medRxiv preprint doi: https://doi.org/10.1101/2022.04.11.22273677; this version posted April 13, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

Germline sequencing of DNA-damage-repair genes in two hereditary prostate cancer cohorts reveals rare risk-associated variants

Georgea R. Foley BMedRes (Hons) [1], James R. Marthick BSc (Hons) [1], Sionne E. Lucas, PhD [1], Kelsie Raspin, PhD [1], Annette Banks [1], Janet L. Stanford, MPH, PhD [2], Elaine A. Ostrander, PhD [3], Liesel M. FitzGerald, PhD [1*], and Joanne L. Dickinson, PhD [1*^]

[1] Menzies Institute for Medical Research, University of Tasmania, Hobart, TAS, 7000, Australia. Tel +3 61 3 62267700

[2] Fred Hutchinson Cancer Research Center, 1100 Fairview Ave. N., M4-B874, Seattle, WA 98109-1024, USA. Tel. (206) 667-500

[3] Cancer Genetics and Comparative Genomics Branch, National Human Genome Research Institute, NIH, Bethesda, MD 20892, USA. Tel. (301) 594-5284

[*]These authors are joint senior authors

[^]Corresponding Author, Tel: +61 3 6226 7622; Email: jo.dickinson@utas.edu.au

Total word count is 2899, with abstract at 295 words.

Key Words: rare high-risk genes; prostate cancer, DNA repair, genetic susceptibility; germline variants.

ABSTRACT

Background: Rare, inherited variants in DNA damage repair (DDR) genes play an important role in prostate cancer (PrCa) susceptibility. **Objective:** To interrogate two independent high-risk familial PrCa datasets to identify rare DDR variants that contribute to disease risk.

NOTE: This preprint reports new research that has not been certified by peer review and should not be used to guide clinical practice.

medRxiv preprint doi: https://doi.org/10.1101/2022.04.11.22273677; this version posted April 13, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

Design: Massively parallel sequencing data from Australian and North American familial PrCa datasets were examined for rare, likely deleterious variants in 35 DDR genes. Putative high-risk variants were prioritised based on frequency (minor allele frequency <1%), mutation type (nonsense, missense, or splice), segregation with disease, and *in silico* predicted deleteriousness. Six prioritised variants were genotyped in a total of 1,963 individuals (700 familial and 459 sporadic PrCa cases, 482 unaffected relatives and 322 screened controls) and M_{QLS} association analysis performed. **Results and Limitations:** Statistically significant associations between PrCa risk and rare variants in *ERCC3* (rs145201970, p=2.57x10^{-4}) and *BRIP1* (rs4988345, p=0.025) were identified in the combined Australian and North American datasets. A variant in *PARP2* (rs200603922, p=0.028) was significantly associated with risk in the Australian dataset alone, while a variant in *MUTYH* (rs36053993, p=0.031) was significantly associated with risk in the North American dataset. Putative pathogenic variants may have been missed due to their very low frequency in the datasets, which precluded statistical evaluation. **Conclusions:** Our study implicates multiple rare germline DDR variants in PrCa risk, whose functional and/or biological effects and role in inherited risk in other PrCa cohorts should be evaluated. These findings will significantly impact the use of gene-based therapies targeting the DDR pathway in PrCa patients. **Patient Summary**: Here, we looked at genetic changes in a group of genes involved in DNA repair, as testing for such genetic changes is proving important in PrCa diagnosis and treatment. We report, for the first time, several new genetic changes in these genes associated with PrCa.

medRxiv preprint doi: https://doi.org/10.1101/2022.04.11.22273677; this version posted April 13, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

INTRODUCTION

Prostate cancer (PrCa) is responsible for a significant proportion of cancer-related deaths in men worldwide and presents a substantial health burden in terms of morbidity, mental health, and economic costs associated with treatment. A significant proportion of men with advanced disease harbour clinically actionable variants, many of which are aberrations in DNA damage repair (DDR) genes [1-6]. Notably, germline variants in these genes have been observed in 8-16% of metastatic PrCa patients [1,4,7].

Despite recognition of their importance, there remains a significant gap in our understanding of the spectrum of DDR gene variants contributing to PrCa risk [7-11]. Variants in several of these genes, particularly *ATM* and *BRCA1/2,* have been associated with a poorer prognosis, differing responses to treatment, and more aggressive disease [12-16]. Importantly, tumours harbouring loss-of-function mutations in DDR genes exhibit a therapeutic response to poly (ADP-ribose) polymerase inhibition (PARPi) [17] and platinum-based chemotherapy [18]. Thus, screening for clinically actionable germline variants in PrCa patients, particularly those with advanced disease, represents an important strategy to improve disease outcomes. The rarity of these variants in population-based PrCa datasets, which represent the majority of PrCa DDR gene studies to date, has hampered research efforts. In addition, many of these studies have not differentiated between germline and acquired mutations, and those variants that have been identified remain largely of unknown clinical significance.

