NIH Public Access
Author Manuscript
Cancer J. Author manuscript; available in PMC 2014 February 25.

Published in final edited form as:
Cancer J. 2011 ; 17(6): 416–422. doi:10.1097/PPO.0b013e31823e5387.

The Genetics of Cancer Risk

Mark M. Pomerantz, MD and **Matthew L. Freedman, MD**

Abstract

One hundred years ago, decades prior to the discovery of the structure of DNA, debate raged regarding how human traits were passed from one generation to the next. Phenotypes, including risk of disease, had long been recognized as having a familial component. Yet it was difficult to reconcile genetic segregation as described by Mendel with observations exhaustively documented by Karl Pearson and others regarding the normal distribution of human characteristics. In 1918, RA Fisher published his landmark paper, "The Correlation Between Relatives on the Supposition of Mendelian Inheritance," bridging this divide and demonstrating that multiple alleles, all individually obeying Mendel's laws, account for the phenotypic variation observed in nature.

Since that time, geneticists have sought to identify the link between genotype and phenotype. Trait-associated alleles vary in their frequency and degree of penetrance. Some minor alleles may approach a frequency of 50% in the human population while others are present within only a few individuals. The spectrum for penetrance is similarly wide. These characteristics jointly determine the segregation pattern of a given trait, which, in turn, determine the method used to map the trait. Until recently, identification of rare, highly penetrant alleles was most practical. Revolutionary studies in genomics reported over the past decade have made interrogation of most of the spectrum of genetic variation feasible.

The following article will review recent discoveries in the genetic basis of inherited cancer risk and how these discoveries inform cancer biology and patient management. While this article focuses on prostate cancer, the principles are generic for any cancer and, indeed, for any trait.

FAMILY HISTORY and CANCER RISK

The epidemiology of common cancers supports the notion that many malignancies tend to aggregate in families. Family history has been examined extensively as a risk factor for prostate cancer [1-9] and the disease serves as a useful model for studying cancer heritability since the familial contribution to disease risk is high [10,11]. For example, one meta-analysis reviewed 33 epidemiologic studies and determined that subjects with a first-degree relative with prostate cancer are at approximately 2.5-fold lifetime risk of disease[12]. Risk increased to approximately 5-fold for those with two or more affected first-degree family members. These trends were reinforced in more recent analyses from the Health Professionals Follow-up Study (HPFS) and the Swedish Family-Cancer Database[13,14]. Family history is similarly risk factor for other common malignancies such as breast cancer and colon cancer[15,16].

While several series suggest the importance of family history in developing cancer, they are limited in their ability to distinguish genetic from non-genetic factors. The familial

Corresponding Author Matthew L. Freedman, MD Department of Medical Oncology Dana-Farber Cancer Institute Boston, MA freedman@broadinstitute.org.

This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

NIH-PA Author Manuscript
NIH-PA Author Manuscript
NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

component of disease risk is not necessarily inherited, as family members share similar lifestyles and exposures. Twins offer a unique study population for dissecting the genetic and non-genetic components of a given phenotype. Siblings often share similar environmental exposures, while monozygotic twins additionally share genetic make-up. Dizygotic twins, on the other hand, like non-twin siblings, only share half of their genomes. A Veterans Administration study of over 16,000 male twins in the U.S. demonstrated a concordance for Prostate cancer diagnosis of 27.1% for monozygotic twins versus 7.1% for dizygotic twins, suggesting a genetic component to risk[17]. Similarly, in 2000, an analysis of 44,788 pairs of twins listed in Swedish, Danish and Finnish twin registries revealed a significantly higher concordance for monozygotic twins compared with dizygotic twins[18]. Using a model developed to determine the effects of heritable versus environmental factors, heritable factors were estimated to account for 42% of prostate cancer risk. Heritable factors were estimated to account for 27% of breast cancer risk and 35% of colon cancer risk. Cervical cancer on the other hand, appeared to have almost no heritable component, as might be expected in a disease caused by infection. These estimates include the small subset of cases within each disease type that are highly heritable, such as BRCA1- and BRCA2-associated breast cancer and FAP-associated colon cancer. These Mendelian disorders, described below, typically account for 5% cases. The large majority of cancer cases are genetically more complex.

MENDELIAN INHERITANCE

Family history and twin studies strongly implicate heredity in disease susceptibility. The first genes clearly associated with inherited risk of cancer were discovered using a Mendelian approach, focusing on families in which a particular tumor or set of tumors were transmitted in an identifiable pattern: autosomal dominant, autosomal recessive, or sex-linked. These genes have been discovered by linkage analysis, in which sub-chromosomal regions co-segregating with affected family members are identified and then closely examined for deleterious mutations. Causative genes involved in diseases and syndromes such as BRCA-associated breast cancer, xeroderma pigmentosum, familial adenomatous polyposis and Li-Fraumeni syndrome have been discovered using this method[19-23]. These Mendelian disorders are all caused by rare alleles with very high penetrance (Table 1). While linkage mapping has produced several putative cancer loci, the method has not been able to identify genetic variants associated with the most common, sporadic cancer. Linkage analysis has been applied to prostate cancer, for example, and risk loci have been reported[24-36], however few have been consistently validated in independent cohorts[37-39].

Several factors complicate linkage analysis in common, complex diseases such as prostate cancer. Since prostate cancer is a common disease, families may include members who have developed sporadic forms of the disease. These subjects, termed phenocopies, can confound linkage analysis. Also, unlike BRCA-associated breast cancer, which presents relatively early in life[40], or familial adenomatous polyposis, which has a distinctive clinical presentation[41], there is little to clinically or pathologically distinguish prostate cancer densely clustered in families from sporadic disease. Also, since most sporadic cancers occur relatively late in life, it is difficult to obtain DNA samples and clinical data from more than one generation. Investigation into cancer risk clearly indicates that Mendelian segregation of this phenotype is the exception rather than the rule.

