U.PORTO

MESTRADO
MEDICINA E ONCOLOGIA MOLECULAR

Simultaneous detection of germline and somatic mutations in DNA repair genes by next generation sequencing of tumor samples from metastatic prostate cancer patients

Ariane Falconi

M

2020

FACULDADE DE **MEDICINA**

Master's in Medicine and Molecular Oncology

Simultaneous detection of germline and somatic mutations in DNA repair genes by next generation sequencing of tumor samples from metastatic prostate cancer patients

Ariane Falconi

This dissertation was submitted to Faculty of Medicine of University of Porto as part of the master's degree in Medicine and Molecular Oncology

Supervisor: Manuel António Rodrigues Teixeira, MD, PhD
Instituto de Ciências Biomédicas Abel Salazar – ICBAS-UP
Instituto Português de Oncologia do Porto – IPO Porto

Co-supervisor: Paula Cristina Martins Santos Paulo, PhD
Instituto Português de Oncologia do Porto – IPO Porto

Cancer Genetics Group - IPO Porto Research Center (CI-IPOP)

D E D I C A T Ó R I A

Dedico este trabalho aos meus pais, Henrique e Cristina, por serem meus pilares da vida, me ensinarem o caminho da honestidade, do conhecimento e a buscar sempre uma versão melhor de mim mesma.
Dedico também aos meus irmãos e à minha pequena Eva, que completam minha vida e meu time de incentivadores.

“[Genetics] is the most beautiful way of thinking”

Mary-Claire King, PhD

Lasker~Koshland Award interview

A G R A D E C I M E N T O S

Agradeço primeiramente à Deus, pelo dom da vida e pelo privilégio da força e resiliência para enfrentar e ultrapassar todos os obstáculos.

A todos os envolvidos na realização desse projeto, especialmente ao meu orientador, Profº Dr. Manuel Teixeira, pela confiança, mentoria e pela enorme oportunidade que me deu, e à Dra Paula Paulo, pela coorientação, parceria e pelo envolvimento diário em tudo o que diz respeito a esse trabalho. À Andreia, por toda a ajuda e ensinamentos com as análises e por plantar a sementinha da bioinformática aqui dentro. Às minhas colegas Márcia, Anita, Catarina e Carla Pinto, por estarem envolvidas diretamente na execução de parte deste trabalho, e aos meus colegas do Serviço de Genética/Grupo de Oncogenética, Maria, Marta, Ana Barbosa, Isabel, Carla Escudeiro, Manela, Joana, Sara, Inês, Sónia, João e demais integrantes, pelo companheirismo, por me receberem e me acolherem de braços abertos. Obrigada a todos por me fazerem me sentir em casa.

Aos pacientes e às suas famílias, por nos darem permissão de estudar suas doenças e contribuírem com a ciência em um momento tão frágil de suas vidas. A todos os profissionais do IPO Porto que estiveram envolvidos nas consultas, no recrutamento dos pacientes, nas colheitas e transporte das amostras.

Aos meus colegas do MMOM e a todos os professores que passaram por nós e foram essenciais para que nos tornássemos os profissionais que somos hoje.

"Enquanto houver vocês do outro lado, aqui do outro eu consigo me orientar"

Aos meus pais, Henrique e Cristina, por terem me dado todo o amor e apoio necessário para estar aqui e completar esse mestrado. Sem vocês, eu não conseguiria nem imaginar nada disso. Vocês foram essenciais para a realização desse sonho, amo vocês!

Aos meus irmãos, Amarílis, Luís Henrique e Adriano, por completarem minha rede de apoio, me incentivarem em todos os momentos e estarem sempre ao meu lado. À minha pequena Eva, por ter me feito tia-dinda e ter me completado de uma forma que eu nem

imaginaria ser possível. Aos meus avós, Salete, João, Lourdes e Neclair, aos meus tios Neto, Vanessa e Edu, e primos Ivan e Pedro, e à minha afilhada e prima Nathália. Muito obrigada!

À Fernanda, pela nossa amizade, pela paciência, por acreditar em mim, me incentivar, ser meu ombro amigo e meu refúgio nessa minha difícil escolha de morar fora. Obrigada por estar sempre presente, por me fazer presente e "segurar minha onda" mesmo de tão longe. Acho que encontrei o que eu estava à procura!

Às minhas amigas Mari, Ivete, Priscilla, Silvana e Juliana, por serem as melhores amigas que alguém pode ter, por me fazerem rir, me escutarem e estarem presentes em cada etapa vivida aqui. Aos meus sempre companheiros de laboratório, Edu, Bia, Jacque e Luana, morro de saudades de dividir meus dias com todos vocês!

À Vivi, por ser "minha mentora", ter me incentivado a fazer o mestrado, a vir para fora do Brasil, a seguir com a minha carreira e sempre acreditar no meu potencial, mesmo quando eu mesma não acreditava.

À Bianca, por todo apoio e amizade durante todos esses anos e, principalmente, nos últimos meses aqui em Portugal. À Lívia, por ser minha companheira de casa por quase um ano e me aturar durante a quarentena. Obrigada por serem minha família brasileira no Porto.

Às minhas sempre amigas de faculdade e da vida, Laís e Steh, às minhas amigas e irmãs, Amanda e Mariana, e aos nossos mascotes, Guto, Pedro e Lucas, obrigada por todo o carinho da vida! Aos meus colegas brasileiros no Porto e também extensão da minha família: Sabrina, Diego, Fran, Ju, Giu e Julyana.

A todos que não citei, mas sabem que estão no meu coração, meu muito obrigada!

"Tem hora que a gente se pergunta, por que é que não se junta tudo numa coisa só?"