Curation of the full spectrum of DDR genetic variants contributing to PrCa risk has significant potential in the healthcare setting, where precision medicine can be implemented for both

medRxiv preprint doi: https://doi.org/10.1101/2022.04.11.22273677; this version posted April 13, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

diagnosis and treatment. Moreover, the observation that germline and acquired mutations are frequently identified in the same DDR genes underscores the importance of these pathways in tumour development. Here, we interrogated whole-genome and -exome germline data from two high-risk familial PrCa datasets with the view to identifying novel, rare DDR gene variants contributing to PrCa risk.

medRxiv preprint doi: https://doi.org/10.1101/2022.04.11.22273677; this version posted April 13, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

MATERIALS AND METHODS

Study Resources

The Tasmanian Familial and Case Control Prostate Cancer Studies

The Tasmanian Familial Prostate Cancer Cohort included 73 PrCa families comprising DNA from 379 affected men and 471 unaffected male and female relatives (additional details: Supplementary Methods 1). Additionally, the population-based Tasmanian Prostate Cancer Case-Control Study comprised 459 cases and 322 male controls (additional details: Supplementary Methods 1). Controls were selected at random from the Tasmanian electoral roll, frequency matched by five-year age groups to the cases, and periodically checked against the Tasmanian Cancer Registry for PrCa diagnosis. These datasets are derived from the island state of Tasmania, Australia, which is predominantly of Northern European ancestry. Ethics approval for this study was obtained from the Human Research Ethics Committee Tasmania, Australia (H0017040). Written informed consent was gained for all participants. Available clinical data for these cohorts are presented in Table 1.

The Fred Hutchinson Cancer Research Center Prostate Cancer Genetic Research Study

The Fred Hutchinson Cancer Research Center (FHCRC) resource comprised the Prostate Cancer Genetic Research Study (*PROGRESS*) and included 307 PrCa families from across North America (additional details: Supplementary Methods 1). Existing whole-exome sequencing (WES) data were available from 130 families and included 321 affected men with early-onset and/or aggressive PrCa and 11 older unaffected men. The *PROGRESS* study was approved by the FHCRC's Institutional Review Board and informed consent was

medRxiv preprint doi: https://doi.org/10.1101/2022.04.11.22273677; this version posted April 13, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

obtained from all study participants. Clinical characteristics of this cohort are provided in Table 1.

Whole-Genome Sequencing and Bioinformatic Sequence Analysis

Whole-genome sequencing (WGS) data were generated from germline DNA (additional details: Supplementary Method 2) for 54 individuals from eight Tasmanian families (Supplementary Table S1) and seven unaffected men from the Tasmanian Case-Control Study. Details of sequenced individuals and sequence analysis are available in Supplementary Method 3.

Variant Filtering, Prioritisation, and Validation

A panel was compiled of 35 genes involved in DDR pathways (Table 2), in addition to the established PrCa gene, *HOXB13* [19]. Variants located in a genomic window 1000bp up and downstream of the 35 candidate genes were extracted using bcftools [20] and annotated using ANNOVAR [21]. Included genes and the exact genomic positions can be found in Supplementary Table S2.

Variants were filtered to include those with a minor allele frequency (MAF) <1% in gnomAD non-Finnish Europeans (NFE) and Combined Annotation-Dependent Depletion (CADD) score >15, with further prioritisation informed by predicted mutation function (e.g., nonsense > missense > splicing > synonymous). Variants were excluded if present in >1 of the seven screened unaffected male control genomes, or if present only in PrCa unaffected familial individuals.

medRxiv preprint doi: https://doi.org/10.1101/2022.04.11.22273677; this version posted April 13, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

Short-listed variants (MAF <1%, CADD >15, nonsynonymous, and carried by >1 PrCa case), which had been validated by Sanger sequencing (additional details: Supplementary Method 4) on the ABI 3500 Genetic Analyser (Applied Biosystems), were genotyped in additional non-WGS relatives to determine segregation in the relevant discovery family.

Additional Genotyping in Expanded Tasmanian Resources and Statistical Analysis

Six prioritised variants were genotyped in the full Tasmanian familial and case-control resources, using TaqMan™ genotyping assays (ThermoFisher Scientific) (Supplementary Table S4) on the LightCycler® 480 system (Roche). Association between genotyped variants and PrCa risk was tested for using Modified Quasi-Likelihood Score (M_{QLS}) analysis [22] (additional details: Supplementary Method 5). Population prevalence of PrCa was set at one in seven, and the analyses were conducted in the Australian familial and case-control datasets alone, the FHCRC *PROGRESS* cohort alone, and all datasets combined.

medRxiv preprint doi: https://doi.org/10.1101/2022.04.11.22273677; this version posted April 13, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

RESULTS

Clinical Characteristics of Australian and North American PrCa Resources

Clinical characteristics of the study resources are presented in Table 1. Age at diagnosis, the time interval between diagnosis and death, and the proportion of PrCa-specific deaths were similar across the datasets.

Identification of Candidate Rare DDR PrCa Risk Variants

WGS data were interrogated for rare, potentially pathogenic variants in 35 DDR genes (Table 2). Initial filtering identified 30 variants in 20 genes, of which two have previously been shown to be significantly associated with PrCa risk in our Australian cohort [23,24], providing proof-of-principle for our approach. Of the 28 remaining variants, four failed to validate via Sanger sequencing and were excluded from further investigation.

Additional non-WGS affected and unaffected relatives with DNA from each of the Australian discovery families underwent Sanger sequencing to determine segregation with disease of the remaining 24 variants (Table 3). An additional five variants were subsequently excluded: three variants that were each only present in a single affected man and two variants that were only present in a single affected man and one unaffected relative. The remaining 19 variants, *ATM* rs56128736, *BARD1* rs3738888, *BRCA1* rs28897673, *BRCA2* rs28897727, *BRCA2* rs55639415, *BRCA2* rs786202915, *BRIP1* rs4988345, *ERCC2* rs142568756, *ERCC3* rs145201970, *MRE11* rs777373591, *MSH6* rs142254875, *MUTYH* rs36053993, *PARP2* rs200603922, *PMS2* rs1554304601, *POLE* chr12: 133219216, *POLE* rs41561818, *PTEN*

medRxiv preprint doi: https://doi.org/10.1101/2022.04.11.22273677; this version posted April 13, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

rs587779989, *PTEN* rs773513402, and *RECQL4* rs780723602, were present in at least two affected relatives from the Australian discovery cohort.