COMPLEX TRAITS and GWAS

Unlike Mendelian diseases, which are governed by highly penetrant variants that segregate according to clear patterns within families (e.g., autosomal dominant, autosomal recessive, sex-linked, mitochondrial), complex diseases result from the interplay of genetic,

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

environmental and stochastic factors. The genetic risk in complex disease is comprised of multiple alleles, with no single allele being fully deterministic for driving tumorigenesis (i.e., modestly penetrant). To identify alleles associated with complex phenotypes, focus shifted from highly penetrant alleles clustered within families to more common variants present in larger, unrelated populations (Table 1).

Initial efforts to identify modestly penetrant alleles associated with cancer risk relied on resequencing candidate genes predicted to play a role in disease risk. Associations were sought by measuring differences in allele frequencies at polymorphisms between cases and controls. While convincing findings have been reported for certain common malignancies, such as bladder cancer[42,43], the candidate gene approach has yielded few associations robustly validated in independent cohorts. In prostate cancer for example, the gene for the androgen receptor warranted significant attention given its known role in prostate carcinogenesis. However, extensive annotation of variation across the gene in prostate cancer cases and matched controls yielded no inherited variants associated with risk[44].

A less biased approach was needed to identify the alleles associated with complex disease. Genome-wide association studies (GWAS) scan the genome for polymorphisms, usually single nucleotide polymorphisms (SNPs), which are associated with a trait of interest. GWAS compare allele frequencies among individuals with a phenotype of interest to frequencies among unaffected individuals. Over the past 10 years several advances made possible the implementation of GWAS: the sequencing of the human genome; the publication of the initial phases of the International HapMap Project, which has catalogued common genetic polymorphisms and their correlations with one another[45-47]; the emergence of technologies that allow high-throughput genotyping of hundreds of thousands of polymorphisms simultaneously; and the development of statistical methods for interpreting the massive amounts of data generated by GWAS and imputing genotypes based on genetic correlation.

GWAS take an unbiased approach in the search for genetic polymorphisms associated with disease, evaluating a substantial portion of the variation across the genome. While the International HapMap Project has catalogued over 10 million SNPs, it is not necessary to genotype and analyze all SNPs in order to achieve genome-wide coverage for common alleles. Nearby SNPs are co-inherited more often than would be expected by chance. A single SNP can serve as proxy for much of the variation in the surrounding genetic region, and due to this linkage disequilibrium (LD), the number of genotypes necessary to conduct a GWAS is greatly reduced. LD must be empirically determined and differs across ethnic groups. Genotyping 500,000 to 1 million “tagged” SNPs can capture roughly 80% of all common SNPs in a given population[48]. Nonetheless, testing up to a million independent SNPs raises important statistical considerations[49]. Due to the potential for a large number of false positives, strict statistical thresholds are necessary to identify true positives rather than associations observed merely by chance. Because of this, a stringent p-value threshold $<5\times10^{-8}$ is commonly applied. In order to achieve this statistical threshold, large datasets, comprised of thousands of cases and controls, are necessary.

Since 2006, over 150 bona fide risk alleles have been discovered for dozens of cancers, including approximately 40 polymorphisms associated with prostate cancer risk[50-62] (see http://www.genome.gov/gwastudies/ for catalogue of GWAS findings reported to date). An encouraging observation is the reproducibility in independent cohorts of most findings[63]. Odds ratios associated with risk alleles for common, polygenic diseases tend to be modest, generally less than 1.5. The power to detect an effect of this size requires very large study populations. Assembling adequately sized cohorts can be extremely challenging. In part due to these power considerations, most GWAS to date report associations with SNPs whose

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

minor allele frequencies are >10%. As larger cohorts are collected and GWAS combine data in meta-analyses, more trait-associated variants with smaller minor allele frequencies may emerge.

Despite the large number of cancer risk loci reported and validated to date, these variants only explain a fraction of the estimated heritability. For example, the 40 risk variants for prostate cancer are estimated to explain 25% of the heritability. Where is the rest of the genetic contribution to disease? There are several explanations for this gap between what has been achieved by GWAS and what remains to be found. Most GWAS have not been adequately powered to capture associations between disease and alleles with minor allele frequencies of 1- 5%. It is hypothesized that alleles even rarer in the population, <1%, may account for much of this gap. Very rare alleles associated with disease may have greater impact. Rather than odds ratios of 1.1-1.5, they may influence disease with higher odds ratios. The 1000 Genomes Project, a cataloguing of human genetic variation based on whole genome sequencing presents the opportunity to explore this possibility given suitably large cohorts[64]. Another possibility is that genome-wide surveys of structural variants, such as copy number variation, will account for some of the heritability gap. These variants are poorly represented in the arrays used for most GWAS. Finally, it is possible that gene-gene and gene-environment interactions play a significant role in inherited risk. The complexities involved in the study of these factors are daunting, but strides are being made[65].

Certain trends have emerged in cancer-related GWAS. There are regions across the genome containing inherited variants for more than one disease. One of these regions is chromosome 8q24, first identified in 2006 as a prostate cancer risk locus in both European and African American populations[51,66]. The region includes the well-known oncogene MYC. Several other prostate cancer GWAS converged on 8q24, and, to date, a total of at least nine SNPs, all independently associated with prostate cancer risk, reside at 8q24[50,61,67]. Intriguingly, risk markers for breast, colon, bladder cancer and chronic lymphocytic leukemia have been discovered at this chromosomal locus.[50,51,66,68-71]. Similarly, chromosome 5p15 harbors multiple risk variants, including SNPs for prostate cancer, glioma, pancreatic cancer, bladder cancer, lung cancer, breast cancer, uterine cancer, melanoma, and basal cell carcinoma[72-78]. The region contains the gene TERT which is involved in telomerase activity. Mutations in this gene have been implicated in bone marrow failure syndromes and hematologic malignancies[79,80]. Two SNPs associated with prostate cancer at another locus, chromosome 17q12, reside at HNF1B. Variants associated with prostate cancer in this region are also associated with type 2 diabetes. However the effects of the risk allele are in the opposite direction for the two phenotypes, raising interesting questions regarding the relationship between prostate tumorigenesis and metabolic processes[55].

