

RESEARCH ARTICLE

Open Access

# African KhoeSan ancestry linked to high-risk prostate cancer



Desiree C. Petersen<sup>1,2,9†</sup>, Weerachai Jaratlerdsiri<sup>1†</sup>, Abraham van Wyk<sup>3</sup>, Eva K. F. Chan<sup>1,2</sup>, Pedro Fernandez<sup>4</sup>, Ruth J. Lyons<sup>1</sup>, Shingai B. A. Mutambirwa<sup>5</sup>, Andre van der Merwe<sup>4</sup>, Philip A. Venter<sup>6</sup>, William Bates<sup>3</sup>, M. S. Riana Bommman<sup>7</sup> and Vanessa M. Hayes<sup>1,2,6,7,8\*</sup> 

## Abstract

**Backgrounds:** Genetic diversity is greatest within Africa, in particular the KhoeSan click-speaking peoples of southern Africa. South African populations represent admixture fractions including differing degrees of African, African-KhoeSan and non-African genetic ancestries. Within the United States, African ancestry has been linked to prostate cancer presentation and mortality. Together with environmental contributions, genetics is a significant risk factor for high-risk prostate cancer, defined by a pathological Gleason score  $\geq 8$ .

**Methods:** Using genotype array data merged with ancestry informative reference data, we investigate the contribution of African ancestral fractions to high-risk prostate cancer. Our study includes 152 South African men of African (Black) or African-admixed (Coloured) ancestries, in which 40% showed high-risk prostate cancer.

**Results:** Genetic fractions were determined for averaging an equal African to non-African genetic ancestral contribution in the Coloured; we found African ancestry to be linked to high-risk prostate cancer ( $P$ -value = 0.0477). Adjusting for age, the associated African ancestral fraction was driven by a significant KhoeSan over Bantu contribution, defined by Gleason score  $\geq 8$  ( $P$ -value = 0.02329) or prostate specific antigen levels  $\geq 20$  ng/ml ( $P$ -value = 0.03713). Additionally, we observed the mean overall KhoeSan contribution to be increased in Black patients with high-risk (11.8%) over low-risk (10.9%) disease. Linking for the first time KhoeSan ancestry to a common modern disease, namely high-risk prostate cancer, we tested in this small study the validity of using KhoeSan ancestry as a surrogate for identifying potential high-risk prostate cancer risk loci. As such, we identified four loci within chromosomal regions 2p11.2, 3p14, 8q23 and 22q13.2 ( $P$ -value = all age-adjusted  $< 0.01$ ), two of which have previously been associated with high-risk prostate cancer.

**Conclusions:** Our study suggests that ancient KhoeSan ancestry may be linked to common modern diseases, specifically those of late onset and therefore unlikely to have undergone exclusive selective pressure. As such we show within a uniquely admixed South African population a link between KhoeSan ancestry and high-risk prostate cancer, which may explain the 2-fold increase in presentation in Black South Africans compared with African Americans.

**Keywords:** African ancestry, Prostate cancer, KhoeSan, High-risk disease, Ancestral fractions, Ancestry informative markers

\* Correspondence: [v.hayes@garvan.org.au](mailto:v.hayes@garvan.org.au); [vanessa.hayes@sydney.edu.au](mailto:vanessa.hayes@sydney.edu.au)

<sup>†</sup>Desiree C. Petersen and Weerachai Jaratlerdsiri contributed equally to this work.

<sup>1</sup>Laboratory for Human Comparative and Prostate Cancer Genomics, Genomics and Epigenetics Division, Garvan Institute of Medical Research, The Kinghorn Cancer Centre, 370 Victoria Street, Darlinghurst, NSW 2010, Australia

<sup>2</sup>Faculty of Medicine, University of New South Wales, Randwick, NSW, Australia

Full list of author information is available at the end of the article



## Background

High-risk prostate cancer (HRPCa) accounts for approximately 15% of diagnoses in Western countries, with significant potential for associated lethality [1]. Although a number of HRPCa classifications have been proposed, including variations in the requirement for clinical tumor staging and serum prostate specific antigen (PSA) levels, HRPCa is typically defined as pathological Gleason score (GS)  $\geq 8$  or PSA  $\geq 20$  ng/ml at diagnosis. In the United States, African American men are disproportionately affected by HRPCa. Overall, mortality rates are 2-fold higher in American men of African versus European ancestry, while reaching as much as a 4.2-fold increase among younger men [2]. Further support for a bias towards aggressive prostate cancer presentation within African American men includes: elevated PSA levels and younger age at diagnosis, a shorter PSA doubling time prior to surgery, higher tumor grade and volume at surgery, higher rates of biochemical relapse post-surgery and reduced rates of curative therapy [3].

