

AD _____

Award Number: W81XWH-09-1-0694

TITLE: Assessing the role of copy number variants in prostate cancer risk and progression using a novel genome-wide screening method

PRINCIPAL INVESTIGATOR: Donna Lehman, Ph.D.

CONTRACTING ORGANIZATION: University of Texas Health Science Center
San Antonio
San Antonio, TX 78229-3901

REPORT DATE: October 2013

TYPE OF REPORT: Annual

PREPARED FOR: U.S. Army Medical Research and Materiel Command
Fort Detrick, Maryland 21702-5012

DISTRIBUTION STATEMENT:

Approved for public release; distribution unlimited

The views, opinions and/or findings contained in this report are those of the author(s) and should not be construed as an official Department of the Army position, policy or decision unless so designated by other documentation.

REPORT DOCUMENTATION PAGE				Form Approved OMB No. 0704-0188	
Public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing this collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden to Department of Defense, Washington Headquarters Services, Directorate for Information Operations and Reports (0704-0188), 1215 Jefferson Davis Highway, Suite 1204, Arlington, VA 22202-4302. Respondents should be aware that notwithstanding any other provision of law, no person shall be subject to any penalty for failing to comply with a collection of information if it does not display a currently valid OMB control number. PLEASE DO NOT RETURN YOUR FORM TO THE ABOVE ADDRESS.					
1. REPORT DATE October 2013		2. REPORT TYPE Annual		3. DATES COVERED 15 September 2012 - 14 September 2013	
4. TITLE AND SUBTITLE Assessing the role of copy number variants in prostate cancer risk and progression using a novel genome-wide screening method.				5a. CONTRACT NUMBER	
				5b. GRANT NUMBER W81XWH-09-1-0694	
				5c. PROGRAM ELEMENT NUMBER	
6. AUTHOR(S) Donna Lehman Robin Leach E-Mail: lehman@uthscsa.edu				5d. PROJECT NUMBER	
				5e. TASK NUMBER	
				5f. WORK UNIT NUMBER	
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) University of Texas Health Science Center San Antonio, TX 78229-3901				8. PERFORMING ORGANIZATION REPORT NUMBER	
9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES) U.S. Army Medical Research and Materiel Command Fort Detrick, Maryland 21702-5012				10. SPONSOR/MONITOR'S ACRONYM(S)	
				11. SPONSOR/MONITOR'S REPORT NUMBER(S)	
12. DISTRIBUTION / AVAILABILITY STATEMENT Approved for Public Release; Distribution Unlimited					
13. SUPPLEMENTARY NOTES					
14. ABSTRACT Individual copy number variations in the genome may play a substantial role in influencing trait variation, yet due to technical limitations they have been understudied. We have performed the first genome-wide association of copy number variants and risk for prostate cancer in Mexican Americans. We found a highly protective deletion on 8q24 which is present in Mexican Americans but extremely rare in Caucasians. Due to the strong effect of this deletion, this discovery has implications for prostate cancer risk assessment and for understanding the etiology of prostate cancer. This variant warrants further study. We have also identified a rare 900 bp deletion in the PTEN gene to be associated with increased risk for prostate cancer and have provided confirmatory data that a rare heritable deletion on 2p24.3 is associated with prostate cancer risk in non-Hispanic Caucasians. These data support our hypothesis that heritable structural variation may affect risk for PCa and/or its progression. Moreover, these variants may be unique to ethnic population and underscores the need to investigate genetic risk in multiple populations. As genes are identified from these studies, they may prove to be both useful biomarkers for early diagnosis and/or novel therapeutic targets for both prevention and treatment of prostate cancer.					
15. SUBJECT TERMS Heritable copy number variation, prostate cancer, Hispanic					
16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT	b. ABSTRACT	c. THIS PAGE			USAMRMC
U	U	U	UU	12	19b. TELEPHONE NUMBER (include area code)

Table of Contents

	<u>Page</u>
Introduction.....	1
Body.....	1-2
Key Research Accomplishments.....	2
Reportable Outcomes.....	2
Conclusion.....	2-3
References.....	3
Appendices.....	4-9

INTRODUCTION

Prostate cancer is known to have a strong genetic component. Thus, the identification of the heritable genetic alteration(s) that precedes or increases susceptibility to somatic cancerous changes in the prostate could likely lead to improved identification of high risk individuals for early screening and possibly to new treatment strategies. Recently, it has become apparent that structural variation comprises similar diversity of human genomes as SNPs and may play a significant role in disease susceptibility and resistance. The goal of this research project is to screen the entire autosomal genome for copy number variation (CNVs) in constitutional DNA to assess their role in risk of development of prostate cancer and then evaluate any direct effect on the prostate. Growing evidence from our study and others indicates that these genetic factors may be rare and numerous and, importantly, unique to the different ethnic populations. The differences in these genetic risk factors may partially explain the ethnic disparities in incidence of prostate cancer. This study is one of few examining genetic risk factors among the Mexican American population.