TABLE OF CONTENTS

FIGURE INDEX

TABLE INDEX

ABBREVIATIONS

ACMG – American College of Medical Genetics and Genomics

ADT – Androgen deprivation therapy

AMP – Association for Molecular Pathology

ASCO – American Society of Clinical Oncology

BAM – Binary alignment map

BER – Base excision repair

BPH – Benign prostatic hyperplasia

CAP – College of American Pathologists

CCD – Charge-coupled device

CE – Capillary electrophoresis

cfDNA – Cell-free DNA

CNV – Copy-number variation

CRPC – Castration-resistant prostate cancer

CTC – Circulating tumor cell

ctDNA – Cell-free circulating tumor DNA

CZ – Central zone (of prostate gland)

ddPCR – Digital droplet PCR

DDR – DNA-damage repair

DIN – DNA Integrity Number

dNTPs – Deoxyribonucleotide triphosphates

DRE – Digital rectal examination

DSB – Double strand break

EBRT – External beam radiation therapy

EDTA – Ethylenediamine tetra acetic acid

EMA – European Medicines Agency

ESCAT – ESMO Scale for Clinical Actionability of Molecular Targets

ESMO – European Society for Medical Oncology

ExAC – Exome Aggregation Consortium

FDA – U.S. Food and Drug Administration

FFPE – Formalin-fixed paraffin-embedded

gDNA – germline DNA

gnomAD – Genome Aggregation Database

HBOC – Hereditary breast and ovary cancer

HE – Hematoxylin-eosin

HgPIN – High grade PIN

HGVS – Human Genome Variation Society

HRR – Homologous recombination repair

IGV – Integrative Genomics Viewer

INFARMED – Autoridade Nacional do Medicamento e Produtos de Saúde, I.P.

IPO-Porto – Instituto Português de Oncologia do Porto

IQN Path – International Quality Network for Pathology

LFL – Li-Fraumeni-Like Syndrome

LFS – Li-Fraumeni Syndrome

LHRH – Luteinizing hormone-releasing hormone

LOF – Loss of function

LOH – Loss of heterozygosity

MAF – Minor allele frequency

mCRPC – Metastatic castration-resistant prostate cancer

MRD – Minimal residual disease

NCCN – National Comprehensive Cancer Network

NGS – Next Generation Sequencing

OR – Odds ratio

OS – Overall survival

PARP – poly-ADP ribose polymerase

PARPi – poly-ADP ribose polymerase inhibitor

PBS – Phosphate-buffered saline

PCR – Polymerase chain reaction

PF – Peripheral zone (of prostate gland)

PIN – Prostatic intraepithelial neoplasia

PK – Proteinase K

PrCa – Prostate cancer

PSA – Prostate Specific Antigen

RF – Replication fork

RP – Radical prostatectomy

rPFS – Radiographic progression-free survival

SBS – Sequencing by synthesis

SNP – Single nucleotide polymorphism

SNV – Single nucleotide variant

SSB – Single strand break

tDNA – tumor DNA

TLB – DNA tissue lysis buffer

TOPARP – Trial of PARP Inhibition in Prostate Cancer

TZ – Transitional zone (of prostate gland)

UTR – Untranslated region

VCF – Variant call format

A B S T R A C T

Prostate cancer is the second most incident type of cancer in men worldwide and the sixth cause of death. In Portugal, it is the most incident type of cancer in men and the third cause of death. However, inevitably a high percentage of these patients develops therapy resistance, progressing of localized disease to metastatic disease. Recently approved by the FDA, PARP inhibitors (PARPi) are an alternative therapy for patients with metastatic disease that harbors deleterious or likely deleterious variants in homologous recombination repair genes, somatic or germline.

The main aim of this study was to evaluate the percentage of patients with metastatic prostate cancer, referenced to IPO Porto, that could benefit from targeted therapy with PARPi. To achieve this, an NGS method was stablished, using a customized gene panel composed by homologous recombination DNA repair genes that were included in the phase III clinical trial PROfound and other DNA repair genes that were included in phase II clinical trials.

For this study, 40 tumor samples of metastatic prostate cancer patients were analyzed. After DNA extraction, the library preparation and next generation sequencing were performed according to the manufacturer instructions. NGS data was analyzed using DRAGEN and VarAFT software. Variant filtering was performed according to international guidelines for variant filtering and interpretation.

A total of 1688 variants were found, with 136 of them being potentially oncogenic. Considering the 14 genes of PROfound clinical trial, 8 patients (20%) presented deleterious variants in 4 actionable genes. *CDK12* is the most frequently alterated gene, with one stop variant found in one patient and 4 variants, two frameshift and two splicing, found in three patients, followed by *BRCA2*, with two frameshift variants found in two patients, *CHEK2*, with one missense variant found in two patients and *ATM*, with one frameshift and one splicing variant found in two patients. Of 11 variants found, 4 of them are germline, confirmed in constitutional DNA, namely the two in *BRCA2* and the two in *CHEK2*, and the other 7 are somatic.

These results show that a significative number of patients with metastatic prostate cancer may benefit from PARPi treatment, a therapy that proved to be effective in increasing overall survival, decreasing the risk of death with few side effects. The NGS approach in tumor samples allows to maximize the identification of patients that carry actionable variants and to identify families in risk of hereditary disease. The possibility of customization enables

the analysis of actionable genes for a given treatment and/or type of cancer, increasing the speed of analysis and the feedback for the patient.