For further prioritisation, we then sought to determine whether any of the 19 variants were present in the North American *PROGRESS* families. Examination of exome data from 332 individuals revealed seven variants in 34 cases from 22 kindreds. Four variants, *ATM* rs56128736*, BRCA2* rs28897727, *ERCC3* rs145201970, and *MUTYH* rs36053993, were present in two or more PrCa cases in a single family (Table 3).

Six DDR variants, *BARD1* rs3738888, *BRCA2* rs28897727, *BRIP1* rs4988345, *ERCC3* rs145201970, *MUTYH* rs36053993, and *PARP2* rs200603922, that segregated with disease in a Tasmanian PrCa family and were present in two or more *PROGRESS* families, were selected for additional investigation (Table 3). These variants were genotyped in the extended Tasmanian familial and case-control resources via TaqMan genotyping. All six variants were identified in additional individuals (n_{range}=9 to 33; Supplementary Table S5) within the Australian datasets, and all except *MUTYH* rs36053993 were observed in additional familial PrCa cases. With the inclusion of the *PROGRESS* dataset, the *BARD1* rs3738888 and *BRIP1* rs4988345 variants were each observed in the most PrCa cases (n=22), which included ten and nine sporadic cases, respectively. The predicted pathogenicity of these variants was determined using multiple bioinformatic tools (additional details: Supplementary Method 6) and outputs are shown in Table 4.

Statistical Analysis

In total, genotypes were available for six variants in 1,963 individuals, including 700 familial and 459 sporadic PrCa cases. M_{QLS} association analysis was utilised, which permits the

medRxiv preprint doi: https://doi.org/10.1101/2022.04.11.22273677; this version posted April 13, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

inclusion of related and unrelated individuals while also appropriately controlling for Type 1 error [22]. Analyses were performed on the combined Australian familial and sporadic datasets, the *PROGRESS* familial dataset alone, and the combined Australian and *PROGRESS* datasets (Table 5). In the Australian dataset, a significant association was observed between *PARP2* rs200603922 and PrCa risk (p=0.028), whilst in the *PROGRESS* dataset, a significant association was observed between *BRIP1* rs4988345 (p=0.034), *ERCC3* rs145201970 (p=0.010), and *MUTYH* rs36053993 (p=0.031) and PrCa risk. In the combined Australian and *PROGRESS* datasets, a significant association with PrCa risk was observed between *BRIP1* rs4988345 (p=0.025) and *ERCC3* rs145201970 ($p=2.57 \times 10^{-4}$). The *ERCC3* variant remains significant following Bonferroni correction for multiple testing. PrCa status of variant carriers is provided in Supplementary Table S5.

medRxiv preprint doi: https://doi.org/10.1101/2022.04.11.22273677; this version posted April 13, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

DISCUSSION

The discovery of rare, high-risk germline variants has long proven challenging due to their very low frequency, which substantially impacts power to detect significant statistical associations. However, there remains considerable impetus to characterise rare risk variants in DDR genes, especially considering the increasing availability of therapies targeting this pathway. In an approach designed to take advantage of large familial PrCa resources, we examined massively parallel sequencing data from two independent datasets to identify rare, likely deleterious DDR variants enriched in PrCa. Subsequent analysis of 1,963 individuals from the Australian and *PROGRESS* datasets revealed statistically significant associations between rare variants in *ERCC3* and *BRIP1* and PrCa risk, with *ERCC3* surviving correction for multiple testing. In addition, a variant in *PARP2* was significantly associated with PrCa risk in the Australian dataset alone, while a variant in *MUTYH* was significantly associated with PrCa risk only in the *PROGRESS* dataset.

ERCC3 encodes one of two ATP-dependent DNA helicases, which are core members of the nucleotide excision repair pathway. The *ERCC3* rs145201970 variant (MAF 0.17%), located in exon 7, causes an amino acid change at position 283 (R283C), which is predicted to disrupt the arginine-aspartic acid salt bridge via the inclusion of a more hydrophobic residue. The variant is located within two domains listed by Interpro as likely required for ERCC3 protein function [25]. Topka *et al.* comprehensively examined germline mutations in the *ERCC2*, *3*, *4,* and *5* genes in 16,712 patients affected by multiple different cancers [26]. Numerous likely pathogenic/pathogenic loss of function (LoF) germline variants were observed in *ERCC3*, with rs145201970 (n=42) representing the second most observed predicted LoF variant in this gene in cancer patients after rs34295337 (n=70) [26]. While there are no previous reports

medRxiv preprint doi: https://doi.org/10.1101/2022.04.11.22273677; this version posted April 13, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

describing rs145201970 as a PrCa risk variant, an intronic *ERCC3* variant has been previously associated with increased risk of biochemical recurrence after low-dose-rate prostate brachytherapy, potentially due to reduced mRNA expression in variant carriers [27], while in breast cancer, a recurrent truncating mutation has been associated with familial disease [28,29]. *In vitro* studies have demonstrated mutations in *ERCC3* impair DNA repair capability and confer a selective sensitivity to Irofulven, a sesquiterpene that has demonstrated some efficacy in clinical trials for metastatic PrCa [26].