BODY

We have genotyped prostate cancer cases and elderly hyper-normal controls of Mexican American descent from the SABOR cohort for ~750,000 markers across the genome. We identified 462 copy number variants (CNVs) which were polymorphic in at least 2 individuals. We have also applied the newly released program IMPUTE2.2, the 1000 Genomes project reference panel, and our SNP genotyping data, to impute known CNVs in our samples. We used allelic association in the program PLINK to test for association with cancer status. In total, 3725 CNVs with minor allele frequency (MAF) >0.01 were identified and tested. None of the 760 common CNVs with MAF >0.1 were significantly associated. Our results from direct testing of CNVs are consistent with the growing consensus that there are not common genetic variants of large additive effects on prostate cancer predisposition, regardless of variant type

This year, we have continued focusing on analyzing rare CNVs using multiple methods. We undertook replication efforts of those novel CNVs reported in the literature[1, 2] to be associated with prostate cancer and published our results in the Journal of Urologic Oncology (attached in Appendix). Our findings are consistent with the reported observation that a heritable deletion on 2p24.3 is associated with prostate cancer risk in non-Hispanic Caucasians. Additionally, our observations indicate that the 2p24.3 variant is associated with risk for high grade prostate cancer in a recessive manner. In our own San Antonio subjects, we had observed 3 other rare CNVs which were biologically interesting and nominally significant in the entire Mexican American SABOR cohort: deletions at 2p32.3, 8q24 and 10q23.31. The deletion at 2p32.3 was the most strongly associated CNV in this study. Chromosomal region 8q24 is of interest due to its consistent implication in GWA studies for prostate cancer. A rare non-recurrent 8486 base pair deletion on 8q24 (distinct from the MYC locus) was associated with decreased prostate cancer risk in 989 Mexican American men (Odds ratio 0.20, $p=0.02$). Only 3 of 1530 Caucasians carried the deletion, indicating that this deletion is not likely to affect risk in the Caucasian population. The deleted sequence contains a putative conserved transcription factor binding site for NKX3.1. NKX3.1 is an androgen regulated homeobox gene involved in prostate cancer development, is required for stem cell maintenance, and marks the luminal epithelial cell that is the cell of origin for prostate cancer.[3] The deletion at 10q23.31 is in the gene PTEN, a known prostate cancer tumor suppressor genes. Sequencing data from our lab last year revealed that this deletion is 896 base pairs long with the closest breakpoint residing 57 base pairs from exon 2. We noted that the controls

bearing the PTEN deletion were significantly younger than controls not harboring the variant (53.3 yr vs. 60.2, respectively; $p=0.002$) which leaves the possibility of conversion to PCa early for these men. Therefore, we modeled time to PCa diagnosis as a function of CNV genotype using Cox proportional hazards model and observed significant increase in risk for bearers of this deletion (OR 2.1 95%CI: 1.14-3.87, $p=0.017$). A manuscript describing these results is currently under review.

We conducted extensive burden testing this year and found no strong evidence to support any form of enrichment of CNVs, either rare or common, despite investigating the issue using multiple approaches; including a gene-centric approach, a pathway driven approach, gene set enrichment based approaches, and investigating the characteristics of the CNVs themselves. We have not tested interactions between these variants, SNPs and diet, however.

This year, we also attempted to proceed with Task 3 to examine prostate tumors of bearers of the PTEN deletion (and 2 others) that appear to increase risk for PCa. We unfortunately experienced a very significant delay in locating and retrieving samples due to loss of personnel in the tissue bank and on the SABOR study team. New staff are now hired and trained. In addition, we are now in an excellent position to maximize the extensive CNV data from this project by combining it with that being generated under the auspices of other projects on which Dr. Robin Leach is an investigator. Whole genome expression and methylation profiles are currently being generated on the prostate cancer samples. These data will be ready for use by December 2013. We will analyze all data in the regions of the CNVs that we and others have identified to be associated with prostate cancer in order to investigate their potential function.