R E S U M O

O cancro da próstata é o segundo tipo de cancro mais comum em homens e a sexta maior causa de morte no mundo. Em Portugal, representa o tipo de cancro mais incidente e a terceira maior causa de morte por cancro. No entanto, inevitavelmente, grande parte dos pacientes desenvolve resistência à terapia, progredindo para cancro metastático. Recentemente aprovados pela FDA, os inibidores da PARP (iPARP) são uma alternativa de tratamento para pacientes com cancro da próstata metastático portadores de variantes deletérias em genes de reparação do DNA envolvidos na recombinação homóloga, tanto a nível germinativo como somático.

O principal objetivo deste estudo foi avaliar a percentagem de pacientes com cancro da próstata metastático, referenciados ao IPO do Porto, que possam beneficiar de tratamento específico com iPARP. Para tal foi estabelecido um método de análise baseado em sequenciação de nova geração (NGS), usando um painel de genes customizado, de forma a identificar doentes que são portadores de mutações em genes de reparação do DNA a nível germinativo e/ou somático.

Neste estudo analisaram-se amostras tumorais de 40 casos de cancro metastático da próstata. Após extração de DNA, a preparação e sequenciação das librarias de NGS foi efetuada de acordo com as recomendações. Os dados de NGS foram analisados usando os *software* DRAGEN e VarAFT. As variantes foram filtradas de acordo com as *guidelines* internacionais para interpretação e reporte de variantes em cancro.

No total, foram identificadas 1688 variantes, das quais 136 são potencialmente oncogénicas. De acordo com os genes analisados no ensaio clínico PROfound (fase III), 8 pacientes (20%) apresentaram variantes deletérias em 4 genes de resposta à terapia com iPARP. O *CDK12* é o gene mais frequentemente alterado, com uma variante *stop* presente em 1 paciente e quatro variantes, duas *frameshift* e duas *splicing*, presentes em três pacientes, seguido do *BRCA2*, com duas variantes *frameshift* presentes em dois pacientes, *CHEK2*, com uma variante identificada em dois pacientes, e *ATM*, com uma variante *frameshift* e uma variante *splicing* presentes em dois pacientes. Das 11 variantes com potencial resposta a iPARP, 4 são germinativas, validadas em DNA constitucional, sendo duas *frameshift* no *BRCA2* e duas *missense* no *CHEK2*.

Estes resultados mostram que um número significativo de doentes com cancro da próstata metastático pode beneficiar de terapia dirigida com iPARP, terapia esta que se mostrou efetiva no aumento da sobrevida, diminuição do risco de morte e com poucos efeitos secundários. A abordagem de NGS por painel customizado ao nível do DNA tumoral

permite, por um lado, maximizar a identificação de doentes portadores de variantes preditivas de resposta e, por outro, identificar famílias em risco de doença hereditária. A possibilidade de customização permite a análise de genes acionáveis específicos para tratamento e/ou tipo de cancro, aumentando a velocidade de análise e resposta ao doente.

Prostate anatomy and histology

The prostate is an exocrine gland that belongs to the male reproductive system. Its role is to produce a thin, milky and slightly alkaline fluid that constitutes the seminal fluid, responsible for maintaining the nutrition of sperm. Its alkaline pH is also responsible for neutralizing the acidity of other seminal fluids and the vagina, increasing the mobility and longevity of sperm (1,2).

The prostate is a retroperitoneal organ, normally weighing 30 to 40g in an adult, and has the shape of an inverted cone, with the base near the bladder neck and the apex distally at the urogenital diaphragm. The prostatic urethra passes through the center of the gland (3,4).

Despite the distinction of right and left lobes still being used as a reference of the palpation of a midline furrow, recent anatomic theories divide the prostate into inner or central and outer or peripheral regions (Figure 1). Benign prostatic hyperplasia (BPH) occurs mostly in the inner zone, while the outer zone is the primary site for carcinomas to occur. The prostate is divided into four zones: anterior fibromuscular stroma (smooth muscle and dense fibrous tissue), central zone (surrounds the ejaculatory ducts), peripheral zone and pre-prostatic region. The peripheral zone is the largest zone of the prostate and contains most of the glandular tissue. The pre-prostatic region has an area known as transition zone, which is most affected by BPH. The central or inner zone of the prostate has an area that histologically mimics high-grade prostatic intraepithelial neoplasia (HGPIN), making the identification of this area of extremely importance in needle biopsy scenario, in order to avoid false positive results (3,4).

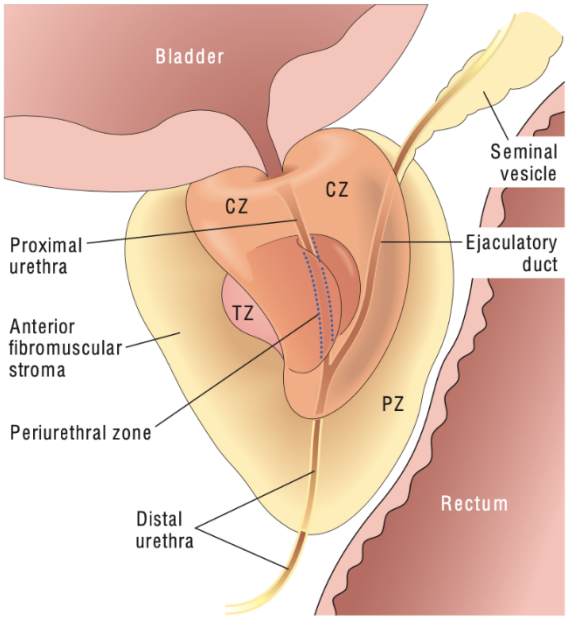


Figure 1. Adult prostate. The normal prostate contains several distinct regions, including a central zone (CZ), a peripheral zone (PZ), a transitional zone (TZ), and a periurethral zone (4).