Another trend in cancer GWAS is the differences in risk allele discovery across diseases. GWAS in prostate cancer, for example, have yielded more associated variants compared to other common cancer such as lung cancer. There are several possible reasons for this. Due to its ubiquity and the relative good health of men with disease, large cohorts have been assembled more readily. Also, prostate cancer has a stronger inherited component compared with other common cancers[18]. Prostate cancers may also be more homogenous than other cancers. For example, case series of lung cancer, for which fewer than 10 associated variants have been found, may include genetically distinct subtypes of disease, affecting the statistical power of finding an association. Breast cancer GWAS results demonstrate certain polymorphisms that appear specific for estrogen receptor (ER)-positive and others for ER-negative disease[81].

GWAS data for populations other than those of European ancestry are generally lacking. While many risk alleles replicate across ethnic groups, there may be cases where the genetic

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

architecture of disease risk differs. This can have substantial implications in any personalized approach to patient care. A prostate cancer GWAS has recently been performed using African American cases and controls, a novel risk SNP in this population has been identified[82]. Further work across multiple ethnic groups should be pursued in order to have a composite picture of disease risk.

FINE MAPPING

SNPs discovered in GWAS are likely not to be the causative polymorphisms. Most SNPs reside in LD blocks with multiple other polymorphisms. The risk SNP described in a GWAS may merely be a proxy for the true causal variant. Fine mapping is a method used to home in on the allele or alleles truly responsible for a given phenotype. A strategy used to comprehensively interrogate a newly discovered disease risk locus begins by resequencing the region in a set of cases and controls to ascertain the full complement of germline variants in the population[74,83,84]. Each variant is then analyzed in a larger set of cases and control for association with the trait. Statistical models are used to determine the allele or set of alleles that most exhaustively accounts for the association.

INSIGHT INTO the MECHANISMS OF INHERITED RISK

An intriguing and perhaps unexpected outcome from GWAS has been the finding that almost 90% of reported disease-associated SNPs occur in non-coding regions of the genome[85,86]. Over 40% of these have been found in intergenic regions. The functional consequences of inheriting a risk allele are not readily apparent. Insight into the mechanisms underlying associations between risk loci and cancer will increase understanding of the genes and pathways mediating tumorigenesis.

Inherited variants can influence phenotype in several ways: by directly altering gene transcription and amino acid sequence, by disrupting transcription of non-coding RNAs, or by affecting regulation of gene activity (influencing transcript abundance or gene splicing, for example)[87]. Because a majority of cancer-related variants resides in non-coding regions, most experience to date comes from examining the role of risk SNPs in gene regulation. It is well established that certain germline variants, referred to as expression quantitative trait loci (eQTLs), can affect transcription locally or at considerable genomic distances[88-91]. Post-GWAS analysis of risk variant function is often based on the premise that the non-coding disease risk loci act as eQTLs. This appears to be the case for a breast cancer risk polymorphism discovered by GWAS residing in an intron of the gene FGFR2; homozygotes for the risk allele exhibit increased FGFR2 expression[92]. Interrogation of two cancer risk loci discovered by GWAS-8q24 and 10q11- illustrate the ways in which the mechanisms of inherited risk may be revealed.

Independent GWAS converged on SNP rs10993994 at chromosome 10q11 as highly associated with prostate cancer risk[52,59]. Fine mapping across the risk locus demonstrated that rs10993994 is the variant most strongly associated with risk[93]. The SNP resides in the promoter region of the MSMB gene, which encodes PSP94, a purported biomarker for prostate cancer. Decreased levels of PSP94 are associated with prostate cancer risk[94]. Electromobility shift assays and luciferase transfection studies showed that genotype at the locus influences MSMB activity[93,95]. Associations between genotype at rs10993994 and expression of nearby genes were measured in 84 human prostate tissue specimens[96]. The 10q11 risk allele was associated with decreased MSMB RNA abundance. The allele is also associated with decreased MSMB expression in urine, a proposed biomarker[97]. Strikingly, it was also highly associated in prostate tissue with increased RNA abundance at NCOA4, an androgen receptor coactivator residing 10 kilobases downstream of MSMB[96]. The activity at

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

this locus demonstrates how alleles may be associated with expression of nearby and/or distal candidate genes.

Notably, all 8q24 risk polymorphisms reside in intergenic, non-coding regions of the genome. The nearest genes to 8q24 risk loci is MYC, located more than 250 kilobases (kb) from the nearest risk SNP. As with the 10q11 prostate cancer risk allele and the FGFR2 breast cancer allele, it was hypothesized that the 8q24 risk loci are eQTLs. However, there does not appear to be an association between risk allele status and MYC expression[98]. Evidence has accumulated implicating 8q24 colon and prostate cancer risk alleles in the activity of genetic enhancers, elements capable of affecting expression of one or more genes from long-range[99,100]. Further evidence suggests that these enhancer elements are in long-range contact with MYC across hundreds of kb[100-103]. These findings suggest involvement by MYC in prostate cancer risk and may provide a paradigm for investigating other risk regions. The discoveries at 8q24 demonstrate the potential for GWAS results to elucidate the underpinnings of inherited risk.

