University, Baltimore, Maryland) emphasized that accumulation of mutations in key genes is a critical aspect of our current views of PC progression. The importance of somatic gene fusions has emerged as paramount (2), and understanding germline factors that predispose to fusion events could be important for initiation or for early events in the carcinogenic process.

Family history is a known risk factor, and studies have reported familial concordance in PC survival. Isaacs cited studies in families that identified several chromosomal regions linked to PC aggressiveness (3,4). An important association with aggressiveness is found in carriers of *BRCA1* and *BRCA2* mutations (5). Therefore, even though rare variants may account for a small fraction of PCs, they may be more important determinants of aggressive disease. Isaacs and other speakers suggested that most of the PC susceptibility loci identified so far probably modulate early stages of disease rather than disease progression.

(ii) GWAS Replications for Aggressive Disease

A study by Liesel FitzGerald and colleagues (Fred Hutchinson Cancer Research Center, Seattle, Washington) found only 1 GWAS risk allele (on chromosome *X*) associated with Gleason score. From a review of published studies from this group, more SNPs from candidate genes than from GWAS were associated with aggressiveness, and only rarely were the candidate gene SNPs for PC risk also associated with aggressiveness. John Witte (University of California San Francisco, California) reported on a study that attempted to replicate GWAS risk SNPs for PC aggressiveness. They found that PC risk alleles on chromosomes 17 and *X* were also associated with biochemical failure after treatment; however, many other susceptibility variants were not associated with biochemical failure (6).

(iii) Proposed International GWAS of Aggressive PC

Janet Stanford (Fred Hutchinson Cancer Research Center, Seattle, Washington) discussed preliminary plans for an international collaborative GWAS focused on aggressive PC. The projected sample size is about 10,000 aggressive PC cases and 10,000 age-matched controls without diagnosed PC.

(iv) Early Age-at-Onset PC

Kathleen Cooney (University of Michigan, Ann Arbor, Michigan) reported evidence for a significant genetic component and more aggressive PC in young men. Their GWAS scan of early-onset cases vs. publicly available controls revealed an association on *8q24* ($P = 1.2 \times 10^{-8}$) and 3 other SNPs on chromosomes *10q11, 11q13, and 11p15*. The average number of PC risk alleles was higher in the early-onset cases compared to population controls.

(v) GWAS on Variation in PSA Values

Julius Gudmundsson (deCODE genetics, Inc., Reykjavik, Iceland) presented results from a GWAS performed on variation in PSA values, which have high heritability. They identified six genome-wide significant loci associated with PSA levels; all but 2 were previously associated with PC risk (7). New loci on *10q* and *12q* were associated with basal PSA levels but not with PC risk. Genetic variants that impact basal PSA levels may affect the frequency of recommending a prostate biopsy and cause PC diagnosis to be delayed among men with genetically low PSA secretion; nevertheless, the improvement from incorporating these SNPs is modest (8).

(vi) Familial PC

William Isaacs (Johns Hopkins) reported on the International Consortium for Prostate Cancer Genetics (ICPCG) group's linkage signals on chromosomes *6, 11, and 20* in one large combined analysis of families with more aggressive disease (9) and on chromosomes *1, 4, 8, and 12* in a second set of such families. Follow-up fine-mapping (e.g., see Ostrander below) and whole-exome sequencing studies (e.g., see Thibodeau below) of some of these linkage signals are being pursued by ICPCG groups.

(vii) Ethnic Heterogeneity

Rick Kittles (University of Illinois, Chicago, Illinois) noted that ethnic genetic heterogeneity must be considered in GWAS studies, because regions associated with PC in one population cannot be assumed to confer the same level of association in other ethnic groups. Variation exists in allele frequency and in effect size across ethnic groups. Candidate genes and pathways in PC also have variants that differ significantly across different populations. Risk alleles in regions on 8q24 near *MYC* are more common in men of African descent and may account for much of the higher risk among men of African versus European descent. Most of the genotyping platforms used for GWAS to date have poor coverage of variants for African-descent populations. Timothy Rebbeck (University of Pennsylvania, Philadelphia, Pennsylvania) reported that the MADCaP consortium validated the regions on *8q24* in African-Americans, but could only validate 2 other loci detected from GWAS on European descent populations (*MSMB* and *JAZF1*). GWAS of prostate cancer in African-American men are currently underway (Principal Investigators: Brian E. Henderson, University of Southern California, Los Angeles, California, and John A. Witte).