Histologically, the prostate consists of epithelial and stromal cells. Epithelial cells form glands and are divided in urothelial cells, secretory cells (secrete prostatic-specific antigen), basal cells (extremely important in differential diagnosis), and neuroendocrine cells. Stromal cells are skeletal and smooth muscle cells, fibroblasts, nerves and endothelial cells (2,3).

Prostate cancer epidemiology

Prostate cancer (PrCa) is the second most common cancer in men, after lung cancer, and the sixth cause of death among men worldwide. In 2018, 1,276,106 new cases were registered, representing 7.1% of all cancer cases in men, and 358,989 deaths, corresponding to 3.8% of all male cancer deaths (Figure 2). According to GLOBOCAN projections, in 2040 there will be 2,293,818 new PrCa cases and 737,994 PrCa deaths, representing increases of 79.8% and 105.6%, respectively, in comparison with 2018 (5–7).

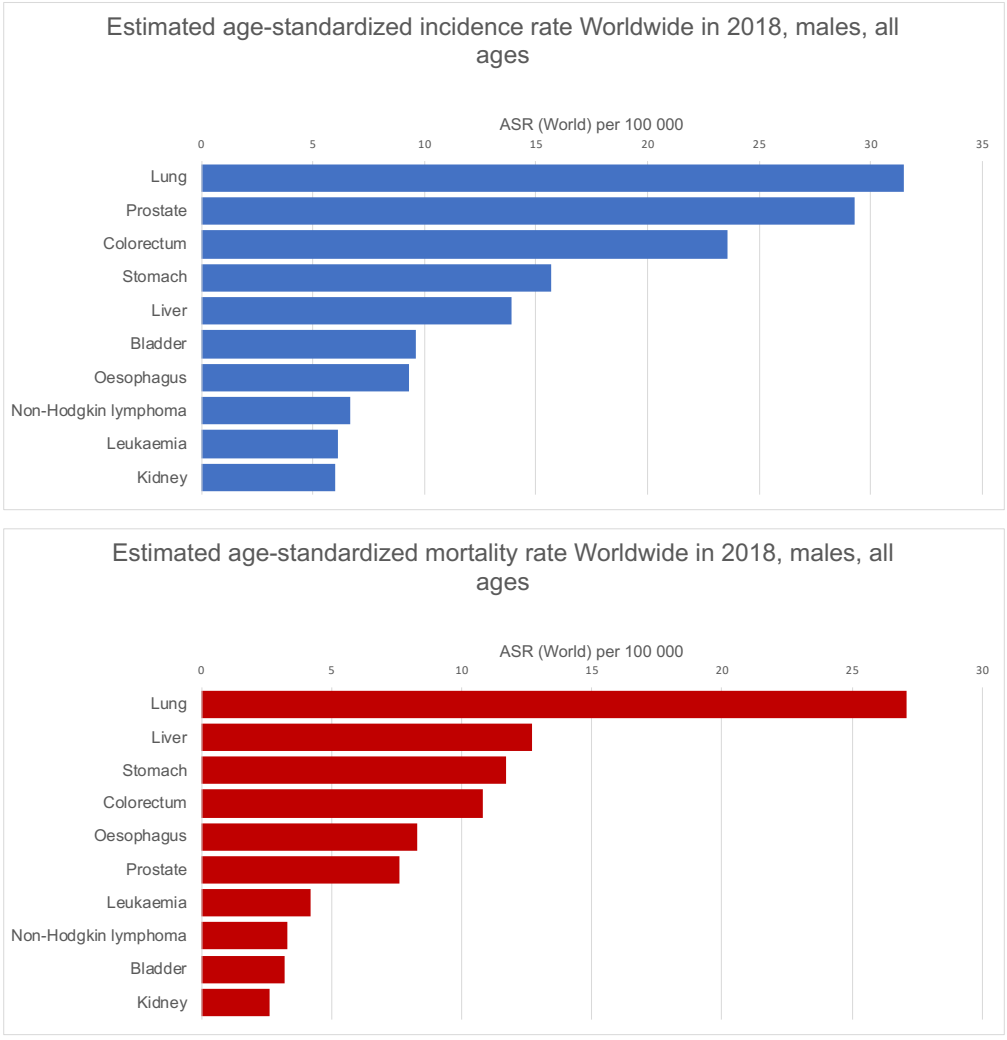


Figure 2. Estimated age-standardized incidence and mortality rates in 2018, worldwide, males, all ages. Data obtained from GLOBOCAN 2018

In Portugal, PrCa is the most incident type of cancer and the third cause of death in males, with 6,609 new cases and 1,879 deaths estimated in 2018 (Figure 3). Projections for 2040 point to 8,362 new cases and 2,937 deaths, an increase of 26.5% and 56.3%, respectively (5,7,8).