KEY RESEARCH ACCOMPLISHMENTS and REPORTABLE OUTCOMES

- Blackburn A, Wilson D, Gelfond J, Yao L, Hernandez J, Thompson IM, Leach RJ, Lehman DM (2013). Validation of copy number variants associated with prostate cancer risk and prognosis. *Urol Oncol*. Sep 17. S1078-1439(13)00253-6. Epub ahead of print. PMID:24054869
- August Blackburn, PhD, graduate of Integrated Multidisciplinary Graduate Program, August 2013, UT Health Science Center San Antonio. A major portion of Dr. Blackburn's dissertation project focused on and was funded by this Idea Development Award. This project provided him with necessary training to lead to his PhD degree and to his authorship of a book chapter in press on the topic of CNVs and complex diseases.
- August Blackburn, PhD, has been selected for a position as Postdoctoral Scientist at the Texas Biomedical Research Institute, San Antonio TX, based upon experience and training supported by this award.

CONCLUSION

We have performed the first genome-wide association of copy number variants and risk for prostate cancer in Mexican Americans. We found a highly protective deletion on 8q24 which is present in Mexican Americans but extremely rare in Caucasians. Due to the strong effect of this deletion, this discovery has implications for prostate cancer risk assessment and for understanding the etiology of prostate cancer. This variant warrants further study. We have also identified a rare 900 bp deletion in the PTEN gene to be associated with increased risk for prostate cancer and have provided confirmatory data that a rare heritable deletion on 2p24.3 is associated with prostate cancer risk in non-Hispanic Caucasians. These data support our hypothesis that heritable structural variation may affect risk for

PCa and/or its progression. Moreover, these variants may be unique to ethnic population and underscores the need to investigate genetic risk in multiple populations. As genes are identified from these studies, they may prove to be both useful biomarkers for early diagnosis and/or novel therapeutic targets for both prevention and treatment of prostate cancer.

REFERENCES

1. Demichelis, F., et al., *Identification of functionally active, low frequency copy number variants at 15q21.3 and 12q21.31 associated with prostate cancer risk*. Proc.Natl.Acad.Sci.U.S.A, 2012. **109**(17): p. 6686-6691.
2. Liu, W., et al., *Association of a germ-line copy number variation at 2p24.3 and risk for aggressive prostate cancer*. Cancer Res., 2009. **69**(6): p. 2176-2179.
3. Kojima, C., Y. Zhang, and W.E. Zimmer, *Intronic DNA elements regulate androgen-dependent expression of the murine Nkx3.1 gene*. Gene Expr., 2010. **15**(2): p. 89-102.

Personnel: Donna Lehman, Robin Leach, Jon Gelfond, August Blackburn



Original article

Validation of copy number variants associated with prostate cancer risk and prognosis ☆, ☆☆

August Blackburn^{1,a}, Desiree Wilson^{1,a}, Jonathan Gelfond, M.D., Ph.D.^{b,e}, Li Yao, Ph.D.^c,
Javier Hernandez, M.D.^{d,e}, Ian M. Thompson, M.D.^{d,e}, Robin J. Leach, Ph.D.^{a,d,e},
Donna M. Lehman, Ph.D.^{c,e,*}

^a Department of Cellular and Structural Biology, University of Texas Health Science Center at San Antonio, San Antonio, TX

^b Department of Epidemiology and Biostatistics, University of Texas Health Science Center at San Antonio, San Antonio, TX

^c Department of Medicine, Division of Clinical Epidemiology, University of Texas Health Science Center at San Antonio, San Antonio, TX

^d Department of Urology, University of Texas Health Science Center at San Antonio, San Antonio, TX

^e Cancer Therapy and Research Center, University of Texas Health Science Center at San Antonio, San Antonio, TX

Received 12 April 2013; received in revised form 3 June 2013; accepted 11 June 2013

Abstract

Objective: Two recent studies have reported novel heritable copy number variants on chromosomes 2p, 15q, and 12q to be associated with prostate cancer (PCa) risk in non-Hispanic Caucasians. The goal of this study was to determine whether these findings could be independently confirmed in the Caucasian population from the South Texas area.

Methods and materials: The study subjects consisted of participants of the San Antonio Biomarkers of Risk for PCa cohort and additional cases ascertained in the same metropolitan area. We genotyped all 7 of the reported copy number variants using real-time quantitative polymerase chain reaction in 1,536 (317 cases and 1,219 controls) non-Hispanic Caucasian men, and additionally, we genotyped 632 (191 cases and 441 controls) Hispanic Caucasian men for one of these variants, a deletion on 2p24.3.

Results: Association of the deletion on 2p24.3 with overall PCa risk did not meet our significance criteria but was consistent with previous reports (odds ratio, 1.40; 95% confidence interval 0.99–2.00; $P = 0.06$). Among Hispanic Caucasians, this deletion is much less prevalent (minor allele frequencies of 0.059 and 0.024 in non-Hispanic and Hispanic Caucasians, respectively) and did not show evidence of association with risk for PCa. Interestingly, among non-Hispanic Caucasians, carrying a homozygous deletion of 2p24.3 was significantly associated with high-grade PCa as defined by Gleason score sum ≥ 8 (odds ratio, 27.99; 95% confidence interval 1.99–392.6; $P = 0.007$ [the Fisher exact test]). The remaining 6 copy number variable regions either were not polymorphic in our cohort of non-Hispanic Caucasians or showed no evidence of association.