BRIP1 is a member of the Fanconi Anaemia gene family and functions in the double-strand break repair pathway, interacting closely with *BRCA1*. The rare rs4988345 variant (MAF 0.43%) is in exon 5. As a result of the R173C amino acid change, there is a loss of positive charge and a more hydrophobic residue introduced within a helicase ATP-binding domain and a region annotated as a nuclear localization signal. *BRIP1* rs4988345 has been previously identified in a study enriched for familial PrCa but was only observed in a single PrCa case (0.52%) [30]. The only other previous study to specifically screen *BRIP1* in hereditary PrCa cases detected five other rare missense mutations (MAF <1%) [31], however, no statistical analyses were performed due to their low frequency. A study in breast cancer has linked the rs4988345 variant to disease susceptibility through impairing protein translocation to the nucleus [32]. *BRIP1* has been included on screening panels for several clinical trials investigating the response of metastatic PrCa patients with DDR defects to Olaparib, a PARPi (ClinicalTrials.gov Identifier: NCT02987543) [33]. A cohort of that study comprised men harbouring mutations in 12 DDR genes, including *BRIP1,* however, only four individuals were identified as carriers of a variant in this gene, below the pre-set threshold for statistical analysis. Evaluation of *BRIP1* has also been included in the Phase 2 TRITON2

medRxiv preprint doi: https://doi.org/10.1101/2022.04.11.22273677; this version posted April 13, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

trial (ClinicalTrials.gov Identifier: NCT02952534) where one patient with a *BRIP1* variant responded to the PARPi, Rucaparib [34].

PARP2 is a poly (ADP-ribose) polymerase involved in the base excision repair pathway (BER), and rs200603922 is located in the first exon (R15G). This variant (MAF 0.12%) has previously been observed to partially segregate with PrCa in familial cases who tested negative for *BRCA1* and *BRCA2* mutations [35]. Although several bioinformatic tools predict the variant allele to be benign (Table 4), the R15G amino acid change introduces a more hydrophobic residue, which may impact protein interactions and the phosphorylation of distal residues. There is one other report of a *PARP2* variant, rs3093926 (MAF 4.2%), segregating in a PrCa pedigree, but the contribution of this variant to PrCa risk remains undetermined [36] and though common, it was not observed in our Australian discovery families. *PARP2* mutations have been associated with breast cancer risk [37], but similar to PrCa, their functional impact remains unclear. *PARP2* remains of interest given the ongoing development of PARPi. Though most primarily target *PARP1*, some, such as Niraparib [38], also affect *PARP2,* which may be relevant when assessing therapeutic PARPi in men with *PARP2* mutations.

MUTYH encodes a DNA glycosylase involved in oxidative DDR and the BER pathway. The rs36053993 variant (MAF 0.47%) results in an amino acid change from a neutral residue to a negatively charged, less hydrophobic residue (G368D), with this change located in the highly conserved nudix hydrolase domain. The NCBI human variant database, ClinVar, lists this variant as pathogenic/likely pathogenic arising from its association with MUTYH-associated polyposis, a hereditary condition typified by the development of colorectal carcinomas. While no previous PrCa studies specifically report on rs36053993, one prior study detected a pathogenic/likely pathogenic variant in *MUTYH* in 1.8% of 1,351 PrCa cases [39], while another

medRxiv preprint doi: https://doi.org/10.1101/2022.04.11.22273677; this version posted April 13, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

reported 2.4% of 3,607 PrCa cases as carrying pathogenic variants in this gene [40]. Furthermore, reduced gene and protein expression of *MUTYH* in prostate tumours has been associated with an increase in total somatic mutations, which may result from impaired DDR capacity .

In this study, the strategy for filtering and prioritisation of variants was developed to detect highly penetrant, rare DDR gene variants that may contribute to familial PrCa risk. The innate rarity of these variants presents a significant challenge; therefore, our approach was designed to maximise power by seeking to identify rare variants enriched in two large familial PrCa cohorts. However, it is possible that additional rare, disease-associated variants were not detected due to not being present in the Australian WGS discovery cases. In addition, restricting follow-up to only those candidate variants observed in more than one North American family may have resulted in risk variants associated with disease in the Australian cohort being missed, e.g., *ATM* rs56128736. Furthermore, instances where prioritised variants were subsequently not found to be associated with PrCa could be attributed to their rarity. Thus, examination of additional populations and/or larger familial and sporadic datasets is required.

Conclusion: This study implicates several additional DDR genes as contributors to inherited genetic risk in PrCa. The existing evidence that rare DDR gene variants are associated with aggressive disease and the growing use of cancer therapies targeting this pathway highlights the potential significance of these findings.

medRxiv preprint doi: https://doi.org/10.1101/2022.04.11.22273677; this version posted April 13, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

Data Access and Responsibility

JLD and LMF had full access to all the data in the study and take responsibility for the integrity of the data and the accuracy of the data analysis.

Data Sharing

The data underlying this study are available in the article and its online supplementary material, or from the corresponding author upon reasonable request. The genome and exome sequencing data underlying this article cannot be shared publicly for ethical/privacy reasons.

Funding

This work was supported by grants from the Cancer Council Tasmania and IMPACT Perpetual Trustees (Betty Lowe Memorial Trust, grant number IPAP20210253), as well as the Royal Hobart Hospital Research Foundation (RHHRF); Cancer Australia; The Mazda Foundation; Max Bruce Trust; The Estate of Dr RA Parker; the Tasmanian Community Fund; the Robert Malcom Familial Prostate Cancer Bequest; the Fred Hutchinson Cancer Research Center (grant number P30 CA015704); and the Institute for Prostate Cancer Research of the University of Washington Medicine and Fred Hutchinson Cancer Research Center. Individual support includes a Cancer Council Tasmania/College of Health and Medicine Scholarship to GRF; the National Cancer Institute of the National Institutes of Health (grant numbers R01 CA080122, U01 CA089600, K05 CA175147) to JLS; the National Human Genome Research Institute of the National Institutes of Health to EAO; a previous Cancer Council Tasmania/College of Health and Medicine Senior Research Fellowship and a current

medRxiv preprint doi: https://doi.org/10.1101/2022.04.11.22273677; this version posted April 13, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

Williams Oncology RHHRF grant and Gerald Harvey University of Tasmania Senior Research Fellowship to LMF; and a previous Australian Research Council Future Fellowship and current Select Foundation Cancer Research Fellowship to JLD.