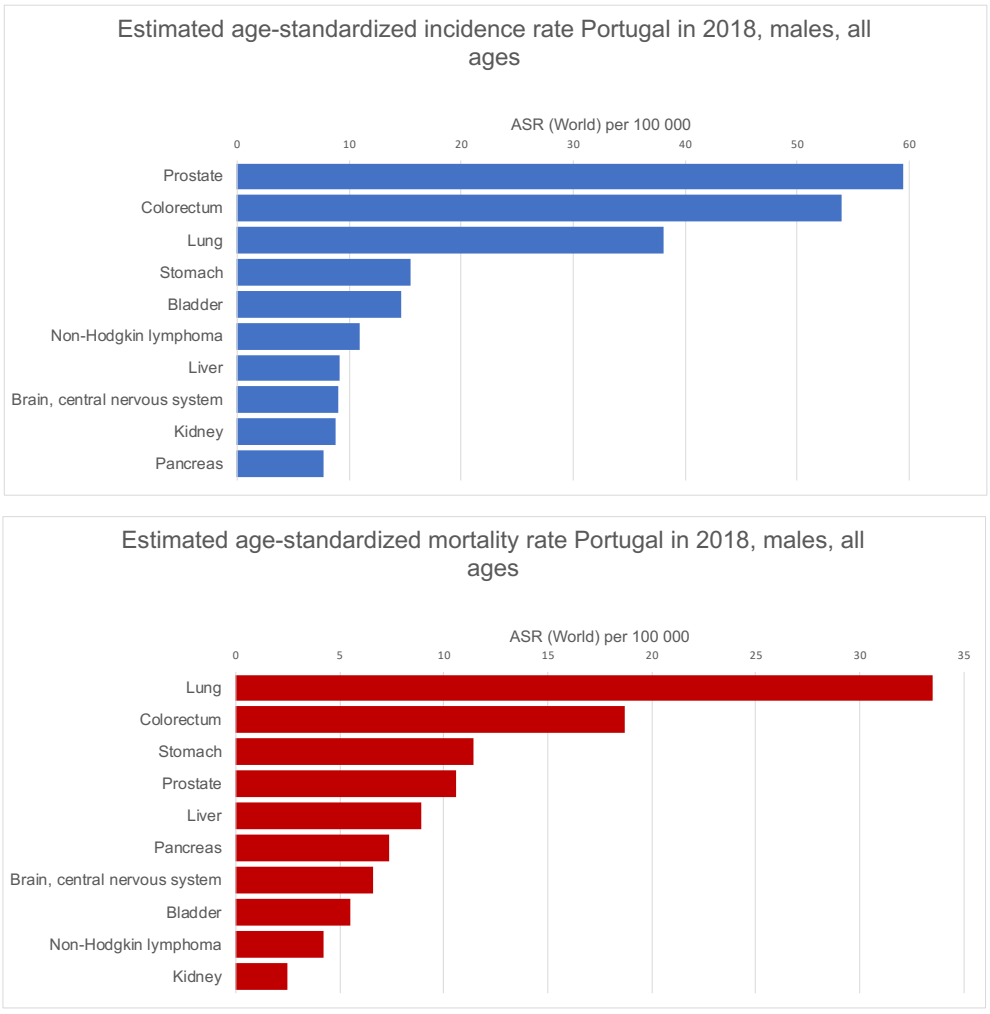


Figure 3. Estimated age-standardized incidence and mortality rates in 2018, Portugal, males, all ages. Data obtained from GLOBOCAN 2018

Prostate cancer etiology

Prostate cancer is a disease most frequently diagnosed in men aged over 50 years and rising sharply in incidence with each decade thereafter (9). However, studies have shown that the prostate of man with 20s to 40s already have histologic foci of prostate cancer, but, usually, these foci do not progress to clinically detectable disease. In fact,

prostate cancer is a highly heterogeneous disease, with several independent histologic foci commonly found in the prostate gland (10).

The early stages of the disease are, usually, asymptomatic and is common to have an indolent course, as the most frequent symptoms are similar to benign prostatic hyperplasia, including difficulty and increased frequency of urination and nocturia (6).

The malignant disease is thought to initiate as a prostatic intraepithelial neoplasia (PIN), followed by localized PrCa, which then progresses to advanced prostate adenocarcinoma with local invasion, with its final step as metastatic disease (Figure 4). The first organs to which cancer cells metastasize are lymph nodes adjacent to the primary tumor, followed by liver, lungs and bones. Bone metastases are often painful, cause hypercalcemia and frequent fractures, due to the presence of osteoblastic lesions and osteolytic features (11). Molecular and cytogenetic analyses of metastatic foci show that multiple metastases of the same patient are clonally related, indicating that advanced prostate cancer is monoclonal. These findings suggest that metastatic prostate cancer may arise from selective advantages of one clone during cancer progression (10).

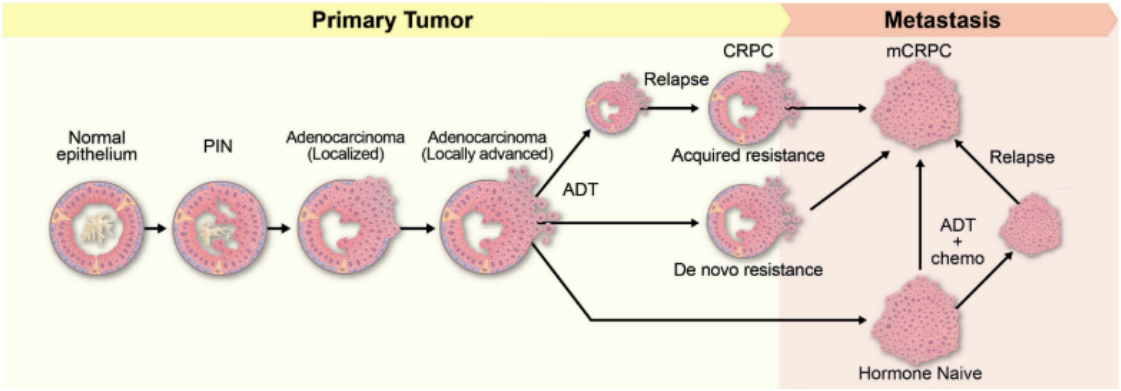


Figure 4. Schematic evolution of prostate cancer. Prostate cancer, presumably, initiates as a PIN lesion, progresses to a localized adenocarcinoma and then to locally advanced adenocarcinoma, after which metastase can occur. For metastatic disease, the best treatment option, known so far, is the androgen deprivation therapy (ADT), which can cause resistance, resulting in castration-resistance prostate cancer (CRPC). This resistance can be due to either relapse (acquired resistance) or "de novo" resistance. On the other hand, some patients can progress to metastatic disease before treatment (hormone naïve metastasis), while others have already metastatic disease at diagnosis, in which scenarios ADT is combined with chemotherapy. Inevitably, ADT therapy fails, and the disease progresses to a metastatic castration-resistant prostate cancer (mCRPC) stage, a condition without curative treatment options (11).

