Heritability of prostate cancer: a tale of rare variants and common single nucleotide polymorphisms

Xifeng Wu, Jian Gu

Department of Epidemiology, The University of Texas MD Anderson Cancer Center, Houston, TX 77030, USA *Correspondence to:* Xifeng Wu. Department of Epidemiology, The University of Texas MD Anderson Cancer Center, Houston, TX 77030, USA. Email: xwu@mdanderson.org.

Submitted Apr 06, 2016. Accepted for publication Apr 13, 2016. doi: 10.21037/atm.2016.05.31 View this article at: http://dx.doi.org/10.21037/atm.2016.05.31

Prostate cancer (PCa) is the most common non-skin cancer men in developed countries. Despite years of research, no strong modifiable risk factor for PCa has been found. The two most significant cancer risk factors, smoking and obesity, do not appear to be strong risk factors for PCa. A recent systematic review and meta-analysis summarized the current literature of tobacco use and PCa mortality and incidence (1). This meta-analysis included 51 cohort studies (50,349 incident cases and 4,082,606 cohort participants) and found a dichotomized association between smoking and PCa risk. Current smoking was associated with an increased risk of PCa [rate ratio (RR): 1.06; 95% confidence interval (CI), 0.98-1.15] in studies completed in 1995 or earlier [before the prostate-specific antigen (PSA) screening era], and a reduced risk of PCa (RR: 0.84, 95% CI, 0.79-0.89) in studies completed afterward (after the wide spread of PSA screening). These data suggest that smoking may reduce the risk of indolent non-aggressive cancers that have predominated in more recent years, while promoting more aggressive cancers. Likewise, there also appears to be a dual effect of obesity on PCa incidence. A metaanalysis of multiple prospective studies found that obesity had a slightly protective effect in localized PCa (RR: 0.94, 95% CI, 0.91-0.97), but was associated with an increased incidence of advanced PCa (RR: 1.09, 95% CI, 1.02-1.16) for every five units of body mass index (BMI) increase (2).

Age is the most important risk factor for PCa: the lifetime The lifetime probability of being diagnosed with PCa was 0.3 (1 in 325) for those aged younger than 50 years, 2.1 (1 in 48) for 50–59 year old men, 5.8 (1 in 17) for 60–69 year old men, and 10.0 (1 in 10) for those aged 70 years and above (3). Another major risk factor for PCa is race: the incidence rate among African Americans is about

70% higher than that among whites (4). In addition, genetic susceptibility plays an important role in PCa etiology. Earlier studies on familial aggregation of cancer clearly showed familial clustering of PCa. A meta-analysis of 33 independent studies demonstrates a pooled RR of 2.48 (95% CI, 2.25–2.74) for men with a first-degree family history (i.e., affected father or brother) and the RR increases to 4.39 (95% CI, 2.61-7.39) for those with two or more affected first-degree family members (5). The large classic Nordic Twin study recently estimated a heritability of about 58% for PCa, much higher than all the other common nonskin cancers (e.g., lung cancer, 18%; breast cancer, 31%; and colon cancer, 15%), whereas shared environment has negligible effect (0%) on PCa development, compared to 24%, 16%, and 16% for lung, breast, and colon cancer, respectively (6,7).

Given the high heritability of PCa, there have been enormous efforts on identifying inherited genetic susceptibility loci for PCa. Many genome-wide association studies (GWAS) have been performed across different ethnicities and over 100 common, low-penetrance PCa predisposing variants have currently been identified through GWAS (8-25). These common single nucleotide polymorphisms (SNPs) have minor allele frequencies (MAF) of at least 5%, and have modest effect sizes [odds ratio (OR) <1.5], and were estimated to explain about 33% of the familial risk (25), leaving the majority of risk unexplained. With the extremely large sample sizes of the published GWAS and meta-analysis of GWAS of PCa, the remaining "low-hanging fruits" of common SNPs are expected to have effect sizes of <1.1 and not likely to increase much of the relative contribution of common SNPs to PCa familial risk. An intriguing hypothesis is that rare

Page 2 of 4

variants (MAF <1%) may explain the substantial "missing heritability" of PCa and other complex diseases. The commonly used SNP arrays in GWAS have very limited representation of variants of this low MAF and GWAS is limited in detecting the association between rare variants and cancer risks. The association of rare variants with PCa has been largely unexplored.