CLINICAL UTILITY OF INHERITED RISK MARKERS

GWAS have revealed bona fide cancer risk factors. The variants also can lend insight into cancer biology. However, it is less clear whether the newly discovered risk marker have clinical utility. Clinical utility is a measure of the potential benefits of a test relative to its risks and costs. A biomarker for cancer risk should be affordable, accurate, and easily interpretable by health care providers and patients[104]. As predictors of risk, germline genetic markers have a natural advantage over many current biomarkers because they are static; they are ever-present and do not fluctuate with time or clinical condition. For example, markers such as PSA only reach clinical attention when prostate cancer has presumably already developed, whereas inherited risk SNPs are testable at any time prior to the presence of disease. These considerations must be balanced, however, against the quantity of information gained by the addition of genetic risk factors.

Much work in this area has involved prostate cancer. Prostate cancer is the second leading cause of cancer-related death among men in the U.S. and is highly curable if detected early. PSA is widely used as a biomarker for disease, but is imperfect[105]. It is not adequately specific for most men with abnormal levels and does a poor job of distinguishing aggressive from indolent disease. Other variables, such as family history and ethnicity are predictive, but not clinically useful. As GWAS reported polymorphisms associated with prostate cancer risk, several groups investigated the possibility that these markers could help identify men with disease more accurately and/or distinguish aggressive from non-aggressive disease.

Zheng et al demonstrated that risk of prostate cancer correlates with increasing number of risk alleles[92]. For men with a family history of prostate cancer who carry five risk SNPs plus family history), the odds ratio was 9.46 for developing prostate cancer compared to men with no risk factors. However, this category represents a small proportion of patients. When the authors constructed receiver operating characteristic (ROC) curves to measure the sensitivity and specificity of genomic profiling, the area under the curve (AUC) was 0.63 for a profile involving age, geographic region, family history and genotype of the five risk SNPs, a very modest improvement over an AUC of 0.61 without genotypic information, both below 0.8, the threshold generally considered to represent accurate prediction[106]. Adding SNPs to the model as they have been discovered has not appreciably improved the ROC curves[107]. Another series used a panel of four prostate cancer-related SNPs and added them to a set of known risk variables. The AUC improved from 0.72 to 0.74 with the addition of the SNP data.

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

Similarly modest results have been reported for predicting breast cancer. The Gail model, incorporating age, family history, reproductive history, and breast biopsy history, has been used for decades to estimate risk of invasive breast cancer. As breast risk SNPs have been discovered by GWAS, they have been added to the model to determine whether they improve prediction. Seven breast cancer SNPs reported in 2007 and 2008 were genotyped in one series of over 1600 cases and 1600 controls, and AUC improved to 0.594 compared with 0.557 for Gail risk alone[108]. A second, larger series of 5590 cases and 5998 controls examined the addition of 10 established breast cancer risk SNPs[109]. In this study, the Gail model demonstrated an AUC of 0.58. The addition of the 10 SNPs improved the AUC to 0.62.

A major determinant limiting the potential of genetic profiling for common, complex cancers are their polygenic nature. The contribution from each risk variant is modest and as a result a substantial majority of subjects will be at average risk of disease. In Zheng et al, 85.5% of prostate cancer cases and 86.2% of controls genotyped harbored 1-3 risk markers[92]. As a result, a finding a proper “cutoff” for declaring high or low risk becomes impossible without accepting unreasonable numbers of false positives or false negatives. Given minor allele frequencies of 5-30% and odds ratios for disease association of 1.25-1.50 (parameters similar to those reported for prostate cancer GWAS), it is estimated that between 23 and 320 markers would be needed to achieve an AUC of 0.8[104].

It has been proposed, then, that genetic profiling based on GWAS findings should be aimed at finding those at highest risk of disease. Yet, even for this small subset, <5% of men, the clinical benefit of profiling is not clear. Those considered high risk, for example, could be recommended for earlier PSA screening, as is often advised for those with a family history disease. There are no known modifiable risk factors for prostate cancer that could be recommended to high-risk patients, though chemoprevention strategies, such as finasteride, may be reasonable. Finasteride, a 5-alpha reductase inhibitor, has been demonstrated to prevent or delay the onset of prostate cancer in men over age 55[110]. Future clinical trials could target carriers of multiple risk alleles and determine whether this population particularly benefits from chemoprevention.

If genetic profiling were able to predict clinical course, clinicians would have a useful tool to help guide treatment decisions. Germline genetic markers that accurately distinguish aggressive from non-aggressive disease could have a significant impact on patient care. In the case of prostate cancer, there does appear to be a genetic component to outcome[111], and several groups have examined whether the prostate cancer risk SNPs discovered to date predict outcome.

An example involves a risk locus discovered by GWAS at chromosome 19q13. The risk SNP, rs2735839 resides in the intergenic region of chromosome 19q13, approximately 600 base pairs downstream of the 3’UTR of KLK3, which encodes PSA. In several series rs2735839 was significantly associated aggressive disease, including prostate cancer-specific mortality[87,112-114]. Interestingly, the risk allele has been associated with *less* aggressive disease. In 2011, a GWAS of PSA level in a population of non-prostate cancer patients identified this exact allele-the protective allele for prostate cancer - as associated with high PSA, raising the possibility that the association with prostate cancer outcome is result of ascertainment bias. In this scenario, those inheriting the risk allele would have a slightly higher PSA and would be more likely to receive a referral for biopsy. A man carrying the non-risk allele will be diagnosed later, which may have significant downstream consequences. It may be appropriate to base PSA cutoff for referral for biopsy, in part, on genotype. Further work is needed before such a personalized approach would be recommended, and the region is complex – recent fine mapping and functional work suggest

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

that inherited variants in the region could contribute to both prostate carcinogenesis as well as PSA level[62,74,115].

Ultimately, for all genetically complex cancers, GWAS for aggressive versus non-aggressive disease may be the most effective means of identifying inherited markers associated with clinically relevant subtypes.