Established risk factors

Age and ethnicity

PrCa incidence increases with age and has an exponential rise after the age of 50 in white men and after the age of 40 in African American men. African American men, West African ancestry from the Caribbean and South American men have higher risk for PrCa than Caucasian men. The lowest incidence of PrCa is in Asian men, not only due to their genetic susceptibility, but also related with diet, lifestyle and environmental factors (9,12).

Family history

The risk for PrCa increases with the number of affected relatives and with closer genetic degree, being higher with an affected first degree relative than with an affected second degree relative, and further increases with younger age at onset of the closest relative (13,14). Among first degree relatives, a man with an affected brother is at higher risk of PrCa than a man with an affected father (15).

Prostate cancer diagnosis

According to the NCCN Clinical Practice Guidelines in Oncology – Prostate Cancer Early Detection version 2.2020, prostate cancer early diagnosis begins with baseline evaluation, including family history of cancer, family history of prostate disease and cancer early detection, race and personal or family history of high-risk mutations. After the baseline evaluation, the clinician has to assess the risk for prostate cancer based on Prostate-Specific Antigen (PSA) levels and digital rectal examination (DRE), complemented or not by other screening tests. Prostate biopsy is recommended for PSA levels higher than 3 ng/mL, confirmed by a second exam (imaging-based and/or biomarker-positive), and/or suspicious or very suspicious DRE (16).

According to the ESMO Clinical Practice Guidelines (2015), screening for prostate cancer should be based on age, ethnicity, family history, PSA level, free/total PSA ratio and findings on digital rectal examination, which may lead to a needle biopsy for diagnosis (17).

The NCCN guidelines (Version 2.2020) for Prostate Cancer advise the clinician to obtain and review diagnostic prostate biopsies and to inquire about known germline mutations and family history (16). If the family history and/or germline mutations are known

or the histology is suspicious/revealing, the recommendation is to perform genetic testing with later genetic counselling if necessary. If family history is unknown or not significant and the histology is not suspicious, the clinician may consider genetic testing, with genetic counselling if necessary. The workflow is demonstrated in figure 5.

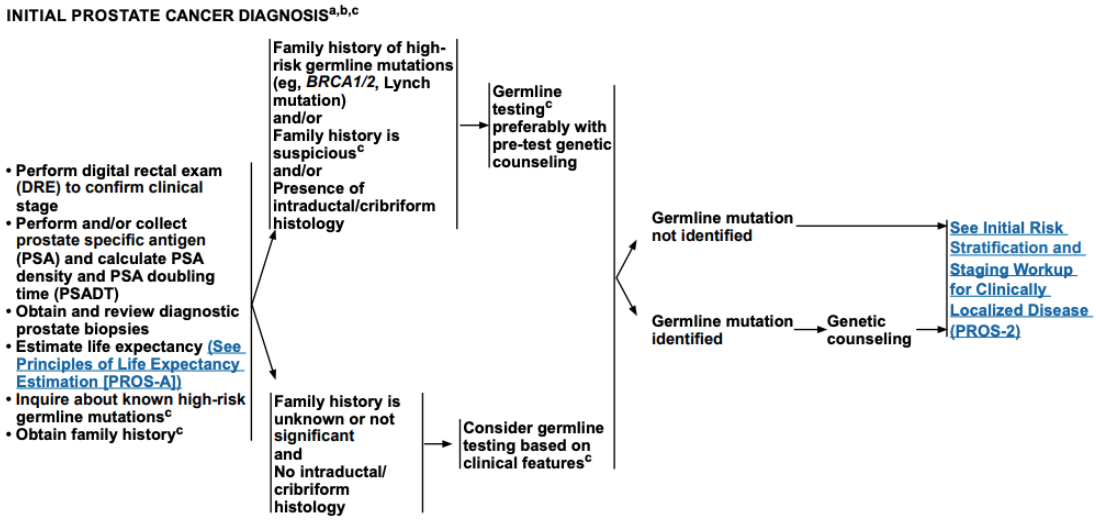


Figure 5. Initial prostate cancer diagnosis according to NCCN Guidelines Version 2.2020 (16)

Prostate-specific antigen

The prostate-specific antigen (PSA) is a glycoprotein synthesized in the ductal epithelium and prostatic acini and is secreted into the lumina to become a component of seminal plasma, helping to trap spermatozoa at ejaculation. It is expressed in normal prostate tissue, as well as in hyperplastic and malignant tissue (3,4).

Total serum PSA has been used as a biomarker for prostate cancer, where an increase of those levels is interpreted as indicator of malignant disease. However, increase in PSA levels are not specific of prostate cancer (3). As mentioned above, hyperplastic prostate, a benign prostate alteration, also produces high levels of serum PSA, as well as prostatitis, ejaculation, manipulations of the prostate and other benign conditions. Populational screening based on PSA levels prevents PrCa at the expenses of over-diagnosis and overtreatment (6,17).

Although PSA levels are not specific of PrCa, they are important for monitoring the disease and guide treatment decisions (4,17).