Conclusions: Our findings are consistent with the reported observation that a heritable deletion on 2p24.3 is associated with PCa risk in non-Hispanic Caucasians. Additionally, our observations indicate that the 2p24.3 variant is associated with risk for high-grade PCa in a recessive manner. We were unable to replicate any association with PCa for the variants on chromosomes 15q and 12q, which may be explained by regional population differences in low frequency variants and disease heterogeneity. Published by Elsevier Inc.

Keywords: Prostate; Cancer; Risk; Deletion; Prognosis

1. Introduction

Prostate cancer (PCa) is the most common nonskin cancer among American men. The majority of PCa cases, however, is indolent and may not require treatments that are associated with significant rates of voiding and sexual function complications. There is great impetus to identify markers that distinguish indolent from aggressive disease. Given that risk for PCa may be attributed to a heritable component in as many as 42% of cases [1], it may be possible to identify genetic polymorphisms that can be used

☆This work was funded in part by DOD grant PC081025, by grant U01CA086402 from the Early Detection Research Network of the National Cancer Institute, and by the CTRC at UTHSCSA grant P30CA054174. Data were generated in the Genomics Core Shared Resource, which is supported by UTHSCSA, and NIH-NCI P30CA054174 (CTRC of UTHSCSA).

☆☆All other authors can be reached through correspondence with the corresponding author.

* Corresponding author: Tel.: +1-210-567-6714.

E-mail address: lehman@uthscsa.edu (D.M. Lehman).

¹These authors contributed equally to the work described in this manuscript

to better predict risk and prognosis. Multiple studies have attempted to identify genetic variants that are associated with risk for PCa.

The contribution of copy number variants (CNVs) to risk for complex diseases has not yet been fully elucidated, partially because of the difficulty in identification and accurate genotyping [2]. Efforts to characterize common and rare copy number variation in various human populations are still underway. Common CNVs have been shown to be in linkage disequilibrium with adjacent single nucleotide polymorphisms (SNPs) [3–5], which indicates that some common CNVs have already been indirectly assessed for association with traits in SNP-based genome-wide association studies (GWAS). However, recurrent variants and risk-bearing alleles with low minor allele frequencies (MAF) may not be well tagged by SNPs in GWAS, and these variants would likely require direct assessment.

Few studies have investigated risk for PCa attributed to CNVs directly. Liu et al. [6] published the first genome-wide investigation of germ line CNVs and risk for PCa, in which they report a germ line deletion spanning 5,947 base pairs at 2p24.3 associated with risk for PCa and aggressive PCa among men of European descent. Demichelis et al. [7] reported 6 deletions associated with risk for PCa with a false discovery rate <0.2 . Two of these variants, deletions on 15q and 12q, were further supported in this same publication with evidence from bioinformatics analysis and functional assays [7].

GWAS for any genetic variant type are subject to a range of errors and biases, and therefore, replication of association signals in independent samples are necessary to confirm true associations [8]. One of the primary purposes of the San Antonio Center for Biomarkers of Risk for Prostate Cancer (SABOR), a Clinical and Validation Center of the Early Detection Research Network of the National Cancer Institute, is to independently confirm PCa biomarkers, including genetic variants that are predictive of risk. With this motivation, we report the investigation of these reported variants in non-Hispanic and Hispanic Caucasians from San Antonio, Texas.

2. Materials and methods

2.1. Study subjects

Study subjects consisted of 1,372 (153 cases and 1,219 controls) non-Hispanic Caucasian and 516 (75 cases and 441 controls) Hispanic Caucasian participants of SABOR. SABOR is a prospective longitudinal study that examines behavioral, genetic, and other markers of risk of PCa. All study participants are screened annually for PCa using prostate-specific antigen (PSA) serum measurements, with exception to those with PSA <1 , who are screened every other year. In addition, 164 non-Hispanic Caucasian and 116 Hispanic Caucasian men were recruited from a parallel cohort

study of prevalent cases in the same metropolitan area. This study consists of men who were diagnosed with PCa prior to enrollment in the study and were recruited using the same methods as SABOR (local advertisement). Age for this study was calculated in the following manner: for prevalent cases, self-reported age of diagnosis was used; for incident cases, age of diagnosis was used; noncancer (control) participants were censored at their most recent SABOR examination age. Self-reported age of diagnosis for most of the prevalent cases was confirmed through medical records. Institutional review board approval was obtained from the University of Texas Health Science Center at San Antonio. Informed consent was obtained from all participants of both cohorts. We refer to the total population as the SABOR cohort.