Role of the Funders

The funders of this study had no role in the design; the collection, analysis, and interpretation of the data; the writing of the manuscript; and the decision to submit the manuscript for publication.

Author Disclosures

The authors report no conflict-of-interest disclosures.

Author Contributions

Conceptualization – LMF, JLD; Data curation – GRF, JRM, SEL, KR; Formal Analysis - GRF; Funding acquisition – LMF, JLD; Investigation – GRF, JRM, SEL, LMF, JLD; Methodology – GRF, JRM, SEL, KR, LMF, JLD; Project administration – JLS, LMF, JLD; Resources – JRM, AB, JLS, EAO, LMF, JLD; Supervision – LMF, JLD; Validation – GRF, JRM, JLS, EAO; Writing – original draft – GRF, JRM, LMF, JLD; Writing – review & editing - all authors.

Acknowledgements

We are greatly indebted to the participants of our prostate cancer studies, the Tasmanian Cancer Registry staff, Tasmanian urologists, pathologists, and the wider Tasmanian clinical and research community for their ongoing support. We also thank the individuals participating in the FHCRC *PROGRESS* study.

medRxiv preprint doi: https://doi.org/10.1101/2022.04.11.22273677; this version posted April 13, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

MAIN TABLES & FIGURES

Table 1: Clinical characteristics of study resources

	Australian Familial Cases, n (%)	Australian Sporadic Cases, n (%)	*PROGRESS* Familial Cases, n (%)
Age at Diagnosis			
<60	93 (24.54%)	137 (29.85%)	108 (33.64%)
60-64	98 (25.86%)	170 (37.04%)	75 (23.36%)
65-69	97 (20.32%)	125 (27.23%)	77 (23.99%)
≥70	107 (28.23%)	26 (5.66%)	61 (19.00%)
Missing	4 (1.06%)	1 (0.22%)	n.a.
Age at Diagnosis, Median (IQR)	64.82 (60.06-71.50)	62.59 (59.26-66.05)	63 (57.0-68.0)
Years between Diagnosis and Death			
<5	19 (5.01%)	25 (5.45%)	11 (3.43%)
5-9	39 (10.29%)	54 (11.76%)	36 (11.21%)
10-19	81 (21.37%)	103 (22.44%)	65 (20.25%)
≥20	18 (4.75%)	19 (4.14%)	7 (2.18%)
Missing	2 (0.53%)	n.a.	n.a.
n.a.	220 (58.05%)	258 (56.21%)	202 (62.93%)
Years between Diagnosis and Death, Median (IQR)	11.62 (8.75-16.10)	11.82 (7.64-15.52)	11.00 (6.50-15.0)
Cause of Death			
PrCa	49 (12.93%)	68 (14.81%)	41 (12.77%)
Other	91 (24.01%)	133 (28.98%)	72 (22.43%)
Not Processed	2 (0.53%)	n.a.	n.a.
Missing	17 (4.49%)	1 (0.22%)	6 (1.87%)
n.a.	220 (58.05%)	257 (55.99%)	202 (62.93%)
Total	379	459	321

IQR = Interquartile Range; n.a. = Not Applicable; PrCa = Prostate Cancer.

medRxiv preprint doi: https://doi.org/10.1101/2022.04.11.22273677; this version posted April 13, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

Table 2: Included DNA repair pathway genes

Pathway	Number of Genes	Gene List
Base Excision Repair	2	*MUTYH, PARP2*
Cell Cycle Regulation	5	*CDH1, CDKN1B, CDKN2Z, PTEN, STK11*
DNA Damage Response	3	*ATM, CHEK2, TP53*
Fanconi Anaemia	5	*BRCA2, BRIP1, PALB2, RAD51C, SLX4*
Homologous Recombination	9	*BARD1, BLM, BRCA1, GEN1, MRE11A, NBN, RAD50, RAD51D, RECQL4*
Mismatch Repair	7	*MLH1, MSH2, MSH5, MSH6, PMS2, POLD1, POLE*
Nucleotide Excision Repair	3	*ERCC2, ERCC3, XPC*
Included as proof-of-principle	1	*HOXB13*
Total	36	

Table 3: Putative pathogenic mutations identified in DDR genes in Australian discovery and North American families