Digital rectal examination

The findings of DRE that are suspicious for cancer are asymmetry, induration, and discrete hard nodules. However, most prostate cancers detected by needle biopsy are non-palpable, due to localization (anteriorly) or small size, leading to a low sensitivity of DRE. The reproducibility of DRE it is also low, with poor interobserver reproducibility among urologists (3,4).

Needle biopsy and grading system

The needle biopsy of the prostate is guided by transrectal ultrasound, where 10 to 12 tissue samples are systematically collected in a grid-like pattern, using the sextant-biopsy technique (3,18).

In histological analysis, the system used to measure the aggressiveness of prostate cancer is the Gleason grading system, or Gleason score. The Gleason score was created by Donald F. Gleason in 1966 and is based on histological patterns, divided in five different patterns. Initially, the Gleason score of a prostate carcinoma was the worst grade of that tumor. Over time, the grading system has changed, and nowadays, the Gleason score is defined as the sum of the two most prevalent patterns in the tissue sample. For each one of the two patterns, a score of 1 to 5 is attributed, giving the Gleason score a range from 2 to 10. The first grade of Gleason score indicates de most prevalent pattern and the second grade the second most prevalent. A score of 4 + 3 = 7 means that the tumor is heterogeneous, and the most prevalent pattern is graded by 4 (3,18).

Recently, the Gleason system has been compressed into Gleason groups, varying from 1 to 5, according to Gleason score, as shown in Table 1 (16). In 2014, a consensus conference revised the Gleason grading system, implementing some major changes. One of these most important changes takes into account the tumor behavior and consists in the differentiation between Gleason scores 3+4 and 4+3. Tumors with score 3+4 (7a) are classified as grade 2, while tumors with score 4+3 (7b) are classified as grade 3. Scores 7b or higher are associated with increased pathologic stage and disease progression (19,20).

Table 1. AJCC grade group for grading prostate cancer, based on Gleason Score

Gleason grading group		
Grade Group	**Gleason Score**	**Gleason Pattern**
1	≤6	≤3+3
2	7	3+4
3	7	4+3
4	8	4+4, 3+5, 5+3
5	9 or 10	4+5, 5+4, 5+5

Genetics of prostate cancer

Prostate cancer has a high and heterogeneous genetic component. The Cancer Genome Atlas (TCGA) Network research group, in a study with 333 cases of localized prostate cancer (21), defined 7 molecular subtypes of prostate cancer. Four of them are related to fusions in the *ETS*-gene family, which are present in 53% of the tumors. They classified the tumors in *ERG, ETV1, ETV4* and *FLI1* fusions subtypes, with *TMPRSS2* gene being the most common partner. The other three subtypes are related to mutations in *SPOP, FOXA1* and *IDH1* genes, with the *SPOP* mutations being mutually exclusive with all *ETS* fusions. Despite this molecular classification, about 26% of the tumors appear to be driven by non-identified molecular alterations or by one or more alterations that co-occur with the ones that define the subtypes.

The TCGA study also revealed that 19% of the patients had alterations in several DNA repair genes (21) and this observation is in accordance with a study made by Robinson and colleagues (22), where the tumors of 150 patients with mCRPC were submitted to whole-exome sequencing. Of these 150 patients, 22.7% presented alterations in DNA repair genes, with 20% of them being in *BRCA1*, *BRCA2* or *ATM* genes.

Associations of PrCa and hereditary breast and ovary cancer (HBOC) syndrome have been made. A study with 993 individuals who had received genetic counseling and had deleterious or suspected deleterious mutations in *BRCA1* or *BRCA2* showed that, besides breast and ovary cancer, pancreatic and prostate cancer were the most frequent types of cancer found in men with *BRCA2* mutations (23). Another study evaluated the risk of 490 *BRCA1/2* mutation carrier families to develop cancer other than breast and ovarian cancers. This study confirmed that the increased risk is confined to men with *BRCA2* mutations and

that mutations in *BRCA2* are associated with prostate cancer (24). Mutations in *BRCA1/2* have been found in approximately 1.4% of all PrCa cases and are associated with prostate cancer aggressiveness and mortality (25,26). In mCRPC the frequency of mutations in *BRCA2* is considerably higher, of about 3-5% (25,27).

More recently, in a multicenter study gathering 692 men with mCRPC, Pritchard et al (2016) performed genetic screening for germline mutations in 20 DNA repair genes (28). Of 692 men, 82 (11.8%) had at least one germline mutation in one of 16 DNA-repair genes, including *BRCA2* (44% of total mutations), *ATM* (13%), *CHEK2* (12%), *BRCA1* (7%), *RAD51D* (4%) and *PALB2* (4%). All the 73 patients with information regarding the Gleason score of the primary tumor had a score over 6, the majority (77%) with Gleason Score of 8-10 (28,29).

Other studies showed a correlation of prostate cancer with Lynch Syndrome (LS), where men with LS are two to three times more likely to develop prostate cancer, with or without prior colorectal cancer (30). These data were corroborated with other molecular studies, where it was observed that 73% of prostate cancers in patients harboring alterations in MMR genes had MMR deficiency, MSI-H (microsatellite instability – high) profile, and lack of MSH2 and MSH6 protein expression in tumor immunohistochemistry. Furthermore, studies showed that tumors of *MHS2* mutated carriers had more incidence of MMR deficiency compared with other MMR alterated gene carriers (31).

Treatment

Factors affecting treatment decision

Life expectancy

Life expectancy estimation is a critical information to treatment decision-making in prostate cancer. The estimation is possible for groups of men but is challenging for individuals. There are tools recommended by NCCN guidelines that help to estimate life expectancy and how to apply adjustments using the clinician's assessment of overall health (16).