Liu et al. [6] defined cancer aggressiveness as meeting any of the following criteria: $T_{3/4}$, N^+ , M^+ , Gleason score sum ≥ 8 , or PSA >50 ng/ml. Although it would be ideal to use identical criterion to define aggressiveness, PSA serum levels at the time of diagnosis as well as staging were not available for all participants of PREF, and thus, these variables were not utilized in defining aggressiveness. Gleason scores were more broadly available, and thus, in this study, aggressive cases in this study were defined by Gleason score of ≥ 8 .

2.2. CNV genotyping

Deoxyribonucleic acid was isolated from whole blood using QIAamp DNA Blood Maxi Kit (Qiagen, Valencia, CA). For real-time quantitative polymerase chain reaction (qPCR), all primers and probes were purchased from Applied Biosystems (Valencia, CA). The primer/probe pairs were either predesigned or designed for the regions of interest using Primer Express (Applied Biosystems, Valencia, CA). qPCR primer/probe information is summarized in Table 1.

For qPCR, primers and probes for a target sequence and reference sequence (Ribonuclease P) were multiplexed in 384-well plates and all samples were run in duplicate. Fluorescence was detected using the 7900HT real-time PCR System (Applied Biosystems, Valencia, CA). Real-time PCR data were analyzed using the reference-free $\Delta\Delta\text{ct}$ method implemented in CopyCaller software (Applied Biosystems, Valencia, CA). This approach estimates the mean Δct for a copy number of 1 and subsequently uses this value to calculate $\Delta\Delta\text{ct}$. We have found this method broadly consistent with the $\Delta\Delta\text{ct}$ method using a reference sample. Discrete copy number calls were determined by plotting histograms of the raw calculated copy number values, which for polymorphic regions reveals nonoverlapping Gaussian distributions representing integer copy number states [9].

To confirm that CNV calls from qPCR accurately detected the 2p24.3 deletion identified by Liu et al. [6], we conducted PCR genotyping as reported by Liu et al. [6] for all homozygous deletions, 12 heterozygous deletions, and 14 wild-type individuals, including 4 individuals on the Gaussian tails where the distributions approach each other in the non-Hispanic Caucasian samples. We observed 100%

Table 1
Summary of quantitative PCR primer/probe pairs

Deletion region	Primer design	Predesigned assay ID or PubMed ID	Probe location (GRCh37/hg19)	Forward primer	Reverse primer	Probe
2p24.3	Custom		Chr2:14707671	TTGGGTCAGGCTGTCTTCT	TTTCAGCAGAGTGGTGAAGGAA	TCTTTCCCAACTCATTACC
15q21.3	Predesigned	Hs05369169_cn ^a	chr15:54199360			
15q21.3	Predesigned	Hs05345872_cn ^a	chr15:54198190			
19q13.12	Predesigned	Hs01617872_cn ^a	chr19:36829874			
8p23.2	Custom ^a		Chr8:3719510			
8q24.3	Predesigned	Hs03668742_cn ^a	chr8:145686582			
7p22.2	Predesigned	Hs01408227_cn ^a	chr7:4091360			
15q21.3	Previously reported	PMID:22496589				
12q21.31	Previously reported	PMID:22496589				

^aAssays designed and ordered from Applied Biosystems and reported according to MIQE Guidelines.

concordance with the real-time PCR results. An example of the genotyping analysis is shown in Fig. 1.

2.3. Statistical analysis

Power analyses for this study were conducted using power for genetic association analysis [10] with the

estimated disease frequency of 0.16, the MAF, estimated relative risk reported for each variant, and $\alpha = 0.05$.

Tests for deviation from Hardy-Weinberg equilibrium were conducted separately for the non-Hispanic Caucasian and Hispanic Caucasian samples. Frequency differences between cases and controls were tested using logistic regression with adjustment for age. As a result of low counts, tests for