Gene	Variant	Allele Change	MAF (gnomAD NFE)	CADD	Australian Discovery Familial Cohort			North American Familial Cohort		
					PrCa Affected Carriers	Total Carriers	Number of Families	PrCa Affected Carriers	Total Carriers	Number of Families
ATM	rs55801750	T>C	9E-04	22.7	1	2	1	-	-	-
ATM	rs55982963	G>A	1E-04	29.3	1	2	1	-	-	-
ATM	rs56128736	T>C	0.0021	23.1	3	6	2	2	2	1
ATM	rs767507047	A>G	6.48E-05	28.6	1	1	1	-	-	-
BRCA1	rs28897673	T>C	1E-04	23.6	2	4	1	-	-	-
BRCA2	rs55639415	C>T	4.71E-05	15.43	2	4	1	-	-	-
BRCA2	rs56403624	A>G	4.29E-04	19.02	1	1	1	-	-	-
BRCA2	rs786202915	T>A	n.a.	16.65	3	8	1	-	-	-
ERCC2	rs142568756	C>T	0.0005	29.1	2	4	1	-	-	-
MRE11	rs777373591	G>A	1.77E-05	27.2	2	4	1	-	-	-
MSH6	rs142254875	C>T	0.0001	22.2	2	6	1	-	-	-
PMS2	rs1554304601	G>A	n.a.	31	3	7	1	-	-	-
POLE	chr12: 133219216	G>C	n.a.	22.8	3	9	1	-	-	-
POLE	rs36120395	G>C	0.0016	20.3	1	1	1	-	-	-
POLE	rs41561818	G>A	0.0044	21.2	2	3	1	-	-	-
PTEN	rs587779989	G>A	n.a.	19.81	3	6	1	-	-	-
PTEN	rs773513402	C>A	0.0003	15.18	2	5	1	-	-	-
RECQL4	rs780723602	T>C	9E-06	22.3	4	11	1	-	-	-
BARD1	rs3738888	G>A	0.0063	26.4	2	5	1	4	4	4
BRCA2	rs28897727	G>T	0.0098	15.85	3	9	1	4	5	3
BRIP1	rs4988345	G>A	0.0043	25.5	2	5	1	6	6	6
ERCC3	rs145201970	G>A	0.0017	26.7	2	7	1	5	5	3
MUTYH	rs36053993	C>T	0.0047	32	2	5	2	9	9	4
PARP2	rs200603922	A>G	0.0012	15.32	4	6	1	2	2	2

n.a. = Not Applicable

Table 4: Predicted pathogenicity of prioritised DDR gene variants

Gene	Variant	Allele Change	AA Change	CADD	DANN*	SIFT	PROVEAN	PolyPhen	Mutation Taster (Rank Score)	Mutation Assessor	FATHMM† (Coding)
BARD1	rs3738888	G>A	R658C	26.4	0.999	0.008 (D)	-4.02 (De)	1 (P)	0.462 (D)	2.12 (M)	0.9778
BRCA2	rs28897727	G>T	D1420Y	15.85	0.988	0.030 (D)	-6.60 (De)	0.03 (B)	0.09 (N)	2.15 (M)	0.49798
BRIP1	rs4988345	G>A	R173C	25.5	0.999	0.001 (D)	-2.54 (De)	1 (P)	0.81 (D)	2.67 (M)	0.93639
ERCC3	rs145201970	G>A	R283C	26.7	0.999	0.000 (D)	-7.59 (De)	0.995 (P)	0.81 (D)	3.31 (M)	0.99364
MUTYH	rs36053993	C>T	G368D	32	0.998	0.000 (D)	-6.46 (De)	1 (P)	0.81 (D)	4.09 (H)	0.99757
PARP2	rs200603922	A>G	R15G	15.32	0.8	0.153 (T)	-1.04 (N)	0 (B)	0.09 (N)	0.695 (N)	0.00048

D = Damaging; T = Tolerated; De = Deleterious; N = Neutral; P = Probably Damaging; B = Benign; M = Medium; H = High. *DANN predictions use a scoring system between 0 and 1, with scores closer to one indicating greater predicted pathogenicity. †FATHMM predictions use a scoring system between 0 and 1, with scores closer to one indicating greater predicted pathogenicity.

Table 5: Carrier frequency and statistical analysis of variants

Gene	Variant	Australian Familial and Sporadic Prostate Cancer			North American Familial *PROGRESS* Cohort			Tasmanian Familial Prostate Cancer Study & *PROGRESS*		
		Total Carriers (% Cases)*	M_{QLS} P-Value	M_{QLS} Odds Ratio	Total Carriers (% Cases)*	M_{QLS} P-Value	M_{QLS} Odds Ratio	Total Carriers (% Cases)*	M_{QLS} P-Value	M_{QLS} Odds Ratio
BARD1	rs3738888	31 (58.1%)	0.407	1.7	4 (100%)	0.318	n.a.	35 (62.9%)	0.066	1.9
BRCA2	rs28897727	24 (54.2%)	0.063	n.a.	5 (80%)	0.157	n.a.	29 (58.6%)	0.193	n.a.
BRIP1	rs4988345	25 (64.0%)	0.118	3.1	6 (100%)	0.034	n.a.	31 (71.0%)	0.025	3.1
ERCC3	rs145201970	16 (50.0%)	0.554	1	5 (100%)	0.010	n.a.	21 (61.9%)	2.57×10^{-4}	1.7
MUTYH	rs36053993	23 (26.1%)	0.630	0.4	9 (100%)	0.031	n.a.	32 (46.9%)	0.201	0.8
PARP2	rs200603922	14 (71.4%)	0.028	n.a.	2 (100%)	0.388	n.a.	16 (75.0%)	0.162	n.a.

n.a. = Not Applicable, as the odds ratio cannot be calculated when no carriers in controls are identified. *% of variant carriers that are cases.

medRxiv preprint doi: https://doi.org/10.1101/2022.04.11.22273677; this version posted April 13, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