HRPCa is also disproportionately observed in men from sub-Saharan Africa and Southern Africa [4, 5]. Compared with African Americans, the latter study has showed Black South African men are at a 2.1-fold and 4.9-fold greater risk for presenting at diagnosis with GS  $\geq 8$  and PSA  $\geq 20$  ng/ml, respectively. While socio-economic and lifestyle factors, as well as late detection contribute to the disproportionate impact of HRPCa within African Americans, data within Africa is severely lacking. We have previously discussed social, cultural, educational and economic factors contributing to advanced disease presentation within southern Africa, while highlighting the unique opportunity the region holds for leveraging new knowledge with regards to prostate cancer risk and biology, including genetic contribution [6]. The significance of genetic contribution to HRPCa cannot be ignored [3, 7].

In addition to significant HRPCa presentation in Black South Africans [5], HRPCa is also elevated within the African-admixed population from South Africa, the South African Coloured [5, 8]. While Black South Africans represent a uniquely African ancestry, predominantly Bantu, with contributing KhoeSan heritage, the Coloured arose as a result of intermarriage between initial European colonists, Dutch East Indian slaves and indigenous Bantu and KhoeSan Southern Africans [9, 10]. Therefore, the genetic ancestral fractions of the South African Coloured uniquely represent the broad spectrum of prostate cancer racial disparity reported in the United States, specifically African-biased high-risk, European-biased intermediate-risk (GS = 7) and Asian-biased low-risk prostate cancer (LRPCa; GS = 6). In this study we determine if African ancestry, specifically Bantu or KhoeSan African ancestry, is preferentially linked to HRPCa presentation in the region.

## Methods

### Study participants

South African men self-identifying as Black ( $n = 68$ ) or Coloured ( $n = 84$ ) presented at the urology clinics at Polokwane (Limpopo Province), Steve Biko or Dr. George Mukhari (Gauteng Province) or Tygerberg (Western Cape Province) Academic Hospitals. The study was approved and participants consented as required by local ethics approvals, with participant recruitment within Limpopo and Gauteng as part of the previously described Southern African Prostate Cancer Study (SAPCS) [5, 11]. DNA was extracted from whole blood using standard methods (QIAGEN Inc., Germantown, Maryland) and deidentified samples shipped to Australia for further genomic analyses (refer to ethics approvals and permits).

### Clinical and pathological presentation

Presence or absence of prostate cancer was provided by clinicopathological diagnosis. All biopsy cores underwent independent rescoring for the 50 Black cases and 18 Black cancer-free patients as previously described [12] and the 84 Coloured cases (by AvW and WB, see Additional file 1). HRPCa defined as a GS  $\geq 8$ , was confirmed for 33 Black (66%) and 27 Coloured (32%), or PSA  $\geq 20$  ng/ml (irrespective of pathological features), was observed for 36 Black (72%) and 39/81 Coloured (48%). LRPCa defined as a GS = 6, was observed for seven Black (14%) and 12 Coloured (14%), or PSA  $< 10$  ng/ml for six Black (12%) and 23 Coloured (28%). The remaining patients were classified as presenting with intermediate risk disease.

### Genomic data generation

Illumina Infinium HumanCore Beadchip ( $> 250$  K markers) genotype array data was either made available (68 Black) [12] or generated (84 Coloured). Data inclusion was dependant on a GenTrain score (a measure representing the reliability of the genotype calls) of at least 0.5 or more (Illumina GenomeStudio 1.9.4) with further selection of autosomal markers based on a linkage disequilibrium  $r^2$  value  $> 0.2$  within a 50-variant sliding window, advanced by five variants at a time (SNP and Variation Suite 8.3.1, Golden Helix).

### Determining ancestral fractions

Genomic data from population representatives (in brackets) for different African ancestral identifiers were used and defined as: KhoeSan (Ju/'hoansi) [9], West African (Mandinka), Proto-Bantu (Yoruba), West Bantu (Bamoun and Fang), and East Bantu (Luhya) [13], while non-African ancestral identifiers included: Asian (Han Chinese) and European (Utah Americans) (Illumina iControl data). African American data ( $n = 48$ ) was sourced from the International Genome Sample Resource. Ancestral fractions were estimated using STRUCTURE 2.3.3

(5000/10000 burn-in iterations, 10,000/20000 replicates) assuming different ancestral contributions ( $\geq$  five replications) [14].