ON THE HORIZON

GWAS data generated to date do not reflect the full complement of factors associated with inherited risk. Identification of this missing heritability will be the focus of the next generation of GWAS. Future studies will determine if rare variants, polymorphisms with lower allele frequencies, account for a substantial portion of inherited risk. The 1000 Genomes Project, larger study cohorts and the decreasing cost of genome sequencing will enable interrogation of these SNPs and ultimately more effective profiling of patients. In addition, there are other types of genetic polymorphisms, such as copy number variants, that are becoming increasingly amenable to testing. As the full spectrum of alleles associated with disease comes into focus, we can anticipate a more profound understanding of cancer pathogenesis, which may ultimately result in improvements in prevention, diagnosis and treatment of cancer.

REFERENCES

1. Steinberg GD, Carter BS, Beaty TH, Childs B, Walsh PC. Family history and the risk of prostate cancer. Prostate. 1990; 17:337–47. [PubMed: 2251225]
2. Spitz MR, Currier RD, Fueger JJ, Babaian RJ, Newell GR. Familial patterns of prostate cancer: a case-control analysis. J Urol. 1991; 146:1305–7. [PubMed: 1942282]
3. Aprikian AG, et al. Family history and the risk of prostatic carcinoma in a high risk group of urological patients. J Urol. 1995; 154:404–6. [PubMed: 7541856]
4. Hayes RB, et al. Prostate cancer risk in U.S. blacks and whites with a family history of cancer. Int J Cancer. 1995; 60:361–4. [PubMed: 7829245]
5. Isaacs SD, Kiemeney LA, Baffoe-Bonnie A, Beaty TH, Walsh PC. Risk of cancer in relatives of prostate cancer probands. J Natl Cancer Inst. 1995; 87:991–6. [PubMed: 7629886]
6. Gronberg H, Damber L, Damber JE. Familial prostate cancer in Sweden. A nationwide register cohort study. Cancer. 1996; 77:138–43.
7. Lesko SM, Rosenberg L, Shapiro S. Family history and prostate cancer risk. Am J Epidemiol. 1996; 144:1041–7. [PubMed: 8942435]
8. Cerhan JR, et al. Family history and prostate cancer risk in a population-based cohort of Iowa men. Cancer Epidemiol Biomarkers Prev. 1999; 8:53–60. [PubMed: 9950240]
9. Hemminki K, Czene K. Age specific and attributable risks of familial prostate carcinoma from the family-cancer database. Cancer. 2002; 95:1346–53. [PubMed: 12216104]
10. Goldgar DE, Easton DF, Cannon-Albright LA, Skolnick MH. Systematic population-based assessment of cancer risk in first-degree relatives of cancer probands. J Natl Cancer Inst. 1994; 86:1600–8. [PubMed: 7932824]
11. Hemminki K, Sundquist J, Bermejo JL. How common is familial cancer? Ann Oncol. 2008; 19:163–7. [PubMed: 17804474]
12. Zeegers MP, Jellema A, Ostrer H. Empiric risk of prostate carcinoma for relatives of patients with prostate carcinoma: a meta-analysis. Cancer. 2003; 97:1894–903. [PubMed: 12673715]
13. Chen YC, Page JH, Chen R, Giovannucci E. Family history of prostate and breast cancer and the risk of prostate cancer in the PSA era. Prostate. 2008; 68:1582–91. [PubMed: 18646000]
14. Brandt A, Bermejo JL, Sundquist J, Hemminki K. Age-specific risk of incident prostate cancer and risk of death from prostate cancer defined by the number of affected family members. Eur Urol. 2010; 58:275–80. [PubMed: 20171779]

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

15. Familial breast cancer: collaborative reanalysis of individual data from 52 epidemiological studies including 58,209 women with breast cancer and 101,986 women without the disease. Lancet. 2001; 358:1389–99. [PubMed: 11705483]
16. Fuchs CS, et al. A prospective study of family history and the risk of colorectal cancer. N Engl J Med. 1994; 331:1669–74. [PubMed: 7969357]
17. Page WF, Braun MM, Partin AW, Caporaso N, Walsh P. Heredity and prostate cancer: a study of World War II veteran twins. Prostate. 1997; 33:240–5. [PubMed: 9397195]
18. Lichtenstein P, et al. Environmental and heritable factors in the causation of cancer-analyses of cohorts of twins from Sweden, Denmark, and Finland. N Engl J Med. 2000; 343:7885.
19. Miki Y, et al. A strong candidate for the breast and ovarian cancer susceptibility gene BRCA1. Science. 1994; 266:66–71. [PubMed: 7545954]
20. Wooster R, et al. Identification of the breast cancer susceptibility gene BRCA2. Nature. 1995; 378:789–92. [PubMed: 8524414]
21. Hussussian CJ, et al. Germline p16 mutations in familial melanoma. Nat Genet. 1994; 8:15–21. [PubMed: 7987387]
22. Leppert M, et al. Genetic analysis of an inherited predisposition to colon cancer in a family with a variable number of adenomatous polyps. N Engl J Med. 1990; 322:904–8. [PubMed: 2156161]
23. Malkin D, et al. Germ line p53 mutations in a familial syndrome of breast cancer, sarcomas, and other neoplasms. Science. 1990; 250:1233–8. [PubMed: 1978757]
24. Smith JR, et al. Major susceptibility locus for prostate cancer on chromosome 1 suggested by a genome-wide search. Science. 1996; 274:1371–4. [PubMed: 8910276]
25. Cunningham JM, et al. Genome linkage screen for prostate cancer susceptibility loci: results from the Mayo Clinic Familial Prostate Cancer Study. Prostate. 2003; 57:335–46. [PubMed: 14601030]
26. Janer M, et al. Genomic scan of 254 hereditary prostate cancer families. Prostate. 2003; 57:309–19. [PubMed: 14601027]
27. Chang BL, et al. Genome-wide screen for prostate cancer susceptibility genes in men with clinically significant disease. Prostate. 2005; 64:356–61. [PubMed: 15754351]
28. Stanford JL, et al. Prostate cancer and genetic susceptibility: a genome scan incorporating disease aggressiveness. Prostate. 2006; 66:317–25. [PubMed: 16245279]
29. Camp NJ, Farnham JM, Cannon-Albright LA. Localization of a prostate cancer predisposition gene to an 880-kb region on chromosome 22q12.3 in Utah high-risk pedigrees. Cancer Res. 2006; 66:10205–12. [PubMed: 17047086]
30. Johanneson B, 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
31. Berthon P, et al. Predisposing gene for early-onset prostate cancer, localized on chromosome 1q42.2-43. Am J Hum Genet. 1998; 62:1416–24. [PubMed: 9585607]
32. Xu J, et al. Evidence for a prostate cancer susceptibility locus on the X chromosome. Nat Genet. 1998; 20:175–9. [PubMed: 9771711]
33. Gibbs M, et al. Evidence for a rare prostate cancer-susceptibility locus at chromosome 1p36. Am J Hum Genet. 1999; 64:776–87. [PubMed: 10053012]
34. Stanford JL, et al. Dense genome-wide SNP linkage scan in 301 hereditary prostate cancer families identifies multiple regions with suggestive evidence for linkage. Hum Mol Genet. 2009; 18:1839–48. [PubMed: 19251732]
35. Berry R, et al. Evidence for a prostate cancer-susceptibility locus on chromosome 20. Am J Hum Genet. 2000; 67:82–91. [PubMed: 10820130]
36. Lange EM, et al. Genome-wide linkage scan for prostate cancer susceptibility from the University of Michigan Prostate Cancer Genetics Project: suggestive evidence for linkage at 16q23. Prostate. 2009; 69:385–91. [PubMed: 19035517]
37. Eeles RA, et al. Linkage analysis of chromosome 1q markers in 136 prostate cancer families. The Cancer Research Campaign/British Prostate Group U.K. Familial Prostate Cancer Study Collaborators. Am J Hum Genet. 1998; 62:653–8.