The first identified rare variants that confer increased PCa risk were BRCA2 mutations: BRCA2 mutation carriers exhibited a nearly 5-fold increased risk of PCa and the risk increased to about 7-fold for early-onset PCa (26,27). More recently, a missense mutation (G84E) in HOXB13 was found to be associated with increased PCa risks (28) and a meta-analysis of 24,213 cases and 73, 631 controls of European descendants revealed an OR of 4.07 (95% CI, 3.05–5.45) (29). These data provide supporting for the evidence of rare variants in PCa etiology. However, due to the rare frequencies of these variants and the scarcity of validated such rare variants, it is still unclear to what degree rare variants can explain the "missing heritability". There have been ongoing efforts to identify such rare variants through systemic sequencing rather than limited candidate gene approach, but without success to date. Rand et al. (30) recently performed whole-exome sequencing of 2,165 PCa cases and 2,034 controls of African ancestry at a mean coverage of 10.1× and identified 395,220 coding variants down to 0.05% frequency in 16,751 genes, but did not observe exome-wide significant associations with PCa risk after correcting for multiple testing in single variant or gene-level testing. The data from this first whole-exome sequencing study of PCa did not provide strong support for the hypothesis that rare, non-synonymous coding variants down to 0.5-1.0% frequency have large effects on PCa risk in men of African ancestry. However, this study only examined protein-coding sequence that comprises 1-2% of the whole genome. The GWAS-identified cancer predisposing SNPs are predominantly located in noncoding regions. In analogy, rare cancer predisposing variants may also locate to non-coding regions. One known example is a rare variant (rs183373024, risk allele frequency 0.54% in controls) at a previously known PCa susceptibility region 8q24 that confers an OR of 2.90 (95% CI, 2.44-3.44) in a study of 10,007 cases and 62,027 controls of European origin (31). This variant maps to an enhancer region to which androgen receptor (AR) and transcription factor FoxA1 are bound and the predicted binding specificity for FoxA1 is greatly reduced by the introduction of the risk allele of this variant, which may explain its association

with PCa risk (32). For non-coding rare variant, wholeexome sequencing is not a suitable approach. Highcoverage whole-genome sequencing (WGS) is the ultimate way to tackle this problem. However, given the technical, analytical, and cost issues of WGS, it is not feasible to perform WGS in a large number of cases and controls for systemic investigations of rare variants and cancer risk at current stage. Targeted resequencing of known cancer susceptibility regions is a valuable option to identify such rare cancer susceptibility variants.

In a recently published manuscript in Nature Genetics, Mancuso et al. (33) reported targeted sequencing of 63 known PCa risk regions in a multi-ancestry study of 9,237 men and assessed the contribution of rare variants to PCa risk. They showed that the variance explained by all the sequenced variants is significantly larger than the variance explained by known GWAS variants at the same loci. In addition, they found evidence of genetic heterogeneity by ancestry in risk for PCa. More importantly, they uses variance-components methods to partition the SNP heritability across different variant frequency classes and found that a large amount of SNP heritability comes from the rare variant class in men of African ancestry: specifically, variants with $0.1\% \le$ MAF <1% explain a point estimate of 0.12 of variance PCa risk as compared to an estimate of 0.17 for variants with MAF $\geq 1\%$. This latter statistics was provocative, indicating that rare variants (MAF of 0.1-1%) explain approximately 42% of the variance contributed by all variants (including rare variants and common SNPs) with MAF of 0.1-50%. Previously, although it was hypothesized that rare variants may explain some "missing heritability" of common diseases, the 42% contributed by rare variants to a common disease like PCa exceeds general expectations given the low frequency of rare variants and the lack of confirmed rare variants predisposing to common diseases. It is even more striking that this 42% may be an underestimate since this study only sequenced known regions of PCa susceptibility that were derived from GWAS. If unbiased WGS were performed, the percentage of contribution of rare variants to PCa susceptibility may be even higher. It should be pointed out that variants with a MAF <1% accounts for only a small fraction of genetic variation in the population. This estimated 42% contribution by rare variants to PCa risk accounts for an order of magnitude larger than the heritability of PCa. The only possible explanation for this provocative number is that newly arising PCa predisposing variants are often subject to natural selection pressure that prevents them from becoming common, i.e., natural selection

Annals of Translational Medicine, Vol 4, No 10 May 2016

has driven down the frequency of many PCa risk alleles.