**Statistical analyses**

Statistical analyses were performed in R (<https://www.r-project.org>) using linear regression (lm) of continuous or categorical data. One-way ANOVA was used for establishing significant disease predictors. Two tailed t-test was used to determine an association between African ancestry and risk extremes, namely HRPCa versus LRPCa. RFMix analysis for local ancestry inference was used to estimate admixture across 22 individual pairs of autosomes [15]. Genotyping data of 84 Coloured patients were removed if unmapped to GRCh37, and phased using SHAPEIT2 with the 1000 Genomes Phase 3 reference panel [16]. RFMix was run with two expectation maximization iterations and 0.2 centimorgan (cM) window size and results of each patient along with the population representatives described above were converted to genomic intervals with ancestral identifiers. The intervals where KhoeSan contributions between HRPCa and LRPCa (defined by either GS or PSA) differed greater than three times were compared using Fisher’s exact significance test and then Bonferroni correction (46 and 45 intervals compared based on GS and PSA values, respectively). Significant phased intervals greater than one megabase were chosen for single marker and haplotype block association

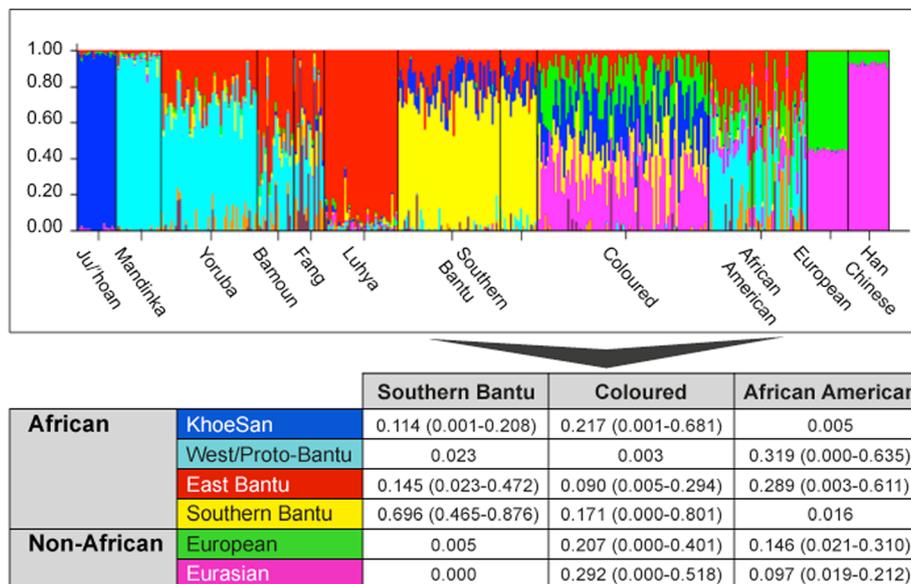
tests using Haploview (<https://www.broadinstitute.org/haploview/haploview>). The RFMix results with posterior probability greater than 0.9 were modelled for migration timing and gene flow estimation using the ancestry tracts analysis (TRACTS) program [17]. The best-fit model assuming KhoeSan, Bantu and Eurasian contributions, was selected based on likelihood values.

**Results**

**Population specific ancestral fractions**

STRUCTURE analysis using 10,295 autosomal markers provided detailed population substructure (Fig. 1 based on eight reference populations). In contrast to African Americans, the African ancestral contributions to the study participants are almost exclusively Bantu and KhoeSan. While African Americans lack KhoeSan contributions, their African ancestral contribution is largely West African (non-Bantu with a lesser West/Proto-Bantu contribution) and East Bantu, with a significant European-biased non-African contribution. The Bantu contribution in our study participants can be defined as uniquely Southern Bantu, 69.6% in the Black and 17.1% in the Coloured, with a smaller East Bantu fraction, 14.5 and 9%, respectively. KhoeSan contributions range from minimal up to 20.8% in the Black and as much as 68.1% in the Coloured.

While the Black participants show exclusive African heritage, the Coloured present overall with an almost equal non-African to African fraction. A 9-fold increase in the number



**Fig. 1** Population substructure of the study participants. (Top Panel) STRUCTURE analysis for 10,295 autosomal markers for eight ancestral populations and including the 68 Black (50 cases and 18 controls) and 84 Coloured South African (SA) study participants compared with African Americans and reference populations from Africa (Ju/'hoansi, Mandinka, Yoruba, Bamoun, Fang and Luhya) and outside of Africa (European and Han Chinese) for a total of 397 subjects. (Middle Panel) Using STRUCTURE analysis we determined the African ancestral fractions, defined as KhoeSan, West/Proto-Bantu, East Bantu and Southern Bantu, as well as the non-African ancestral fractions, defined as European and Eurasian, within our study cohort with comparisons made with the African Americans

of ancestry informative markers through limiting founder population inclusion (91,263 markers) allowed for further separation of the non-African Coloured fractions into European (range 0 to 62.3%) and Asian (range 0.3 to 42.2%) (Fig. 2a). To better understand the extent of African ancestral contributions in our Coloured participants ( $n = 84$ ), we used TRACTS to model their migration history. Consequently, we defined the Coloured as migratory non-African, with significant KhoeSan contributions from 11 (31.5%) to 10 (7.1%) generations ago, followed by Bantu contributions appearing 8 (20.4%) and 7 (11.8%) generations ago (Fig. 2b). In contrast, the KhoeSan contribution to the Black population ( $n = 68$ ) appeared as a single pulse migration event roughly 21 generations ago (11.1%; Optimal likelihoods value:  $-255.7$ ).