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

38. Berry R, et al. Linkage analyses at the chromosome 1 loci 1q24-25 (HPC1), 1q42.2-43 (PCAP), and 1p36 (CAPB) in families with hereditary prostate cancer. Am J Hum Genet. 2000; 66:539–46. [PubMed: 10677314]
39. Xu J. Combined analysis of hereditary prostate cancer linkage to 1q24-25: results from 772 hereditary prostate cancer families from the International Consortium for Prostate Cancer Genetics. Am J Hum Genet. 2000; 66:945–57. [PubMed: 10712209]
40. Welcsh PL, King MC. BRCA1 and BRCA2 and the genetics of breast and ovarian cancer. Hum Mol Genet. 2001; 10:705–13. [PubMed: 11257103]
41. Rustgi AK. The genetics of hereditary colon cancer. Genes Dev. 2007; 21:2525–38. [PubMed: 17938238]
42. Garcia-Closas M, et al. NAT2 slow acetylation, GSTM1 null genotype, and risk of bladder cancer: results from the Spanish Bladder Cancer Study and meta-analyses. Lancet. 2005; 366:649–59. [PubMed: 16112301]
43. Garcia-Closas M, et al. A single nucleotide polymorphism tags variation in the arylamine N-acetyltransferase 2 phenotype in populations of European background. Pharmacogenet Genomics. 2011; 21:231–6. [PubMed: 20739907]
44. Lindstrom S, et al. A large study of androgen receptor germline variants and their relation to sex hormone levels and prostate cancer risk. Results from the National Cancer Institute Breast and Prostate Cancer Cohort Consortium. J Clin Endocrinol Metab. 2010; 95:E121–7.
45. The International HapMap Project. Nature. 2003; 426:789–96. [PubMed: 14685227]
46. A haplotype map of the human genome. Nature. 2005; 437:1299–320. [PubMed: 16255080]
47. Frazer KA, et al. A second generation human haplotype map of over 3.1 million SNPs. Nature. 2007; 449:851–61. [PubMed: 17943122]
48. Kruglyak L. The road to genome-wide association studies. Nat Rev Genet. 2008; 9:314–8. [PubMed: 18283274]
49. Wang WY, Barratt BJ, Clayton DG, Todd JA. Genome-wide association studies: theoretical and practical concerns. Nat Rev Genet. 2005; 6:109–18. [PubMed: 15716907]
50. Al Olama AA, et al. Multiple loci on 8q24 associated with prostate cancer susceptibility. Nat Genet. 2009; 41:1058–60. [PubMed: 19767752]
51. Amundadottir LT, et al. A common variant associated with prostate cancer in European and African populations. Nat Genet. 2006; 38:652–8. [PubMed: 16682969]
52. Eeles RA, et al. Multiple newly identified loci associated with prostate cancer susceptibility. Nat Genet. 2008; 40:316–21. [PubMed: 18264097]
53. Eeles RA, et al. Identification of seven new prostate cancer susceptibility loci through a genome-wide association study. Nat Genet. 2009; 41:1116–21. [PubMed: 19767753]
54. Gudmundsson J, et al. Genome-wide association study identifies a second prostate cancer susceptibility variant at 8q24. Nat Genet. 2007; 39:631–7. [PubMed: 17401366]
55. Gudmundsson J, et al. Two variants on chromosome 17 confer prostate cancer risk, and the one in TCF2 protects against type 2 diabetes. Nat Genet. 2007; 39:977–83. [PubMed: 17603485]
56. Gudmundsson J, et al. Common sequence variants on 2p15 and Xp11.22 confer susceptibility to prostate cancer. Nat Genet. 2008; 40:281–3. [PubMed: 18264098]
57. Gudmundsson J, et al. Genome-wide association and replication studies identify four variants associated with prostate cancer susceptibility. Nat Genet. 2009; 41:1122–6. [PubMed: 19767754]
58. Sun J, et al. Evidence for two independent prostate cancer risk-associated loci in the HNF1B gene at 17q12. Nat Genet. 2008; 40:1153–5. [PubMed: 18758462]
59. Thomas G, et al. Multiple loci identified in a genome-wide association study of prostate cancer. Nat Genet. 2008; 40:310–5. [PubMed: 18264096]
60. Yeager M, et al. Genome-wide association study of prostate cancer identifies a second risk locus at 8q24. Nat Genet. 2007; 39:645–9. [PubMed: 17401363]
61. Yeager M, et al. Identification of a new prostate cancer susceptibility locus on chromosome 8q24. Nat Genet. 2009; 41:1055–7. [PubMed: 19767755]
62. Kote-Jarai Z, et al. Identification of a novel prostate cancer susceptibility variant in the KLK3 gene transcript. Hum Genet. 2011; 129:687–94. [PubMed: 21465221]