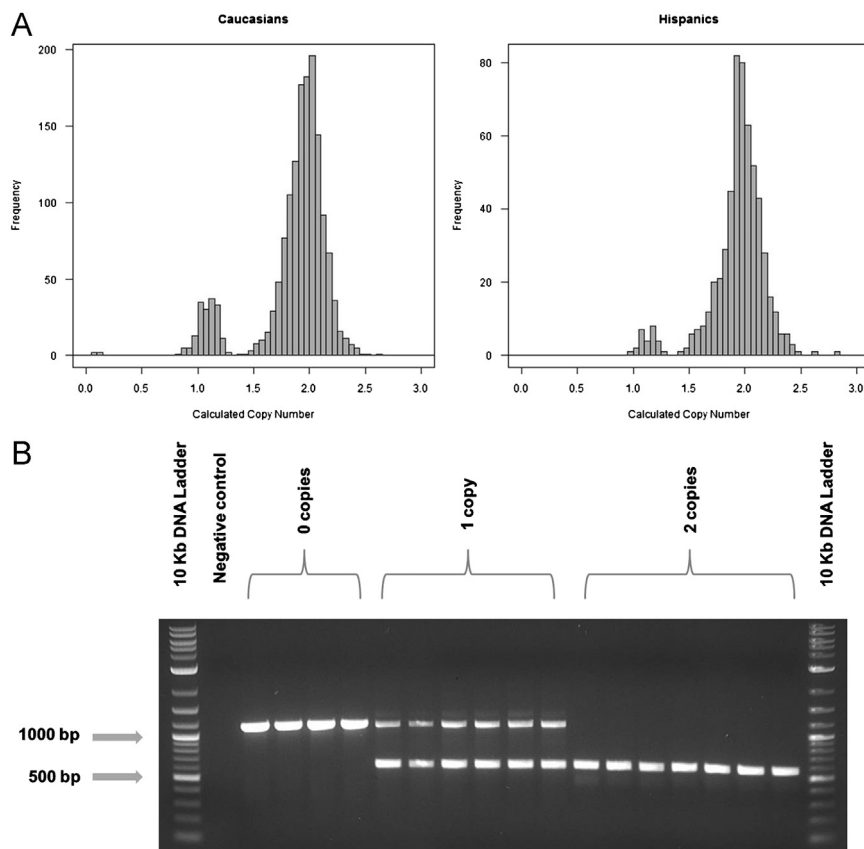


Fig. 1. Genotyping of deletion in San Antonio Cohorts. (A) Histograms of calculated copy number values in Caucasian and Hispanic individuals for a deletion on 2p24.3. Gaussian distributions are present representing measurement variation around integer copy number values of 0, 1, and 2. Discrete values were inferred based on these distributions. (B) Gel electrophoresis of PCR products, as reported by Liu et al., was used to confirm the inferred genotypes in a subset of individuals, including the individuals with calculated copy number values at the edges of these distributions. This gel displays the genotyping results for a subset of the samples and displays all 3 genotypic states of this deletion. Negative control indicates no DNA. DNA = deoxyribonucleic acid.

association with aggressive PCa were conducted using the Fisher exact test for a 2 by 2 contingency table under both dominant and recessive models.

3. Results

We genotyped 7 deletion variants previously reported to be associated with risk for PCa using real-time qPCR in 1,536 (317 cases and 1,219 controls) non-Hispanic Caucasian men, with an average of 1,468 men genotyped per assay after applying strict genotyping criteria. The mean age of diagnosis was 66.7 ± 8.3 , and the mean age of controls was 66.5 ± 9.5 . Using this data set, we estimated power to detect association with 6 variants reported by Demichelis et al. [7] at >0.9 , and ≈ 0.37 for the 2p24.3 variant reported by Liu et al. [6] with an α of 0.05. Four of 6 CNVRs reported to be associated by Demichelis et al. [7] were not observed to be polymorphic. Variants on 12q21.31 and 19q13.12 were polymorphic (MAF ≈ 0.10 , 0.21 respectively), but showed no evidence for association with risk for PCa (12q21.31: odds ratio [OR], 0.89; 95% confidence interval (CI), 0.28–2.77; $P = 0.84$) (19q13.12: OR, 1.08; 95% CI, 0.87–1.34; $P = 0.49$). Genotype counts for these deletions in non-Hispanic Caucasians are presented in Table 2.

The deletion on 2p24.3 reported by Liu et al. [6] had a MAF of 0.059 in non-Hispanic Caucasians and was not associated with risk of PCa given our predetermined α of 0.05 (OR, 1.40; 95% CI, 0.99–2.00; $P = 0.06$). However, these results are consistent with the reported MAF, effect size, and confidence intervals reported by Liu et al. using a dominant model. Among SABOR participants, carrier frequency was not associated with high-grade PCa when compared to controls under a dominant model (OR, 1.28; 95% CI, 0.43–3.12; $P = 0.62$ [the Fisher exact test]), but was significantly associated with high-grade PCa under a recessive model (OR, 27.99; 95% CI, 1.99–392.6; $P = 0.007$). Two (4.44%) of 45 men with high-grade PCa carried homozygous deletions, compared with 2 (0.16%) men carrying homozygous deletions in 1,219 controls. The 2 control carriers for the deletion were aged 43.4 and 55.35 years, which is relatively young for developing PCa.