Technical Aspects of DNA Sequencing

Several presentations dealt with current approaches to follow-up of promising loci and new DNA sequencing and genotyping technologies.

(i) Fine-Mapping Studies

In discussing follow-up studies of promising loci, Elaine A. Ostrander (National Human Genome Research Institute, Bethesda, Maryland) described an example of how her group fine-mapped a susceptibility locus on *22q* (10). She emphasized that they had many advantages (e.g., large amount of data from other groups and the ability to narrow the linkage signal to few hundred kilobases in a gene-poor region that was functionally easy to identify) and warned that with other loci, the challenges will be greater.

(ii) Whole Exome and Genome Sequencing

Stephen Thibodeau (Mayo Clinic, Rochester, Minnesota) presented pilot studies in whole exome sequencing in ICPCG families. They selected families with a large number of affected individuals and available first cousins. He described advances in sequencing and target capture technology and indicated that they are currently obtaining high-quality data containing up to 150 million reads, >70% of the sequences having >40x coverage, identifying ~35,000 filtered on-target SNPs of which ~16% were novel, including an average of ~30 splice and ~100 nonsense changes per individual. The remaining changes were missense and synonymous changes. He emphasized the need for considerable QC and bioinformatics support for these projects. To identify candidate targets, they filter the data initially at sharing across family members for nonsense and splice variants.

Stephen J. Chanock (National Cancer Institute, Bethesda, Maryland) emphasized how quickly the technology is moving and how difficult it is to analyze the vast amounts of data generated. Whole exome sequencing should be performed in families in which multiple members could also be sequenced for comparative purposes. Coverage of sequence is the most important consideration. Current versions of whole exome sequencing capture methods provide approximately 70% of the coverage desired, and up to 2000 potentially important genes are missed, including cancer genes. False negatives are particularly troublesome and validation is problematic. Bioinformatics issues are extremely difficult; high errors generate too many false positives for efficient follow-up with variants; since the bioinformatic tools for predicting actual functionally, important coding shifts or terminations are imprecise, follow-up analysis should include a large set of variants; and an agnostic testing approach is

not feasible because of multiple comparisons. He also questioned the wisdom of attempting whole genome sequencing for the opportunity of examining a favored candidate region.

(iii) Bioinformatics

Elliott Margulies (National Human Genome Research Institute, Bethesda, Maryland) also discussed bioinformatics issues, explaining techniques for reducing false positives and statistical "noise." The sequence reads he has obtained have increased to 95% of the genome with a low false-positive rate. He reported that somatic variants are suppressed in non-coding functional regions and postulated that defining them may be another way of segmenting the genome into various regions.

(iv) New Gene Chip Technology

Pointing out that most of the known PC risk alleles have a MAF ≥10%, Chanock indicated that more sophisticated technology is needed to find rarer variants (MAF <5%). He described a new commercial 2.5 million SNP chip that enables interrogation of less common SNPs, namely those with a MAF between 3% and 10%. The chip includes at least 10%-15% more common variant bins not captured in the previous chips.

PC Genetics Consortia

Rosalind Eeles (Institute of Cancer Research and Royal Marsden Hospital, London, U.K.) reviewed the multiple logistical issues to be considered in setting up genetics research consortia and highlighted the current activities of several related consortia, including PRACTICAL (29 groups with >56, 000 samples for follow-up of genetic variants for validation), and ELLIPSE (NIH-funded U19 [Elucidation of Loci for Prostate Cancer Susceptibility] investigating the role of SNP profiles in different aspects of disease progression and management). The GWG will coordinate its studies with these efforts.