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

63. Chanock SJ, et al. Replicating genotype-phenotype associations. Nature. 2007; 447:655–60. [PubMed: 17554299]
64. Hsiung CA, et al. The 5p15.33 locus is associated with risk of lung adenocarcinoma in never-smoking females in Asia. PLoS Genet. 2010; 6
65. Thomas D. Gene--environment-wide association studies: emerging approaches. Nat Rev Genet. 2010; 11:259–72. [PubMed: 20212493]
66. Freedman ML, et al. Admixture mapping identifies 8q24 as a prostate cancer risk locus in African-American men. Proc Natl Acad Sci U S A. 2006; 103:14068–73. [PubMed: 16945910]
67. Haiman CA, et al. Multiple regions within 8q24 independently affect risk for prostate cancer. Nat Genet. 2007; 39:638–44. [PubMed: 17401364]
68. Tomlinson I, et al. A genome-wide association scan of tag SNPs identifies a susceptibility variant for colorectal cancer at 8q24.21. Nat Genet. 2007; 39:984–8. [PubMed: 17618284]
69. Zanke BW, et al. Genome-wide association scan identifies a colorectal cancer susceptibility locus on chromosome 8q24. Nat Genet. 2007; 39:989–94. [PubMed: 17618283]
70. Kiemeney LA, et al. Sequence variant on 8q24 confers susceptibility to urinary bladder cancer. Nat Genet. 2008; 40:1307–12. [PubMed: 18794855]
71. Crowther-Swanepoel D, et al. Common variants at 2q37.3, 8q24.21, 15q21.3 and 16q24.1 influence chronic lymphocytic leukemia risk. Nat Genet. 2010; 42:132–6. [PubMed: 20062064]
72. Kote-Jarai Z, et al. Seven prostate cancer susceptibility loci identified by a multi-stage genome-wide association study. Nat Genet. 2011; 43:785–91. [PubMed: 21743467]
73. Shete S, et al. Genome-wide association study identifies five susceptibility loci for glioma. Nat Genet. 2009; 41:899–904. [PubMed: 19578367]
74. Petersen GM, et al. A genome-wide association study identifies pancreatic cancer susceptibility loci on chromosomes 13q22.1, 1q32.1 and 5p15.33. Nat Genet. 2010; 42:224–8. [PubMed: 20101243]
75. Rothman N, et al. A multi-stage genome-wide association study of bladder cancer identifies multiple susceptibility loci. Nat Genet. 2010; 42:978–84. [PubMed: 20972438]
76. Landi MT, et al. A genome-wide association study of lung cancer identifies a region of chromosome 5p15 associated with risk for adenocarcinoma. Am J Hum Genet. 2009; 85:679–91. [PubMed: 19836008]
77. Wang Y, et al. Common 5p15.33 and 6p21.33 variants influence lung cancer risk. Nat Genet. 2008; 40:1407–9. [PubMed: 18978787]
78. Rafnar T, et al. Sequence variants at the TERT-CLPTM1L locus associate with many cancer types. Nat Genet. 2009; 41:221–7. [PubMed: 19151717]
79. Yamaguchi H, et al. Mutations in TERT, the gene for telomerase reverse transcriptase, in aplastic anemia. N Engl J Med. 2005; 352:1413–24. [PubMed: 15814878]
80. Calado RT, et al. Constitutional hypomorphic telomerase mutations in patients with acute myeloid leukemia. Proc Natl Acad Sci U S A. 2009; 106:1187–92. [PubMed: 19147845]
81. Garcia-Closas M, et al. Heterogeneity of breast cancer associations with five susceptibility loci by clinical and pathological characteristics. PLoS Genet. 2008; 4:e1000054. [PubMed: 18437204]
82. Haiman CA, et al. Genome-wide association study of prostate cancer in men of African ancestry identifies a susceptibility locus at 17q21. Nat Genet. 2011; 43:570–3. [PubMed: 21602798]
83. Yeager M, et al. Comprehensive resequence analysis of a 136 kb region of human chromosome 8q24 associated with prostate and colon cancers. Hum Genet. 2008; 124:161–70. [PubMed: 18704501]
84. Chung CC, et al. Fine mapping of a region of chromosome 11q13 reveals multiple independent loci associated with risk of prostate cancer. Hum Mol Genet. 2011; 20:2869–78. [PubMed: 21531787]
85. Hindorff LA, et al. Potential etiologic and functional implications of genome-wide association loci for human diseases and traits. Proc Natl Acad Sci U S A. 2009; 106:9362–7. [PubMed: 19474294]
86. Hirschhorn JN. Genomewide association studies--illuminating biologic pathways. N Engl J Med. 2009; 360:1699–701. [PubMed: 19369661]