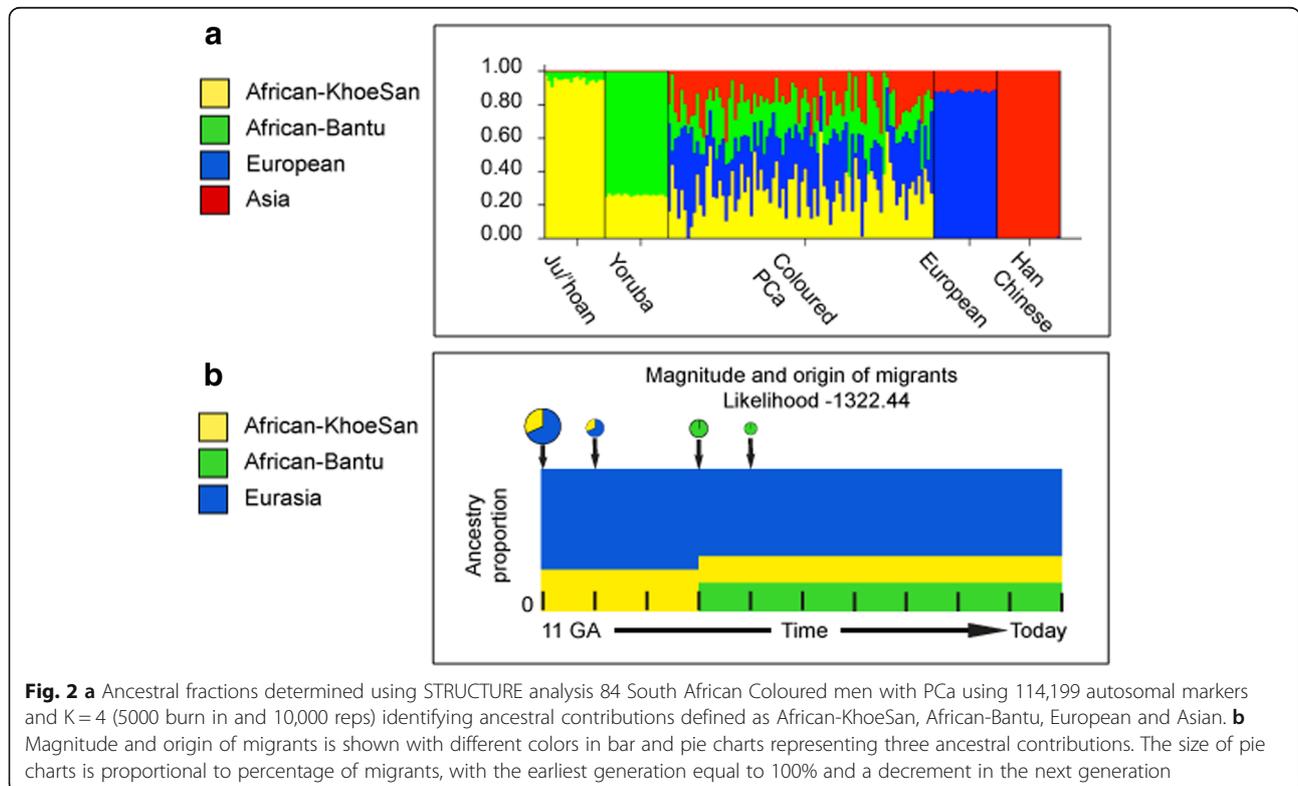
**African ancestral fractions linked to HRPcA**

Presenting with an almost even distribution of African to non-African heritage, the Coloured provide an ideal genetic resource to further evaluate the African ancestral contribution to HRPcA. We observed a significant association between total African ancestry and prostate cancer pathology. Participants with HRPcA ( $GS \geq 8$ ) showed an average of 54.8% African ancestry compared to the 37.3% observed for patients with LRPCa ( $GS = 6$ ) ( $t = 2.0974$ ,  $P$ -value = 0.0477). Furthermore, we observed a significant KhoeSan over Bantu African contribution to HRPcA, specifically the average KhoeSan contributions to  $GS \geq 8$  versus 6 tumors was 31 and

