

Review Article

Diversity on demand: multi-ancestry meta-analysis improves genetic risk prediction in prostate cancer

Andres F Leon, Cindy H Chau, Douglas K Price, William D Figg

Molecular Pharmacology Section, Genitourinary Malignancies Branch, Center for Cancer Research, National Cancer Institute, National Institutes of Health, Bethesda, Maryland, USA

Received February 23, 2021; Accepted March 30, 2021; Epub April 15, 2021; Published April 30, 2021

Abstract: Several genome-wide association studies have been conducted to identify genetic risk factors associated with prostate cancer, but their ability to discover new genetic variants and their applicability across ancestry groups have been limited by their lack of genetic diversity, owing to an underrepresentation of non-European populations. A recent meta-analysis published in *Nature Genetics* by Conti et al. has used a multi-ancestry approach to identify 86 new genetic loci associated with prostate cancer risk, refine leads in known risk regions, and develop a genetic risk score that is transferable across population groups. The findings of this study represent a significant advancement in genetic risk prediction for prostate cancer and their incorporation into standard screening protocols may lead to significant improvements in clinical outcomes.

Keywords: Prostate cancer, genome-wide association study, multi-ancestry, health disparities

Introduction

Over the past two decades, genome-wide association studies (GWAS) have become powerful tools for the analysis of genetic risk factors underlying common diseases and have been used to inform prevention and treatment strategies. However, the applicability of such studies has been plagued by a consistent oversampling of populations of European ancestry. Roughly 78% of individuals included in all GWAS are of European ancestry [1]. Consequently, these strategies may not be transferable to other racial groups, who often exhibit higher incidence and mortality. In prostate cancer, the incidence rate is 1.8 times greater and the mortality rate 2.1 times greater in men of African ancestry than in men of European ancestry [2].

A recent meta-analysis published in *Nature Genetics* by Conti et al. combined the data of 136 GWAS for 107,247 prostate cancer cases and 127,006 controls, including 85,554 cases and 91,972 controls of European ancestry, 10,368 cases and 10,986 controls of African ancestry, 8,611 cases and 18,809 controls of

East Asian ancestry, and 2,714 cases and 5,239 controls from Hispanic populations [3]. It is the most extensive multi-ancestry genetic analysis ever conducted for prostate cancer. Accordingly, the results of the study have allowed for the refinement of signals within genetic regions and broader applicability across ancestry groups. The majority of GWAS for prostate cancer are also focused disproportionately on European populations, limiting their breadth and statistical power for discovering new genetic loci [4]. Across all 136 studies, 5.8-16.8 million single nucleotide polymorphisms (SNPs), insertions, and deletions with $\geq 1\%$ frequency were examined for their association with prostate cancer risk.

Developing the genetic risk score

By combining the summary statistics of single variant tests, the authors identified 86 new independent genetic loci that are associated with prostate cancer risk at genome-wide significance threshold ($P < 5.0 \times 10^{-8}$), bringing the total number of known genetic risk variants to 269 [3]. Further analyses showed that risk allele frequencies of all 269 were similar among

Genetic disparities in prostate cancer risk

Europeans (0.490), Africans (0.494), and Hispanics (0.494) while lower in East Asians (0.479). However, despite similarities in the fraction of each population containing these risk alleles, variants with multi-ancestry odds ratios (ORs) above 1.10 (for 71 variants of 269 or 26.4%) were on average more common in Africans (0.509) than Europeans (0.482), Hispanics (0.483), or East Asians (0.472). Similarly, using a familial risk estimate of 2.5 [5], the 269 risk variants were estimated to explain a higher percentage of familial relative risk in Africans at 43.2% than in Europeans (42.6%), Hispanics (38.5%), or East Asians (33.6%). To capture the cumulative contribution of these genetic risk variants, the authors constructed a genetic risk score (GRS) using the multi-ancestry weights of these variants. While there is no evidence of GRS differentiating risk of aggressive versus non-aggressive disease, 45-51% of all men with aggressive prostate cancer have GRS in the top two decile categories. Therefore, while the GRS does not predict an individual's chances of developing aggressive prostate cancer, it does outline a subset in which a considerable fraction of aggressive disease will develop.