This study only observed significant contribution of rare variation to the heritability of PCa in African Americans, but not in European, Japanese, and Latino Americans. Han et al. (34) recently examined 82 common PCa predisposing SNPs identified by GWAS in European and Asian populations in 4,853 PCa cases and 4,678 controls of African ancestry. Of the 82 variants, 68 (83%) had effects that were directionally consistent in their association with PCa risk, indicating common functional alleles of SNPs that are shared across populations. It is known that African population carry a much larger number of rare variants than European and Asian populations. Whether this observation in the discrepancy of rare PCa predisposition variants in different populations is an artifact due to smaller sample sizes of other populations (less than half of the sample size of African Americans) in this study or reflects true differences in the genetic architecture of different ethnicities remains to be determined.

A caveat of this study by Mancuso et al. (33) that estimated such a high contribution to PCa risk by rare variants is the statistically complex analyses the author performed. Assaying SNP heritability using variance components makes a number of simplifying assumptions and there might be hidden biases that were not appreciated. The identity of such rare PCa predisposition variants, which should be a large number, is unknown. High-coverage targeted sequencing and WGS in larger samples, which will provide sufficient statistical power to allow a direct variantby-variant analysis of rare variants in the MAF range of 0.1-1%, is warranted to find the identity of rare PCa susceptibility variants that together contribute to almost half of the genetic heritability of PCa. Only by then will we be completely confident that rare variants make comparable contribution to the genetic heritability of PCa.

Acknowledgements

None.

Footnote

Provenance: This is a Guest Commentary commissioned by Guest Editor Xiongbing Zu, MD, PhD (Department of Urology, Xiangya Hospital, Central South University, Changsha, China).

Conflicts of Interest: The authors have no conflicts of interest to declare.

Commentary on: Mancuso N, Rohland N, Rand KA, *et al.* The contribution of rare variation to prostate cancer heritability. Nat Genet 2016;48:30-5.

References

- 1. Islami F, Moreira DM, Boffetta P, et al. A systematic review and meta-analysis of tobacco use and prostate cancer mortality and incidence in prospective cohort studies. Eur Urol 2014;66:1054-64.
- Discacciati A, Orsini N, Wolk A. Body mass index and incidence of localized and advanced prostate cancer--a dose-response meta-analysis of prospective studies. Ann Oncol 2012;23:1665-71.
- Siegel RL, Miller KD, Jemal A. Cancer statistics, 2016. CA Cancer J Clin 2016;66:7-30.
- DeSantis CE, Siegel RL, Sauer AG, et al. Cancer statistics for African Americans, 2016: Progress and opportunities in reducing racial disparities. CA Cancer J Clin 2016. [Epub ahead of print].
- Kiciński M, Vangronsveld J, Nawrot TS. An epidemiological reappraisal of the familial aggregation of prostate cancer: a meta-analysis. PLoS One 2011;6:e27130.
- 6. Hjelmborg JB, Scheike T, Holst K, et al. The heritability of prostate cancer in the Nordic Twin Study of Cancer. Cancer Epidemiol Biomarkers Prev 2014;23:2303-10.
- Mucci LA, Hjelmborg JB, Harris JR, et al. Familial Risk and Heritability of Cancer Among Twins in Nordic Countries. JAMA 2016;315:68-76.
- Amundadottir LT, Sulem P, Gudmundsson J, et al. A common variant associated with prostate cancer in European and African populations. Nat Genet 2006;38:652-8.
- Duggan D, Zheng SL, Knowlton M, et al. Two genomewide association studies of aggressive prostate cancer implicate putative prostate tumor suppressor gene DAB2IP. J Natl Cancer Inst 2007;99:1836-44.
- Gudmundsson J, Sulem P, Manolescu A, et al. Genomewide association study identifies a second prostate cancer susceptibility variant at 8q24. Nat Genet 2007;39:631-7.
- Gudmundsson J, Sulem P, Steinthorsdottir V, 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.
- Yeager M, Orr N, Hayes RB, et al. Genome-wide association study of prostate cancer identifies a second risk locus at 8q24. Nat Genet 2007;39:645-9.
- 13. Eeles RA, Kote-Jarai Z, Giles GG, et al. Multiple newly

Page 4 of 4

identified loci associated with prostate cancer susceptibility. Nat Genet 2008;40:316-21.