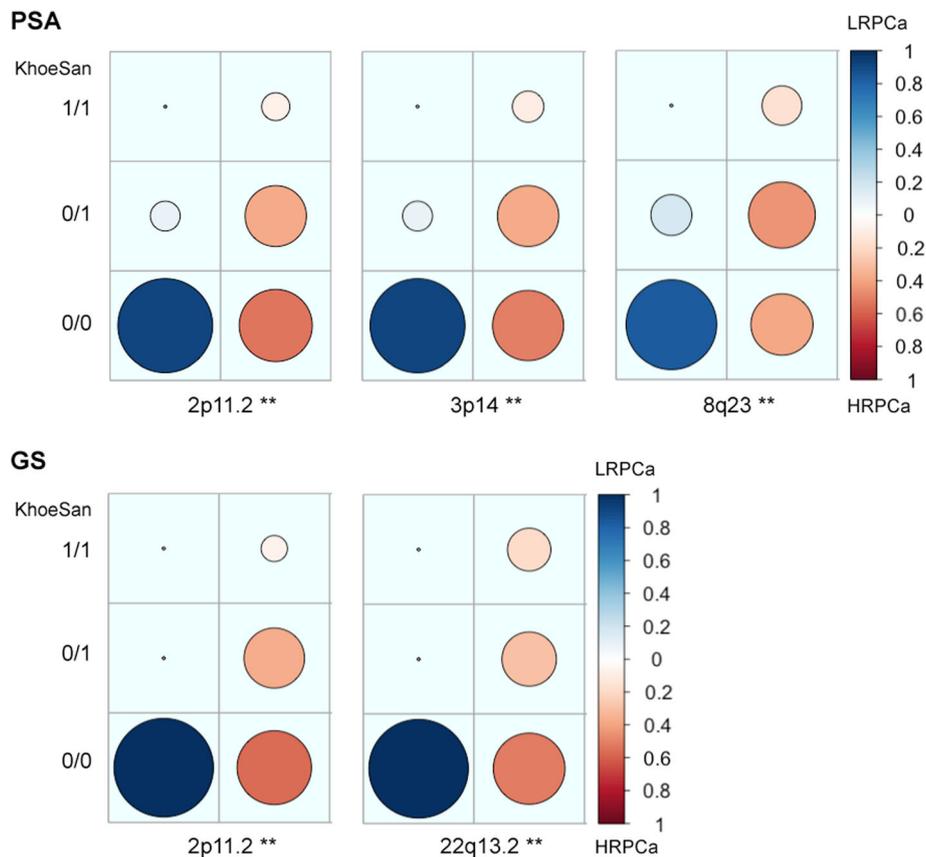
20.1%, respectively ( $t = 2.4491$ ,  $P$ -value = 0.0233) and for  $PSA \geq 20$  versus  $< 10$  ng/ml tumors, 31 and 24.1%, respectively ( $t = 2.1455$ ,  $P$ -value = 0.0371). Although the total KhoeSan contribution to the Black patients was less significant (range 0 to 21%), we did note a slight increase in total KhoeSan ancestral contribution within patients presenting with  $GS \geq 8$  versus 6 tumors (mean 11.8% vs 10.9%;  $t = 0.3249$ ,  $P$ -value = 0.754).

**HRPcA loci enriched for KhoeSan ancestral contribution**

Associating excess KhoeSan contribution within HRPcA presentation in the Coloured, we performed a local-ancestry inference analysis for KhoeSan-specific enrichment, using RFMix [15]. The most significant age-adjusted KhoeSan ancestral association with  $GS \geq 8$  was observed at chromosome 22q13.2 (95 markers; GRCh37 positions 40,178,619-42,552,253; ANOVA  $P$ -value = 0.0062) and chromosome 2p11.2 (332 markers; positions 80,741,406-85,833,046; ANOVA  $P$ -value = 0.0083) (Fig. 3). While KhoeSan ancestry was also associated with an elevated  $PSA \geq 20$  ng/ml at 2p11.2 (ANOVA  $P$ -value = 0.0004), two additional PSA-HRPcA associated loci were identified, including chromosome 3p14 (127 markers; positions 57,971,523-59,436,405; ANOVA  $P$ -value = 0.0026) and 8q23 (79 markers; positions 111,028,667 to 112,656,042; ANOVA  $P$ -value = 0.0052). Performing haplotype and single marker association test we identified two markers, rs10103786 and rs4504665, within 8q23 that remained significant after correcting for multiple



**Fig. 2 a** Ancestral fractions determined using STRUCTURE analysis 84 South African Coloured men with PCa using 114,199 autosomal markers and  $K = 4$  (5000 burn in and 10,000 reps) identifying ancestral contributions defined as African-KhoeSan, African-Bantu, European and Asian. **b** Magnitude and origin of migrants is shown with different colors in bar and pie charts representing three ancestral contributions. The size of pie charts is proportional to percentage of migrants, with the earliest generation equal to 100% and a decrement in the next generation



**Fig. 3** Candidate high-risk prostate cancer (HRPCa) chromosomal regions defined as an over-abundance of KhoeSan heritage. Legends show the proportion of Coloured participants presenting with HRPCa (red) versus low-risk prostate cancer (LRPCa; blue); asterisks (\*\*) indicate regions with age-adjusted *P*-values < 0.01; 1/1, 0/1 or 0/0 represent the presence of KhoeSan ancestry within both DNA strands, a single strand or none, respectively. The local ancestry is defined using RFMix

testing (1000 permutations; Chi-Square = 15.365 and 11.245; *P*-value = 0.007 and 0.048, respectively).