Additionally, carrying a homozygous deletion is also significantly associated with PCa grade with 2 of 45 high-grade cases carrying a homozygous deletion compared with 0 of 272 low-grade cases (OR, Infinite; 95% CI, 1.15–Infinite, $P = 0.02$). The 2 case carriers had observed Gleason scores of 9 and 10. When considering controls and nonaggressive cases together the enrichment for homozygous deletion in aggressive cases is strong (OR, 34.3; 95% CI, 2.43–477; $P = 0.005$).

We genotyped the 2p24.3 deletion in 632 (191 cases and 441 controls) Hispanic Caucasian men, for whom the average age of diagnosis was 63.0 ± 8.0 and the average age of controls was 60.0 ± 8.3 . Among Hispanic Caucasians, the deletion had a MAF of 0.023 and was not associated with PCa risk (OR, 0.91; 95% CI, 0.36–2.31; $P = 0.84$). Of 37 Hispanic Caucasians with high-grade PCa, we observed zero cases carrying the deletion. This would be expected based on the low MAF of this deletion in the Hispanic Caucasian population.

4. Discussion

In this study, we have independently tested reported associations of germ line CVNs with risk for PCa. Of 6 variants reported by Demichelis et al. [7], 4 were not polymorphic in the non-Hispanic Caucasian population from the SABOR cohort. This observation may be explained by regional population differences in low MAF variants. Alternatively, it is possible that although the primers used in this study fall within the reported boundaries of the variants, they do not fall within the variant's true boundaries. To avoid this scenario, primers and probes were designed to fall within variants identified by the 1,000 genomes project if applicable [11]. Additionally, 3 separate primer probe sets, including those reported by Demichelis et al., were used to genotype the deletion at 15q21.3. Taken together, our data provide strong evidence that the non-Hispanic Caucasian population from the SABOR cohort do not harbor these deletion variants.

Table 2
Genotyping counts in Caucasians

Deletion region	PubMed ID	Cases				Controls			
		DD	DW	WW	MAF	DD	DW	WW	MAF
2p24.3	19258504	2	43	272	0.074	2	129	1,088	0.055
15q21.3	22496589	0	0	310	0	0	0	1,184	0
19q13.12	22496589	14	106	184	0.220	48	397	741	0.208
12q21.31	22496589	4	52	243	0.100	16	185	902	0.098
8p23.2	22496589	0	0	305	0	0	0	1,173	0
8q24.3	22496589	0	0	297	0	0	0	1,120	0
7p22.2	22496589	0	0	305	0	0	0	1,154	0

DD = homozygous deletion; DW = heterozygous deletion; WW = normal copy number.

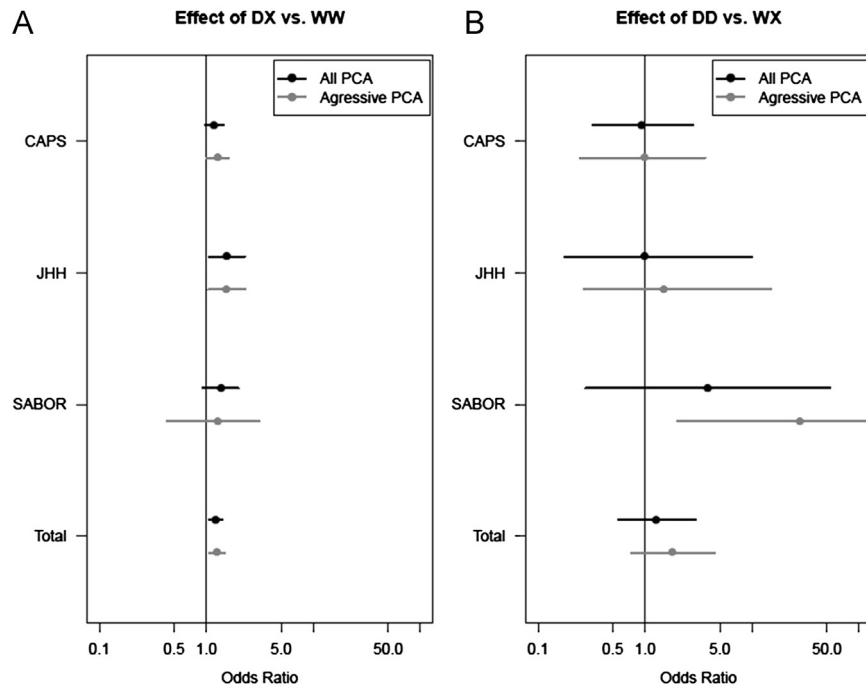


Fig. 2. Comparison between studies of estimated effects for a deletion on 2p24.3. A comparison of odds ratios and 95% confidence intervals for Caucasians in this study (SABOR) and for cohorts previously reported on by Liu et al. Total indicates the results when counts are combined between studies. (A) The results under a dominant model (DX vs. WW) and (B) a recessive model (DD vs. WX). DD = homozygous deletion; DX = hemizygous deletion or homozygous deletion; WW = normal copy number; WX = hemizygous deletion or normal copy number.