- Gudmundsson J, Sulem P, Rafnar T, et al. Common sequence variants on 2p15 and Xp11.22 confer susceptibility to prostate cancer. Nat Genet 2008;40:281-3.
- Sun J, Zheng SL, Wiklund F, et al. Evidence for two independent prostate cancer risk-associated loci in the HNF1B gene at 17q12. Nat Genet 2008;40:1153-5.
- Thomas G, Jacobs KB, Yeager M, et al. Multiple loci identified in a genome-wide association study of prostate cancer. Nat Genet 2008;40:310-5.
- Al Olama AA, Kote-Jarai Z, Giles GG, et al. Multiple loci on 8q24 associated with prostate cancer susceptibility. Nat Genet 2009;41:1058-60.
- Eeles RA, Kote-Jarai Z, Al Olama AA, et al. Identification of seven new prostate cancer susceptibility loci through a genome-wide association study. Nat Genet 2009;41:1116-21.
- Gudmundsson J, Sulem P, Gudbjartsson DF, et al. Genome-wide association and replication studies identify four variants associated with prostate cancer susceptibility. Nat Genet 2009;41:1122-6.
- Takata R, Akamatsu S, Kubo M, et al. Genome-wide association study identifies five new susceptibility loci for prostate cancer in the Japanese population. Nat Genet 2010;42:751-4.
- 21. Haiman CA, Chen GK, Blot WJ, 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.
- 22. Kote-Jarai Z, Olama AA, Giles GG, et al. Seven prostate cancer susceptibility loci identified by a multistage genome-wide association study. Nat Genet 2011;43:785-91.
- Schumacher FR, Berndt SI, Siddiq A, et al. Genomewide association study identifies new prostate cancer susceptibility loci. Hum Mol Genet 2011;20:3867-75.
- 24. Eeles RA, Olama AA, Benlloch S, et al. Identification

Cite this article as: Wu X, Gu J. Heritability of prostate cancer: a tale of rare variants and common single nucleotide polymorphisms. Ann Transl Med 2016;4(10):206. doi: 10.21037/ atm.2016.05.31

of 23 new prostate cancer susceptibility loci using the iCOGS custom genotyping array. Nat Genet 2013;45:385-91, 391e1-2.

- Al Olama AA, Kote-Jarai Z, Berndt SI, et al. A metaanalysis of 87,040 individuals identifies 23 new susceptibility loci for prostate cancer. Nat Genet 2014;46:1103-9.
- Breast Cancer Linkage Consortium. Cancer risks in BRCA2 mutation carriers. J Natl Cancer Inst 1999;91:1310-6.
- 27. Gayther SA, de Foy KA, Harrington P, et al. The frequency of germ-line mutations in the breast cancer predisposition genes BRCA1 and BRCA2 in familial prostate cancer. The Cancer Research Campaign/British Prostate Group United Kingdom Familial Prostate Cancer Study Collaborators. Cancer Res 2000;60:4513-8.
- Ewing CM, Ray AM, Lange EM, et al. Germline mutations in HOXB13 and prostate-cancer risk. N Engl J Med 2012;366:141-9.
- Shang Z, Zhu S, Zhang H, et al. Germline homeobox B13 (HOXB13) G84E mutation and prostate cancer risk in European descendants: a meta-analysis of 24,213 cases and 73, 631 controls. Eur Urol 2013;64:173-6.
- Rand KA, Rohland N, Tandon A, et al. Whole-exome sequencing of over 4100 men of African ancestry and prostate cancer risk. Hum Mol Genet 2016;25:371-81.
- Gudmundsson J, Sulem P, Gudbjartsson DF, et al. A study based on whole-genome sequencing yields a rare variant at 8q24 associated with prostate cancer. Nat Genet 2012;44:1326-9.
- 32. Hazelett DJ, Coetzee SG, Coetzee GA. A rare variant, which destroys a FoxA1 site at 8q24, is associated with prostate cancer risk. Cell Cycle 2013;12:379-80.
- Mancuso N, Rohland N, Rand KA, et al. The contribution of rare variation to prostate cancer heritability. Nat Genet 2016;48:30-5.
- 34. Han Y, Signorello LB, Strom SS, et al. Generalizability of established prostate cancer risk variants in men of African ancestry. Int J Cancer 2015;136:1210-7.