**Discussion**

HRPCa is the major contributor to prostate cancer mortality. The highest mortality rates globally are reported for the Caribbean (29.3 per 100,000), as well as southern, middle, western and eastern Africa (range 24.4 to 18.7 per 100,000) [18]. Asian countries, with overall elevated life expectancy, present with the lowest mortality rates (range 7 to 2.9 per 100,000). Controlling for geography, within the United States, African Americans are at 2.4- and 5-fold greater risk for prostate cancer mortality compared with Americans of European or Asian ancestry, respectively [19]. Elevated mortality rates reported across the Caribbean, United States and Africa, with further implications for familial history of disease as a significant risk factor, raise an important question regarding the contribution of African genetic ancestry to HRPCa.

We determined the contribution of African ancestral contributions defined as Bantu and KhoeSan to increased HRPCa presentation within South Africa. In contrast to African Americans, Black South Africans present with uniquely Bantu, specifically Southern over West Bantu or West non-Bantu contribution, with a single pulse KhoeSan contribution occurring over 550 years ago. The South African Coloured present, on average, with matched non-African to African genetic contributions. Interestingly, the non-African fraction includes both European and Asian contributions, representing intermediate-risk and low-risk populations for HRPCa. Specifically, the African initiating admixture event predates African American admixture by two generations and includes significant KhoeSan contributions followed to a lesser extent by Bantu contribution. We demonstrate that the South African Coloured represents a unique and alternative resource to African American studies for identifying significant African ancestral contributions to elevated HRPCa.

Confirming an African ancestral link to HRPCa within the Coloured, we showed that the observed significance appears to be driven largely by a KhoeSan over Bantu contribution. Although we must caution the need for replication due to a relatively small study size, to the best of our knowledge, this is the first report linking ancient KhoeSan ancestry and prognosis of a common modern condition. It would be reasonable to speculate that prostate cancer risk alleles would not be under negative selection within a hunter-gatherer society with an on average younger overall lifespan. Using KhoeSan ancestry as a surrogate for HRPCa, we have identified four chromosomal regions as potential risk loci for aggressive presentation within the region. The 2p11.2 locus, enriched for both  $GS \geq 8$  and  $PSA \geq 20$  ng/ml, has previously been associated with PCa risk [20, 21]. A recent study, using capture-based Chromosome Conformation Capture (3C) sequencing, identified a significant physical long-range interaction between common variants within the largely non-coding 2p11.2 region and the candidate tumor suppressor gene *CAPG*, with expression quantitative trait locus signals at rs1446669, rs699664 and rs1078004 (absent within our array content) [22]. Additionally, the GS-associated 22q13.2 region has previously been associated with HRPCa in a roughly 1000 strong Swedish genome-wide association study, with independent rs7291691 cross study validation. Located at position 38,778,569, the latter common variant is upstream of the region identified in this study, which may indicate a population specific impact [23]. Notably, the PSA-associated regions, 3p14 and 8q23, are both proximal to known prostate cancer risk loci, including a deletion of the 3p14.1-3p13 region in HRPCa [24] and the common 8q24 prostate cancer risk loci [22].

## Conclusions

In summary, this is the first study to link ancient KhoeSan ancestry to a common modern disease. Specifically, we link KhoeSan ancestry to HRPCa presentation within a uniquely admixed population with African, KhoeSan and Bantu, as well as non-African, European and Asian, ancestries. Using KhoeSan ancestry as a surrogate for HRPCa, we identify potential candidate loci, although one must caution that these regions are only suggestive and require larger study numbers to meet levels of genome-wide significance. However, previously two regions, 2p11 and 22q13 have been suggested as HRPCa risk loci, while two variants at 8q23 remained significant when accounting for multiple testing. Our findings suggest that modern humans earliest ancestors may have been carrying genomic signatures for HRPCa, which would not have been selected against due to later age of onset of prostate cancer. Although largely under-represented in contemporary populations, our study suggests a unique modern application to ancient KhoeSan genetic ancestry.

## Additional files

**Additional file 1: Table S1.** Clinical classification of Coloured patients with prostate cancer. (DOCX 13 kb)

**Additional file 2:** Genotyping data of 68 Black participants (PED and MAP formats). (ZIP 16703 kb)

**Additional file 3:** Genotyping data of 84 Coloured participants (PED and MAP formats). (ZIP 1230 kb)

## Abbreviations

cM: centimorgan; GS: Gleason score; HRPCa: high-risk prostate cancer; LRPCa: low-risk prostate cancer; PSA: prostate specific antigen; SAPCS: Southern African Prostate Cancer Study

## Acknowledgements

The authors acknowledge the study participants, Sister Heather Money and nursing staff at Western Province Blood Transfusion Service (WPBTS), as well as additional urological members of the South African Prostate Cancer Study (SAPCS), Dr. Richard L. Monare and Dr. Smit van Zyl. This research was undertaken with the assistance of resources and services from the National Computational Infrastructure (NCI), which is supported by the Australian Government, and from the Sydney Informatics Hub at the University of Sydney (Artemis facility).