Compared with men at average genetic risk (40-60% GRS category), the estimated OR for men in the top 10% of the GRS was found to be 5.06 (95% CI, 4.84-5.29) for Europeans, 4.47 (95% CI, 3.52-5.68) for East Asians, 4.15 (95% CI, 3.33-5.17) for Hispanics, and 3.74 (95% CI, 3.36-4.17) for Africans. Absolute risk in the top decile was 38% for Europeans (95% CI, 37-39%) and Africans (95% CI, 36-41%), 31% (95% CI, 27-36%) for Hispanics, and 26% (95% CI, 22-30%) for East Asians. The GRS distribution in controls, however, shows an average GRS (standardized around an average GRS of 1.0 in men of European descent) of 2.18 (95% CI, 2.14-2.22) for men of African ancestry, 0.97 (95% CI, 0.94-1.00) for men of Hispanic ancestry, and 0.73 (95% CI, 0.71-0.76) for men of East Asian ancestry, indicating substantial discrepancies in mean genetic risk for men of different ancestry groups, consistent with clinical observations of racial disparities in prostate cancer [3].

This research represents a significant step forward in the pursuit of precision medicine and genetic risk prediction in prostate cancer. While GWAS have led to many important discoveries

of genomic regions associated with common diseases, it is often the case that a SNP found to be associated with the condition is not the causal polymorphism, but a proxy to it as a result of linkage disequilibrium. By incorporating the genetic variability of different populations, one can significantly increase the localization success rate for a causal SNP [6]. In addition to the discovery of 86 new genetic risk variants, this study found stronger markers for 62 of the 183 previously known risk variants and replaced 8 improperly imputed variants with suitable surrogates, strengthening the utility of these risk variants for all population groups.

Diversity on demand

This multi-ancestry approach has led to the development of a genetic risk score that is effective in stratifying prostate cancer across populations of different ancestry and greatly improves upon discriminative models based on age and family history. Based on the more than twofold difference in GRS distribution in controls between men of European and African ancestry, these risk variants are predicted to account for a large fraction of the disparity in prostate cancer incidence. Men at higher risk may benefit from earlier and more frequent screening, and the incorporation of genetic risk prediction capable of transferability across populations may greatly improve early detection of prostate cancer cases, lending to improvements in clinical outcomes. Likewise, the applicability of these findings to other populations, especially men of African ancestry, is likely to improve outcomes for at-risk populations and the reduce clinical health disparities associated with prostate cancer.

Gaps remain in the expansion of genetic analyses to non-European populations and a lack of diversity is a common problem in genetic studies. We accessed the GWAS Diversity Monitor [1], an interactive real-time dashboard that tracks population diversity in GWAS research, to determine the racial composition of GWAS studies conducted in all diseases and specifically in prostate cancer (**Figure 1**). To date, we found that the majority (57%) of GWAS in prostate cancer are exclusively of European ancestry with only 15% of studies covering African, African-American, or African-Caribbean ancestry. These values increase to 78% and less than