For the 2p24.3 variant, our observations are highly consistent with those reported by Liu et al. [6]. Although we obtained a P -value that is not strictly significant, the probability that these results would be observed by chance ($P = 0.06$) is suggestive of an association with overall PCa risk. It is worth noting that the estimated power to detect this association at $\alpha = 0.05$ was not strong (≈ 0.37). A comparison of results under a dominant model (as they reported it) and a recessive model is illustrated in Fig. 2. Interestingly, in Caucasians, we observed statistically significant enrichment of homozygous deletions in high-grade PCa cases. Although all 3 cohorts display overlapping confidence intervals, this observation appears to be somewhat discordant with the observations reported by Liu et al. [6]. However, the criteria that defined aggressiveness varied among the studies, which may explain different associations with aggressive disease. The definition of aggressiveness in this study did not include staging or PSA measurements because data were not available for all prevalent cases. However, the measurement used in this study, Gleason score sum, is a robust predictor of PCa prognosis [12]. Follow-up of this variant in larger cohorts with various definitions of aggressiveness is merited.

5. Conclusions

The initial discovery, and now independent supporting evidence, of the association between the 2p24.3 deletion

reported by Liu et al. and PCa supports the involvement of copy number variation in the etiology of heritable disease. Further studies are merited to confirm the relationship between this deletion and risk for aggressive PCa and to investigate the relationship among this deletion, other linked variants, and nearby genes. The evidence from this study provides additional support that this CVN is a true risk allele for PCa in non-Hispanic Caucasians.

Acknowledgments

The participation and cooperation of all study subjects is gratefully acknowledged. The study could not have been accomplished without the skilled assistance of the SABOR/PREF clinical staff. We thank the expert technical assistance of Yasmin Ench.

References

- [1] Lichtenstein P, Holm NV, Verkasalo PK, Iliadou A, Kaprio J, Koskenvuo M, 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;343:78–85.
- [2] Alkan C, Coe BP, Eichler EE. Genome structural variation discovery and genotyping. *Nat Rev Genet* 2011;12:363–76.
- [3] Conrad DF, Pinto D, Redon R, Feuk L, Gokcumen O, Zhang Y, et al. Origins and functional impact of copy number variation in the human genome. *Nature* 2010;464:704–12.
- [4] Altshuler DM, Gibbs RA, Peltonen L, Dermitzakis E, Schaffner SF, Yu F, et al. Integrating common and rare genetic variation in diverse human populations. *Nature* 2010;467:52–8.

- [5] Craddock N, Hurles ME, Cardin N, Pearson RD, Plagnol V, Robson S, et al. Genome-wide association study of CNVs in 16,000 cases of eight common diseases and 3,000 shared controls. *Nature* 2010;464:713–20.
- [6] Liu W, Sun J, Li G, Zhu Y, Zhang S, Kim ST, et al. Association of a germ-line copy number variation at 2p24.3 and risk for aggressive prostate cancer. *Cancer Res* 2009;69:2176–9.
- [7] Demichelis F, Setlur SR, Banerjee S, Chakravarty D, Chen JY, Chen CX, et al. Identification of functionally active, low frequency copy number variants at 15q21.3 and 12q21.31 associated with prostate cancer risk. *Proc Natl Acad Sci U S A* 2012;109:6686–91.
- [8] McCarthy MI, Abecasis GR, Cardon LR, Goldstein DB, Little J, Ioannidis JP, et al. Genome-wide association studies for complex traits: consensus, uncertainty and challenges. *Nat Rev Genet* 2008;9:356–69.
- [9] Barnes C, Plagnol V, Fitzgerald T, Redon R, Marchini J, Clayton D, et al. A robust statistical method for case-control association testing with copy number variation. *Nat Genet* 2008;40:1245–52.
- [10] Menashe I, Rosenberg PS, Chen BE. PGA: power calculator for case-control genetic association analyses. *BMC Genet* 2008;9:36.
- [11] Consortium TGP. A map of human genome variation from population-scale sequencing. *Nature* 2010;467:1061–73.
- [12] Egevad L. Recent trends in Gleason grading of prostate cancer. II. Prognosis, reproducibility and reporting. *Anal Quant Cytol Histol* 2008;30:254–60.