## Authors' contributions

DCP and VMH conceived and designed the study. DCP, PF, SBAM, AvdM, PAV and MSRB enrolled study subjects and maintained clinical databases. MSRB and VMH direct, manage and fund the SAPCS, with SBAM as urological director. VMH sourced funding for genomic analyses. SBAM, AvdM and MSRB provided clinical revision. AvW and WB performed pathological analyses. DCP and RJL isolated the samples, generated genomic data and provided genetic reports. DCP, WJ, EKFC and VMH performed data analysis and critical interpretation. DCP, WJ and EKFC performed statistical analyses. DCP, WJ and VMH drafted the manuscript. All authors reviewed the manuscript. All authors read and approved the final manuscript.

## Funding

The Southern African Prostate Cancer Study (SAPCS) participant recruitment and biobanking was supported by funds from the Cancer Foundation of South Africa (CANSa), the National Research Foundation (NRF) of South Africa, and the Medical Research Council (MRC) of South Africa. Support for genomic analyses was granted to VMH from the Australian Prostate Cancer Research Centre (APCRC) New South Wales (NSW) and from a Perpetual IMPACT grant to the Garvan Foundation, Australia. EFKC and DCP are supported by the Movember Australia and the Prostate Cancer Foundation Australia (PCFA) Prostate Cancer Bone Metastasis (ProMis) Movember Revolutionary Team Award (MRTA), while VMH is supported by the Petre Foundation and University of Sydney Foundation, Australia.

## Availability of data and materials

All data generated or analysed during this study are included in this published article and its supplementary information files (Additional file 2 and Additional file 3).

## Ethics approval and consent to participate

Participants were recruited and performed written consent according to research ethics approvals granted from the Provincial Government of Limpopo (#32/2008) and the University of Limpopo Medical Research Ethics Committee (#MREC/H/28/2009), the University of Pretoria Human Research Ethics Committee (HREC #43/2010, including US Federal wide assurance FWA00002567 and IRB00002235 IORG0001762), Stellenbosch University HREC (#N08/03/072) or the SANBS HREC (#2012/11). DNA was shipped to Australia under the Republic of South Africa Department of Health Export Permits in accordance with the National Health Act 2003 (J1/2/4/2 #1/10, #1/12 and #3/15) and as per institutional Material Transfer Agreements. Genomic interrogation was performed in accordance with St Vincent's Hospital (SVH) HREC site-specific approval (#SVH15/227).

**Consent for publication**

Not applicable.

**Competing interests**

The authors declare that they have no competing interests.

**Author details**

<sup>1</sup>Laboratory for Human Comparative and Prostate Cancer Genomics, Genomics and Epigenetics Division, Garvan Institute of Medical Research, The Kinghorn Cancer Centre, 370 Victoria Street, Darlinghurst, NSW 2010, Australia. <sup>2</sup>Faculty of Medicine, University of New South Wales, Randwick, NSW, Australia. <sup>3</sup>Division of Anatomical Pathology, NHLS Tygerberg and Stellenbosch University, Tygerberg, South Africa. <sup>4</sup>Division of Urology, Department of Surgical Sciences, Faculty of Medicine and Health Sciences, Stellenbosch University, Tygerberg, South Africa. <sup>5</sup>Department of Urology, Sefako Makgatho Health Science University, Dr George Mukhari Academic Hospital, Medunsa, South Africa. <sup>6</sup>Faculty of Health Sciences, University of Limpopo, Mankweng, South Africa. <sup>7</sup>School of Health Systems and Public Health, University of Pretoria, Pretoria, South Africa. <sup>8</sup>Central Clinical School, University of Sydney, Camperdown, NSW, Australia. <sup>9</sup>Centre for Proteomic and Genomic Research (CPGR), 1st Floor, St. Peters Mall, Cnr. Anzio and Main Road, Observatory, Cape Town 7925, South Africa.