Genetic disparities in prostate cancer risk

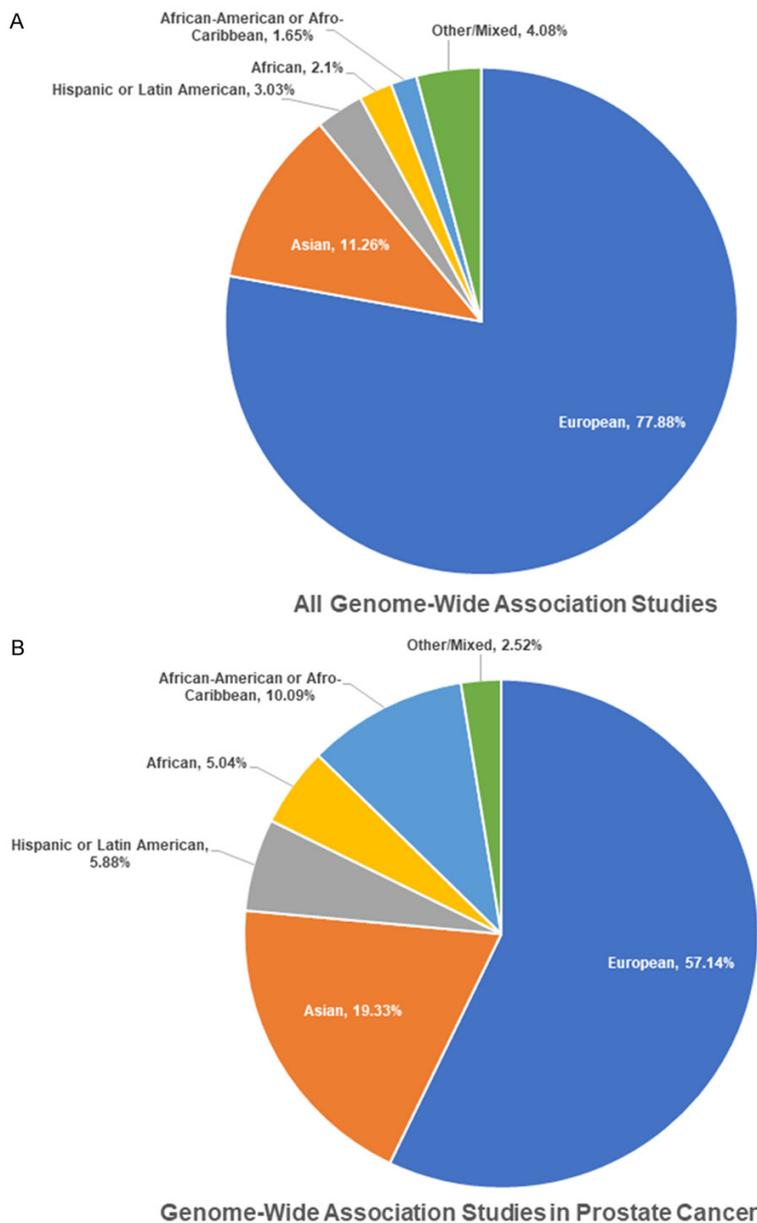


Figure 1. Ancestry distribution of studies in GWAS through February 2021. We show the representation of ancestry groups as percentages included in (A) all GWAS (n=18,211) and (B) prostate cancer GWAS (n=119) through February 2021. The information was obtained using the GWAS Diversity Monitor (<https://gwasdiversitymonitor.com/>) [1].

4% in all GWAS, respectively. There remains a great need for genetic studies with more ancestrally diverse populations and efforts should be made to better include non-European populations. However, it is also important to note that diverse studies such as this may fail to capture specific gene effects within understudied groups. Large meta-analyses identify variants that have similar effects across popu-

lations and often fail to detect population-specific genetic risk factors [7]. It is therefore also necessary to expand the scope of ethnicity-specific genetic analyses in non-European populations. The applicability of the GRS developed by Conti et al. to non-European populations is not yet confirmed and ethnicity-specific studies in men of African descent may provide a valuable cohort for further analysis. For example, the predictive potential of the GRS can be further improved or refined through the RESPOND study (Research on Prostate Cancer in African American Men: Defining the Roles of Genetics, Tumor Markers, and Social Stress), which is expected to be one of the largest genetic studies conducted on prostate cancer in men of African descent.

Furthermore, the predictive utility of genetic risk scores and their implementation in the clinical setting will largely depend on the recently published new standards for reporting and understanding the scores, especially the inclusion of ancestry information of the study population [8]. In addition to standardize reporting of the GRS, incorporation of a multi-ethnic dataset in developing these tools is equally important so that they can perform well in men of all ethnic and racial backgrounds. Conti et al. represents a significant advancement in the pursuit of precision medicine in prostate cancer. In the era of germline testing for prostate cancer, the incorporation of diverse genetic analyses and genetic risk scores to a risk-based screening approach is highly anticipated and may lead to marked improvements in clinical outcomes for at-risk populations through early prevention screening efforts.

Acknowledgements

This work was supported by the Intramural Research Program of the Center for Cancer Research, National Cancer Institute, National Institutes of Health (ZIA BC 010453). The content of this publication does not necessarily reflect the views or policies of the Department of Health and Human Services, nor does mention of trade names, commercial products, or organization imply endorsement by the U.S. Government.

Disclosure of conflict of interest

None.

Address correspondence to: Dr. William D Figg, Molecular Pharmacology Section, Genitourinary Malignancies Branch, Center for Cancer Research, National Cancer Institute, National Institutes of Health, 9000 Rockville Pike, Building 10/Room 5A03, Bethesda 20892, Maryland, USA. Tel: 240-760-6179; Fax: 240-858-3020; E-mail: figgw@mail.nih.gov