REFERENCES

1. Castro E, Romero-Laorden N, del Pozo A, et al. PROREPAIR-B: A Prospective Cohort Study of the Impact of Germline DNA Repair Mutations on the Outcomes of Patients With Metastatic Castration-Resistant Prostate Cancer. *Journal of Clinical Oncology.* 2019/02/20 2019;37(6):490-503. doi: 10.1200/JCO.18.00358.
2. Abida W, Armenia J, Gopalan A, et al. Prospective Genomic Profiling of Prostate Cancer Across Disease States Reveals Germline and Somatic Alterations That May Affect Clinical Decision Making. *JCO Precis Oncol.* Jul 2017;2017. doi: 10.1200/po.17.00029.
3. Abida W, Cyrta J, Heller G, et al. Genomic correlates of clinical outcome in advanced prostate cancer. *Proceedings of the National Academy of Sciences of the United States of America.* Jun 4 2019;116(23):11428-11436. doi: 10.1073/pnas.1902651116.
4. Robinson D, Van Allen Eliezer M, Wu Y-M, et al. Integrative Clinical Genomics of Advanced Prostate Cancer. *Cell.* 2015/05/21/ 2015;161(5):1215-1228. doi: https://doi.org/10.1016/j.cell.2015.05.001.
5. Holt C. Comprehensive Genetic Profiling Unravels Molecular Targets in Prostate Cancer. *Oncology Times.* 2019;41(14):9. doi: 10.1097/01.COT.0000577096.71826.2f.
6. Beltran H, Yelensky R, Frampton GM, et al. Targeted next-generation sequencing of advanced prostate cancer identifies potential therapeutic targets and disease heterogeneity. *Eur Urol.* May 2013;63(5):920-926. doi: 10.1016/j.eururo.2012.08.053.
7. Pritchard CC, Mateo J, Walsh MF, et al. Inherited DNA-Repair Gene Mutations in Men with Metastatic Prostate Cancer. *The New England journal of medicine.* 2016;375(5):443-453. doi: 10.1056/NEJMoa1603144.
8. Lozano R, Castro E, Aragón IM, et al. Genetic aberrations in DNA repair pathways: a cornerstone of precision oncology in prostate cancer. *British Journal of Cancer.* Feb 2021;124(3):552-563. doi: 10.1038/s41416-020-01114-x.
9. Paulo P, Maia S, Pinto C, et al. Targeted next generation sequencing identifies functionally deleterious germline mutations in novel genes in early-onset/familial prostate cancer. *PLOS Genetics.* 2018;14(4):e1007355. doi: 10.1371/journal.pgen.1007355.
10. Leongamornlert DA, Saunders EJ, Wakerell S, et al. Germline DNA Repair Gene Mutations in Young-onset Prostate Cancer Cases in the UK: Evidence for a More Extensive Genetic Panel. *Eur Urol.* Sep 2019;76(3):329-337. doi: 10.1016/j.eururo.2019.01.050.
11. Darst BF, Dadaev T, Saunders E, et al. Germline Sequencing DNA Repair Genes in 5545 Men With Aggressive and Nonaggressive Prostate Cancer. *JNCI: Journal of the National Cancer Institute.* 2021;113(5):616-625. doi: 10.1093/jnci/djaa132.
12. Nombela P, Lozano R, Aytes A, Mateo J, Olmos D, Castro E. BRCA2 and Other DDR Genes in Prostate Cancer. *Cancers.* Mar 12 2019;11(3). doi: 10.3390/cancers11030352.
13. Castro E, Goh C, Leongamornlert D, et al. Effect of BRCA Mutations on Metastatic Relapse and Cause-specific Survival After Radical Treatment for Localised Prostate Cancer. *European Urology.* 2015/08/01/ 2015;68(2):186-193. doi: https://doi.org/10.1016/j.eururo.2014.10.022.
14. Na R, Zheng SL, Han M, et al. Germline Mutations in ATM and BRCA1/2 Distinguish Risk for Lethal and Indolent Prostate Cancer and are Associated with Early Age at Death. *European Urology.* 2017/05/01/ 2017;71(5):740-747. doi: https://doi.org/10.1016/j.eururo.2016.11.033.
15. Mitra A, Fisher C, Foster CS, et al. Prostate cancer in male BRCA1 and BRCA2 mutation carriers has a more aggressive phenotype. *British Journal of Cancer.* Jan 29 2008;98(2):502-507. doi: 10.1038/sj.bjc.6604132.
16. Castro E, Goh C, Olmos D, et al. Germline BRCA mutations are associated with higher risk of nodal involvement, distant metastasis, and poor survival outcomes in prostate cancer.

medRxiv preprint doi: https://doi.org/10.1101/2022.04.11.22273677; this version posted April 13, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

Journal of clinical oncology : official journal of the American Society of Clinical Oncology. May 10 2013;31(14):1748-1757. doi: 10.1200/jco.2012.43.1882.