Received: 16 December 2018 Accepted: 21 May 2019

Published online: 04 June 2019

**References**

- Chang AJ, Autio KA, Roach M, Scher HI. High-risk prostate cancer-classification and therapy. *Nat Rev Clin Oncol*. 2014;11:308–23.
- Kelly SP, Rosenberg PS, Anderson WF, Andreotti G, Younes N, Cleary SD, et al. Trends in the incidence of fatal prostate Cancer in the United States by race. *Eur Urol*. 2017;71:195–201.
- McGinley KF, Tay KJ, Moul JW. Prostate cancer in men of African origin. *Nat Rev Urol*. 2016;13:99–107.
- Rebbeck TR, Devesa SS, Chang BL, Bunker CH, Cheng I, Cooney K, et al. Global patterns of prostate cancer incidence, aggressiveness, and mortality in men of african descent. *Prostate Cancer*. 2013;2013:560857.
- Tindall EA, Monare LR, Petersen DC, van Zyl S, Hardie RA, Segone AM, et al. Clinical presentation of prostate cancer in black south Africans. *Prostate*. 2014;74(8):880–91.
- Hayes VM, Bornman MSR. Prostate Cancer in southern Africa: does Africa hold untapped potential to add value to the current understanding of a common disease? *J Glob Oncol*. 2017;4:1–7.
- Tan DS, Mok TS, Rebbeck TR. Cancer genomics: diversity and disparity across ethnicity and geography. *J Clin Oncol*. 2016;34:91–101.
- Heyns CF, Fisher M, Lecuona A, van der Merwe A. Prostate cancer among different racial groups in the Western cape: presenting features and management. *S Afr Med J*. 2011;101:267–70.
- Petersen DC, Libiger O, Tindall EA, Hardie RA, Hannick LI, Glashoff RH, et al. Complex patterns of genomic admixture within southern Africa. *PLoS Genet*. 2013;9:e1003309.
- Patterson N, Petersen DC, van-der-Ross RE, Sudoyo H, Glashoff RH, Marzuki S, et al. Genetic structure of a unique admixed population: implications for medical research. *Hum Mol Genet*. 2010;19:411–9.
- Tindall EA, Bornman MS, van-Zyl S, Segone AM, Monare LR, Venter PA, et al. Addressing the contribution of previously described genetic and epidemiological risk factors associated with increased prostate cancer risk and aggressive disease within men from South Africa. *BMC Urol*. 2013;13:74.
- McCrow JP, Petersen DC, Louw M, Chan EK, Harmeyer K, Vecchiarelli S, et al. Spectrum of mitochondrial genomic variation and associated clinical presentation of prostate cancer in south African men. *Prostate*. 2016;76:349–58.
- Henn BM, Gignoux CR, Jobin M, Granka JM, Macpherson JM, Kidd JM, et al. Hunter-gatherer genomic diversity suggests a southern African origin for modern humans. *Proc Natl Acad Sci U S A*. 2011;108:5154–62.
- Pritchard JK, Stephens M, Donnelly P. Inference of population structure using multilocus genotype data. *Genetics*. 2000;155:945–59.
- Maples BK, Gravel S, Kenny EE, Bustamante CD. RFMix: a discriminative modeling approach for local-ancestry inference. *Am J Hum Genet*. 2013;93:278–88.
- Delaneau O, Marchini J. 1000-genomes-project-consortium. Integrating sequence and array data to create an improved 1000 genomes project haplotype reference panel. *Nat Commun*. 2014;5:3934.
- Gravel S. Population genetics models of local ancestry. *Genetics*. 2012;191:607–19.
- Taitt HE. Global trends and prostate Cancer: a review of incidence, detection, and mortality as influenced by race, ethnicity, and geographic location. *Am J Mens Health*. 2018;12:1807–23.
- Siegel R, Ma J, Zou Z, Jemal A. Cancer statistics, 2014. *CA Cancer J Clin*. 2014;64:9–29.
- Akamatsu S, Takata R, Haiman CA, Takahashi A, Inoue T, Kubo M, et al. Common variants at 11q12, 10q26 and 3p11.2 are associated with prostate cancer susceptibility in Japanese. *Nat Genet*. 2012;44:426–9.
- Kote-Jarai Z, Olama AA, Giles GG, Severi G, Schleutker J, Weischer M, et al. Seven prostate cancer susceptibility loci identified by a multi-stage genome-wide association study. *Nat Genet*. 2011;43:785–91.
- Du M, Tillmans L, Gao J, Gao P, Yuan T, Dittmar RL, et al. Chromatin interactions and candidate genes at ten prostate cancer risk loci. *Sci Rep*. 2016;6:23202.
- Sun J, Zheng SL, Wiklund F, Isaacs SD, Li G, Wiley KE, et al. Sequence variants at 22q13 are associated with prostate cancer risk. *Cancer Res*. 2009;69:10–5.
- Feik E, Schweifer N, Baierl A, Sommergruber W, Haslinger C, Hofer P, et al. Integrative analysis of prostate cancer aggressiveness. *Prostate*. 2013;73:1413–26.

**Publisher's Note**

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

**Ready to submit your research? Choose BMC and benefit from:**

- fast, convenient online submission
- thorough peer review by experienced researchers in your field
- rapid publication on acceptance
- support for research data, including large and complex data types
- gold Open Access which fosters wider collaboration and increased citations
- maximum visibility for your research: over 100M website views per year

**At BMC, research is always in progress.**

Learn more [biomedcentral.com/submissions](https://biomedcentral.com/submissions)