References

- [1] Mills MC and Rahal C. The GWAS diversity monitor tracks diversity by disease in real time. *Nat Genet* 2020; 52: 242-243.
- [2] Siegel RL, Miller KD, Fuchs HE and Jemal A. Cancer statistics, 2021. *CA Cancer J Clin* 2021; 71: 7-33.
- [3] Conti DV, Darst BF, Moss LC, Saunders EJ, Sheng X, Chou A, Schumacher FR, Olama AAA, Benlloch S, Dadaev T, Brook MN, Sahimi A, Hoffmann TJ, Takahashi A, Matsuda K, Momozawa Y, Fujita M, Muir K, Lophatananon A, Wan P, Le Marchand L, Wilkens LR, Stevens VL, Gapstur SM, Carter BD, Schleutker J, Tammela TLJ, Sipeky C, Auvinen A, Giles GG, Southey MC, MacInnis RJ, Cybulski C, Wokolorczyk D, Lubirski J, Neal DE, Donovan JL, Hamdy FC, Martin RM, Nordestgaard BG, Nielsen SF, Weischer M, Bojesen SE, Røder MA, Iversen P, Batra J, Chambers S, Moya L, Horvath L, Clements JA, Tilley W, Risbridger GP, Gronberg H, Aly M, Szulkin R, Eklund M, Nordström T, Pashayan N, Dunning AM, Ghousaini M, Travis RC, Key TJ, Riboli E, Park JY, Sellers TA, Lin HY, Albanes D, Weinstein SJ, Mucci LA, Giovannucci E, Lindstrom S, Kraft P, Hunter DJ, Penney KL, Turman C, Tangen CM, Goodman PJ, Thompson IM Jr, Hamilton RJ, Fleshner NE, Finelli A, Parent MÉ, Stanford JL, Ostrander EA, Geybels MS, Koutros S, Freeman LEB, Stampfer M, Wolk A, Håkansson N, Andriole GL, Hoover RN, Machiela MJ, Sørensen KD, Borre M, Blot WJ, Zheng W, Yeboah ED, Mensah JE, Lu YJ, Zhang HW, Feng N, Mao X, Wu Y, Zhao SC, Sun Z, Thibodeau SN, McDonnell SK, Schaid DJ, West CML, Burnet N, Barnett G, Maier C, Schnoeller T, Luedeke M, Kibel AS, Drake BF, Cussenot O, Cancel-Tassin G, Menegaux F, Truong T, Koudou YA, John EM, Grindedal EM, Maehle L, Khaw KT, Ingles SA, Stern MC, Vega A, Gómez-Caamaño A, Fachal L, Rosenstein BS, Kerns SL, Ostrer H, Teixeira MR, Paulo P, Brandão A, Watya S, Lubwama A, Bensen JT, Fontham ETH, Mohler J, Taylor JA, Kogevinas M, Llorca J, Castaño-Vinyals G, Cannon-Albright L, Teerlink CC, Huff CD, Strom SS, Multigner L, Blanchet P, Brureau L, Kaneva R, Slavov C, Mitev V, Leach RJ, Weaver B, Brenner H, Cuk K, Holleczeck B, Saum KU, Klein EA, Hsing AW, Kittles RA, Murphy AB, Logothetis CJ, Kim J, Neuhausen SL, Steele L, Ding YC, Isaacs WB, Nemesure B, Hennis AJM, Carpten J, Pandha H, Michael A, De Ruyck K, De Meerleer G, Ost P, Xu J, Razack A, Lim J, Teo SH, Newcomb LF, Lin DW, Fowke JH, Neslund-Dudas C, Rybicki BA, Gamulin M, Lessele D, Kulis T, Usmani N, Singhal S, Parliament M, Claessens F, Joniau S, Van den Broeck T, Gago-Dominguez M, Castela JE, Martinez ME, Larkin S, Townsend PA, Aukim-Hastie C, Bush WS, Aldrich MC, Crawford DC, Srivastava S, Cullen JC, Petrovics G, Casey G, Roobol MJ, Jenster G, van Schaik RHN, Hu JJ, Sanderson M, Varma R, McKean-Cowdin R, Torres M, Mancuso N, Berndt SI, Van Den Eeden SK, Easton DF, Chanock SJ, Cook MB, Wiklund F, Nakagawa H, Witte JS, Eeles RA, Kote-Jarai Z and Haiman CA. Trans-ancestry genome-wide association meta-analysis of prostate cancer identifies new susceptibility loci and informs genetic risk prediction. *Nat Genet* 2021; 53: 65-75.
- [4] Al Olama AA, Kote-Jarai Z, Berndt SI, Conti DV, Schumacher F, Han Y, Benlloch S, Hazelett DJ, Wang Z, Saunders E, Leongamornlert D, Lindstrom S, Jugurnauth-Little S, Dadaev T, Tymrakiewicz M, Stram DO, Rand K, Wan P, Stram A, Sheng X, Pooler LC, Park K, Xia L, Tyrer J, Kolonel LN, Le Marchand L, Hoover RN, Machiela MJ, Yeager M, Burdette L, Chung CC, Hutchinson A, Yu K, Goh C, Ahmed M, Govindasami K, Guy M, Tammela TL, Auvinen A, Wahlfors T, Schleutker J, Visakorpi T, Leinonen KA, Xu J, Aly M, Donovan J, Travis RC, Key TJ, Siddiq A, Canzian F, Khaw KT, Takahashi A, Kubo M, Pharoah P, Pashayan N, Weischer M, Nordestgaard BG, Nielsen SF, Klarskov P, Røder MA, Iversen P, Thibodeau SN, McDonnell SK, Schaid DJ, Stanford JL, Kolb S, Holt S, Knudsen B, Coll AH, Gapstur SM, Diver WR, Stevens VL, Maier C, Luedeke M, Herkommer