Summary

Multiple lines of evidence suggest a genetic component to PC aggressiveness, but it has not been possible to define it using variants identified from current studies. The clinical importance of unraveling the molecular genetics of PC aggressiveness cannot be overstated and is urgently needed, as distinguishing between indolent and aggressive PC would be of tremendous clinical benefit. We propose undertaking genotyping-based projects here, because while undertaking sequencing is appealing, it remains expensive and computationally complicated.

The success of this workshop was evidenced by the number and quality of scientists in attendance and by the strength of their presentations. The information presented and the thoughtful discussion will be valuable in planning the development of the GWG consortium, the establishment of the large PC patient cohort, and the implementation of the large case-case-control association study of aggressive PC.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

We thank the many people who helped organize the meeting, including Drs. Robin Leikin and Chung Lee. We also appreciate the assistance and guidance of Drs. Toby Hecht and Andrew Hruszkewycz and other NCI staff members,

as well as Dennis Murphy (Northwestern Memorial Hospital), Michelle Pattison and Daniel O'Brien (Northwestern University), Mahogany Hopkins (Nova), Pauline Minhinnett (American Society of Human Genetics).

Financial support; This workshop was supported by resources from The National Cancer Institute.

References

1. Penney KL, Pyne S, Schumacher FR, Sinnott JA, Mucci LA, Kraft PL, et al. Genome-wide association study of prostate cancer mortality. Cancer Epidemiol Biomarkers Prev. 2010; 19:2869–76. [PubMed: 20978177]
2. Palanisamy N, Ateeq B, Kalyana-Sundaram S, Pflueger D, Ramnarayanan K, Shankar S, et al. Rearrangements of the RAF kinase pathway in prostate cancer, gastric cancer and melanoma. Nat Med. 2010; 16:793–8. [PubMed: 20526349]
3. Witte JS, Goddard KA, Conti DV, Elston RC, Lin J, Suarez BK, et al. Genomewide scan for prostate cancer-aggressiveness loci. Am J Hum Genet. 2000; 67:92–9. [PubMed: 10825281]
4. Stanford JL, McDonnell SK, Friedrichsen DM, Carlson EE, Kolb S, Deutsch K, et al. Prostate cancer and genetic susceptibility: a genome scan incorporating disease aggressiveness. Prostate. 2006; 66:317–25. [PubMed: 16245279]
5. Mitra AV, Bancroft EK, Barbachano Y, Page EC, Foster CS, Jameson C, et al. Targeted prostate cancer screening in men with mutations in BRCA1 and BRCA2 detects aggressive prostate cancer: preliminary analysis of the results of the IMPACT study. BJU Int. 2011; 107:28–39. [PubMed: 20840664]
6. Cheng I, Plummer SJ, Neslund-Dludas C, Klein EA, Casey G, Rybicki BA, Witte JS. Prostate cancer susceptibility variants confer increased risk of disease progression. Cancer Elpidemiol Biomarkers Prev. 2010; 19:2124–32.
7. Gudmundsson J, Besenbacher S, Sulem P, Budbjartsson DF, Olafsson I, Arinbjamarson S, et al. Genetic correction of PSA values using sequence variants associated with PSA levels. Sci Transl Med. 2010 Dec 15.2(62):62ra92.
8. Witte JS. Personalized prostate cancer screening: improving PSA tests with genomic information. Sci Transl Med. 2010; 2:62.
9. Schaid DJ, McDonnell SK, Zarfas KE, Cunningham JM, Hebbring S, Thibodeau SN, et al. Pooled genome linkage scan of aggressive prostate cancer: results from the International Consortium for Prostate Cancer Genetics. Hum Genet. 2006; 120:471–85. [PubMed: 16932970]
10. Johanneson B, McDonnell SK, Karyadi DM, Quignon P, McIntosh L, Riska SM, et al. Family-based association analysis of 42 hereditary prostate cancer families identifies the Apolipoprotein L3 region on chromosome 22q12 as a risk locus. Hum Mol Genet. 2010; 19:3852–62. [PubMed: 20631155]