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

87. Pomerantz MM, et al. Association of prostate cancer risk Loci with disease aggressiveness and prostate cancer-specific mortality. Cancer Prev Res (Phila). 2011; 4:719–28. [PubMed: 21367958]
88. Morley M, et al. Genetic analysis of genome-wide variation in human gene expression. Nature. 2004; 430:743–7. [PubMed: 15269782]
89. Stranger BE, et al. Genome-wide associations of gene expression variation in humans. PLoS Genet. 2005; 1:e78. [PubMed: 16362079]
90. Schadt EE, et al. Genetics of gene expression surveyed in maize, mouse and man. Nature. 2003; 422:297–302. [PubMed: 12646919]
91. Cheung VG, Spielman RS. Genetics of human gene expression: mapping DNA variants that influence gene expression. Nat Rev Genet. 2009; 10:595–604. [PubMed: 19636342]
92. Zheng SL, et al. Cumulative association of five genetic variants with prostate cancer. N Engl J Med. 2008; 358:910–9. [PubMed: 18199855]
93. Lou H, et al. Fine mapping and functional analysis of a common variant in MSMB on chromosome 10q11.2 associated with prostate cancer susceptibility. Proc Natl Acad Sci U S A. 2009; 106:7933–8. [PubMed: 19383797]
94. Nam RK, et al. A novel serum marker, total prostate secretory protein of 94 amino acids, improves prostate cancer detection and helps identify high grade cancers at diagnosis. J Urol. 2006; 175:1291–7. [PubMed: 16515983]
95. Chang BL, et al. Fine mapping association study and functional analysis implicate a SNP in MSMB at 10q11 as a causal variant for prostate cancer risk. Hum Mol Genet. 2009; 18:1368–75. [PubMed: 19153072]
96. Pomerantz MM, et al. Analysis of the 10q11 cancer risk locus implicates MSMB and NCOA4 in human prostate tumorigenesis. PLoS Genet. 2010; 6:e1001204. [PubMed: 21085629]
97. Whitaker HC, et al. The rs10993994 risk allele for prostate cancer results in clinically relevant changes in microseminoprotein-beta expression in tissue and urine. PLoS One. 2010; 5:e13363. [PubMed: 20967219]
98. Pomerantz MM, et al. Evaluation of the 8q24 prostate cancer risk locus and MYC expression. Cancer Res. 2009; 69:5568–74. [PubMed: 19549893]
99. Jia L, et al. Functional enhancers at the gene-poor 8q24 cancer-linked locus. PLoS Genet. 2009; 5:e1000597. [PubMed: 19680443]
100. Pomerantz MM, et al. The 8q24 cancer risk variant rs6983267 shows long-range interaction with MYC in colorectal cancer. Nat Genet. 2009; 41:882–4. [PubMed: 19561607]
101. Ahmadiyeh N, et al. 8q24 prostate, breast, and colon cancer risk loci show tissue-specific long-range interaction with MYC. Proc Natl Acad Sci U S A. 2010; 107:9742–6. [PubMed: 20453196]
102. Wright JB, Brown SJ, Cole MD. Upregulation of c-MYC in cis through a large chromatin loop linked to a cancer risk-associated single-nucleotide polymorphism in colorectal cancer cells. Mol Cell Biol. 2010; 30:1411–20. [PubMed: 20065031]
103. Yochum GS, Sherrick CM, Macpartlin M, Goodman RH. A beta-catenin/TCF-coordinated chromatin loop at MYC integrates 5′ and 3′ Wnt responsive enhancers. Proc Natl Acad Sci U S A. 2010; 107:145–50. [PubMed: 19966299]
104. Kraft P, et al. Beyond odds ratios--communicating disease risk based on genetic profiles. Nat Rev Genet. 2009; 10:264–9. [PubMed: 19238176]
105. Andriole GL, et al. Mortality results from a randomized prostate-cancer screening trial. N Engl J Med. 2009; 360:1310–9. [PubMed: 19297565]
106. Janssens AC, et al. A critical appraisal of the scientific basis of commercial genomic profiles used to assess health risks and personalize health interventions. Am J Hum Genet. 2008; 82:593–9. [PubMed: 18319070]
107. Zheng SL, et al. Genetic variants and family history predict prostate cancer similar to prostate-specific antigen. Clin Cancer Res. 2009; 15:1105–11. [PubMed: 19188186]
108. Mealiffe ME, et al. Assessment of clinical validity of a breast cancer risk model combining genetic and clinical information. J Natl Cancer Inst. 2010; 102:1618–27. [PubMed: 20956782]

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

109. Wacholder S, et al. Performance of common genetic variants in breast-cancer risk models. N Engl J Med. 2010; 362:986–93. [PubMed: 20237344]

110. Thompson IM, et al. The influence of finasteride on the development of prostate cancer. N Engl J Med. 2003; 349:215–24. [PubMed: 12824459]

111. Hemminki K, Ji J, Forsti A, Sundquist J, Lenner P. Concordance of survival in family members with prostate cancer. J Clin Oncol. 2008; 26:1705–9. [PubMed: 18375899]

112. Gallagher DJ, et al. Susceptibility loci associated with prostate cancer progression and mortality. Clin Cancer Res. 2010; 16:2819–32. [PubMed: 20460480]

113. Kader AK, et al. Individual and cumulative effect of prostate cancer risk-associated variants on clinicopathologic variables in 5,895 prostate cancer patients. Prostate. 2009; 69:1195–205. [PubMed: 19434657]

114. Xu J, et al. Association of prostate cancer risk variants with clinicopathologic characteristics of the disease. Clin Cancer Res. 2008; 14:5819–24. [PubMed: 18794092]

115. Parikh H, et al. Fine mapping the KLK3 locus on chromosome 19q13.33 associated with prostate cancer susceptibility and PSA levels. Hum Genet. 2011; 129:675–85. [PubMed: 21318478]

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

Table 1

Comparison of Medelian and common, sporadic (non-Mendelian) cancers

	Medelian	non-Mendelian
Mode of inheritance	monogenic	polygenic
Incidence	+	+++
Effect size of each risk allele	+++	+
Minor allele frequency	+	++ to +++
Penetrance of risk allele	+++	+
Method used to map the trait	linkage analysis	GWAS