Genetic disparities in prostate cancer risk

- K, Rinckleb AE, Strom SS, Pettaway C, Yeboah ED, Tettey Y, Biritwum RB, Adjei AA, Tay E, True-love A, Niwa S, Chokkalingam AP, Cannon-Albright L, Cybulski C, Wokolorczyk D, Kluźniak W, Park J, Sellers T, Lin HY, Isaacs WB, Partin AW, Brenner H, Dieffenbach AK, Stegmaier C, Chen C, Giovannucci EL, Ma J, Stampfer M, Penney KL, Mucci L, John EM, Ingles SA, Kittles RA, Murphy AB, Pandha H, Michael A, Kierzek AM, Blot W, Signorello LB, Zheng W, Albanes D, Virtamo J, Weinstein S, Nemesure B, Carpten J, Leske C, Wu SY, Hennis A, Kibel AS, Rybicki BA, Neslund-Dudas C, Hsing AW, Chu L, Goodman PJ, Klein EA, Zheng SL, Batra J, Clements J, Spurdle A, Teixeira MR, Paulo P, Maia S, Slavov C, Kaneva R, Mitev V, Witte JS, Casey G, Gillanders EM, Seminara D, Riboli E, Hamdy FC, Coetzee GA, Li Q, Freedman ML, Hunter DJ, Muir K, Gronberg H, Neal DE, Southey M, Giles GG, Severi G; Breast and Prostate Cancer Cohort Consortium (BPC3); PRACTICAL (Prostate Cancer Association Group to Investigate Cancer-Associated Alterations in the Genome) Consortium; COGS (Collaborative Oncological Gene-environment Study) Consortium; GAMEON/ELLIPSE Consortium, Cook MB, Nakagawa H, Wiklund F, Kraft P, Chanock SJ, Henderson BE, Easton DF, Eeles RA and Haiman CA. A meta-analysis of 87,040 individuals identifies 23 new susceptibility loci for prostate cancer. *Nat Genet* 2014; 46: 1103-1109.
- [5] Kicinski M, Vangronsveld J and Nawrot TS. An epidemiological reappraisal of the familial aggregation of prostate cancer: a meta-analysis. *PLoS One* 2011; 6: e27130.
- [6] Zaitlen N, Pasaniuc B, Gur T, Ziv E and Halperin E. Leveraging genetic variability across populations for the identification of causal variants. *Am J Hum Genet* 2010; 86: 23-33.
- [7] Sirugo G, Williams SM and Tishkoff SA. The missing diversity in human genetic studies. *Cell* 2019; 177: 26-31.
- [8] Wand H, Lambert SA, Tamburro C, Iacocca MA, O'Sullivan JW, Sillari C, Kullo IJ, Rowley R, Dron JS, Brockman D, Venner E, McCarthy MI, Antoniou AC, Easton DF, Hegele RA, Khera AV, Chatterjee N, Kooperberg C, Edwards K, Vlessis K, Kinnear K, Danesh JN, Parkinson H, Ramos EM, Roberts MC, Ormond KE, Khoury MJ, Janssens ACJW, Goddard KAB, Kraft P, MacArthur JAL, Inouye M and Wojcik GL. Improving reporting standards for polygenic scores in risk prediction studies. *Nature* 2021; 591: 211-219.