17. Mateo J, Carreira S, Sandhu S, et al. DNA-Repair Defects and Olaparib in Metastatic Prostate Cancer. *New England Journal of Medicine.* 2015/10/29 2015;373(18):1697-1708. doi: 10.1056/NEJMoa1506859.
18. Mota JM, Barnett E, Nauseef JT, et al. Platinum-Based Chemotherapy in Metastatic Prostate Cancer With DNA Repair Gene Alterations. *JCO Precis Oncol.* 2020;4:355-366. doi: 10.1200/po.19.00346.
19. Ewing CM, Ray AM, Lange EM, et al. Germline Mutations in HOXB13 and Prostate-Cancer Risk. *The New England journal of medicine.* Jan 12 2012;366(2):141-149. doi: 10.1056/NEJMoa1110000.
20. Danecek P, Bonfield JK, Liddle J, et al. Twelve years of SAMtools and BCFtools. *Gigascience.* 2021;10(2):giab008. doi: 10.1093/gigascience/giab008.
21. Wang K, Li M, Hakonarson H. ANNOVAR: functional annotation of genetic variants from high-throughput sequencing data. *Nucleic Acids Research.* 2010;38(16):e164-e164. doi: 10.1093/nar/gkq603.
22. Thornton T, McPeek Mary S. Case-Control Association Testing with Related Individuals: A More Powerful Quasi-Likelihood Score Test. *American Journal of Human Genetics.* 2007;81(2):321-337. doi: 10.1086/519497.
23. Marthick JR, Raspin K, Foley GR, et al. Massively parallel sequencing in hereditary prostate cancer families reveals a rare risk variant in the DNA repair gene, RAD51C. *European Journal of Cancer.* 2021/12/01/ 2021;159:52-55. doi: https://doi.org/10.1016/j.ejca.2021.09.038.
24. FitzGerald LM, Raspin K, Marthick JR, et al. Impact of the G84E variant on HOXB13 gene and protein expression in formalin-fixed, paraffin-embedded prostate tumours. *Scientific Reports.* 2017;7. doi: 10.1038/s41598-017-18217-w.
25. Blum M, Chang H-Y, Chuguransky S, et al. The InterPro protein families and domains database: 20 years on. *Nucleic Acids Research.* 2021;49(D1):D344-D354. doi: 10.1093/nar/gkaa977.
26. Topka S, Steinsnyder Z, Ravichandran V, et al. Targeting Germline- and Tumor-Associated Nucleotide Excision Repair Defects in Cancer. *Clinical cancer research : an official journal of the American Association for Cancer Research.* Apr 1 2021;27(7):1997-2010. doi: 10.1158/1078-0432.Ccr-20-3322.
27. Carignan D, Lessard T, Villeneuve L, et al. DNA repair gene polymorphisms, tumor control, and treatment toxicity in prostate cancer patients treated with permanent implant prostate brachytherapy. *The Prostate.* 2020/05/01 2020;80(8):632-639. doi: https://doi.org/10.1002/pros.23975.
28. Feliubadaló L, Tonda R, Gausachs M, et al. Benchmarking of Whole Exome Sequencing and Ad Hoc Designed Panels for Genetic Testing of Hereditary Cancer. *Scientific Reports.* 2017;7:37984-37984. doi: 10.1038/srep37984.
29. Vijai J, Topka S, Villano D, et al. A Recurrent ERCC3 Truncating Mutation Confers Moderate Risk for Breast Cancer. *Cancer Discov.* Nov 2016;6(11):1267-1275. doi: 10.1158/2159-8290.Cd-16-0487.
30. Kote-Jarai Z, Jugurnauth S, Mulholland S, et al. A recurrent truncating germline mutation in the BRIP1/FANCJ gene and susceptibility to prostate cancer. *British Journal of Cancer.* 2009/01/01 2009;100(2):426-430. doi: 10.1038/sj.bjc.6604847.
31. Ray AM, Zuhlke KA, Johnson GR, et al. Absence of truncating BRIP1 mutations in chromosome 17q-linked hereditary prostate cancer families. *British Journal of Cancer.* 2009/12/01 2009;101(12):2043-2047. doi: 10.1038/sj.bjc.6605433.
32. Lei H, Vorechovsky I. BACH1 517C→T transition impairs protein translocation to nucleus: A role in breast cancer susceptibility? *International Journal of Cancer.* 2003/04/10 2003;104(3):389-391. doi: https://doi.org/10.1002/ijc.10947.

medRxiv preprint doi: https://doi.org/10.1101/2022.04.11.22273677; this version posted April 13, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

33. de Bono J, Mateo J, Fizazi K, et al. Olaparib for Metastatic Castration-Resistant Prostate Cancer. *New England Journal of Medicine.* 2020/05/28 2020;382(22):2091-2102. doi: 10.1056/NEJMoa1911440.
34. Abida W, Campbell D, Patnaik A, et al. Non-BRCA DNA Damage Repair Gene Alterations and Response to the PARP Inhibitor Rucaparib in Metastatic Castration-Resistant Prostate Cancer: Analysis From the Phase II TRITON2 Study. *Clinical cancer research : an official journal of the American Association for Cancer Research.* 2020;26(11):2487-2496. doi: 10.1158/1078-0432.CCR-20-0394.
35. Hunter SM, Rowley SM, Clouston D, et al. Searching for candidate genes in familial BRCAX mutation carriers with prostate cancer. *Urologic Oncology: Seminars and Original Investigations.* 2016/03/01/ 2016;34(3):120.e129-120.e116. doi: https://doi.org/10.1016/j.urolonc.2015.10.009.
36. Johnson AM, Zuhlke KA, Plotts C, et al. Mutational landscape of candidate genes in familial prostate cancer. *The Prostate.* 2014/10/01 2014;74(14):1371-1378. doi: https://doi.org/10.1002/pros.22849.
37. Popanda O, Seibold P, Nikolov I, et al. Germline variants of base excision repair genes and breast cancer: A polymorphism in DNA polymerase gamma modifies gene expression and breast cancer risk. *International Journal of Cancer.* 2013/01/01 2013;132(1):55-62. doi: https://doi.org/10.1002/ijc.27665.
38. Thorsell A-G, Ekblad T, Karlberg T, et al. Structural Basis for Potency and Promiscuity in Poly(ADP-ribose) Polymerase (PARP) and Tankyrase Inhibitors. *J Med Chem.* 2017;60(4):1262-1271. doi: 10.1021/acs.jmedchem.6b00990.
39. Kwon DH-M, Borno HT, Cheng HH, Zhou AY, Small EJ. Ethnic disparities among men with prostate cancer undergoing germline testing. *Urologic Oncology: Seminars and Original Investigations.* 2020/03/01/ 2020;38(3):80.e81-80.e87. doi: https://doi.org/10.1016/j.urolonc.2019.09.010.
40. Nicolosi P, Ledet E, Yang S, et al. Prevalence of Germline Variants in Prostate Cancer and Implications for Current Genetic Testing Guidelines. *JAMA Oncology.* 2019;5(4):523-528. doi: 10.1001/jamaoncol.2018.6760.