Risk stratification for localized disease

Measures the risk of localized disease to progress to metastatic disease. General health information and comorbidities that could affect treatment should be assessed.

Patients that are not suitable for curative treatment due to poor health conditions normally do not require staging investigation (16,17).

Initial risk stratification is divided in five groups, namely “very low”, “low”, “intermediate”, “high” and “very high” (Table 2). The “intermediate” group is divided in “favorable intermediate” and “unfavorable intermediate” groups. Each group is characterized by specific clinical and pathologic features, which will influence the imaging exams, germline testing and molecular analysis to be performed, overall conditioning the choice of the initial therapy (16).

Table 2. NCCN initial risk stratification and staging workup for clinically localized disease

Risk Group	Clinical/Pathologic Features			Imaging[f,g]	Germline Testing[c]	Molecular/ Biomarker Analysis of Tumor[c]
Very low[d]	Has all of the following: • T1c • Grade Group 1 • PSA <10 ng/mL • Fewer than 3 prostate biopsy fragments/cores positive, ≤50% cancer in each fragment/core[e] • PSA density <0.15 ng/mL/g			Not indicated	Recommended if family history positive or intraductal/cribriform histology See PROS-1	Not indicated
Low[d]	Has all of the following but does not qualify for very low risk: • T1–T2a • Grade Group 1 • PSA <10 ng/mL			Not indicated	Recommended if family history positive or intraductal/cribriform histology See PROS-1	Consider if life expectancy ≥10 y[j]
Intermediate[d]	Has all of the following: • No high-risk group features • No very-high-risk group features • Has one or more intermediate risk factors (IRF): ▸ T2b–T2c ▸ Grade Group 2 or 3 ▸ PSA 10–20 ng/mL	Favorable intermediate	Has all of the following: • 1 IRF • Grade Group 1 or 2 • <50% biopsy cores positive[e]	• Bone imaging[h]: not recommended for staging • Pelvic ± abdominal imaging[i]: recommended if nomogram predicts >10% probability of pelvic lymph node involvement • If regional or distant metastases are found, see PROS-8	Recommended if family history positive or intraductal/cribriform histology See PROS-1	Consider if life expectancy ≥10 y[j]
		Unfavorable intermediate	Has one or more of the following: • 2 or 3 IRFs • Grade Group 3 • ≥50% biopsy cores positive[e]	• Bone imaging[h]: recommended if T2 and PSA >10 ng/mL • Pelvic ± abdominal imaging[i]: recommended if nomogram predicts >10% probability of pelvic lymph node involvement • If regional or distant metastases are found, see PROS-8	Recommended if family history positive or intraductal/cribriform histology See PROS-1	Consider if life expectancy ≥10 y[j]
High	Has no very-high-risk features and has at least one high-risk feature: • T3a OR • Grade Group 4 or Grade Group 5 OR • PSA >20 ng/mL			• Bone imaging[h]: recommended • Pelvic ± abdominal imaging[i]: recommended if nomogram predicts >10% probability of pelvic lymph node involvement • If regional or distant metastases are found, see PROS-8	Recommended	Consider if life expectancy ≥10 y[j]
Very high	Has at least one of the following: • T3b–T4 • Primary Gleason pattern 5 • 2 or 3 high-risk features • >4 cores with Grade Group 4 or 5			• Bone imaging[h]: recommended • Pelvic ± abdominal imaging[i]: recommended if nomogram predicts >10% probability of pelvic lymph node involvement • If regional or distant metastases are found, see PROS-8	Recommended	Not routinely recommended

Treatment options

There are some curative options for treating localized disease, such as radical prostatectomy (RP), external beam radiation therapy (EBRT) and brachytherapy. In general terms, patients with very low, low and favorable intermediate risk have the option of active surveillance besides RP, EBRT and brachytherapy. On the other hand, patients with unfavorable intermediate risk, high and very high risk follow a more intensive protocol, which includes RP, EBRT, brachytherapy and Androgen Deprivation Therapy (ADT), with a closer monitoring (16,17).

Patients must be aware of the benefits and risks of the treatment options, such as sexual dysfunction, infertility, and bowel and urinary problems. It is advised that men consult both a urologist and a radiation oncologist for different treatment options and be informed of benefits and risks of each treatment (16,17).

An important characteristic of prostate cancer is its hormone responsiveness. The use of luteinizing hormone-releasing hormone (LHRH) drugs (chemical castration), and orchiectomy (surgical castration) are approaches for androgen deprivation therapy (ADT). Patients may develop resistance to this therapy, originating castration-resistant prostate cancer (CRPC) or metastatic castration-resistant prostate cancer (mCRPC), for whom treatment is difficult and bears significant side-effects. In these cases, ADT is indicated to maintain castrate serum levels of testosterone, and it is used in addition to other treatment options. The first line treatment includes Abiraterone or Enzalutamide, two drugs that also decrease testosterone levels, Docetaxel, a chemotherapeutic drug, and Sipuleucel-T, an immunotherapy drug. Under certain circumstances, Radium-223 and Mitoxantrone can be used, as well as other secondary hormone therapy (16,17).

The leading cause of death associated with prostate cancer is metastatic disease, for state treatment options are limited (11). Unfortunately, about 7-15% of the patients diagnosed with PrCa have already metastatic disease (32), for which the 5-year survival rate is only about 30% (33). Therefore, efforts have been made to understand the metastatic process and improve treatment options (11). As a result of these efforts, PARPi therapy was recently approved as a second line treatment for patients with mCRPC.

PARP inhibitors as treatment for HRR-deficient mCRPC

Mechanism of action of PARP inhibitors

Members of the poly-ADP ribose polymerase (PARP) enzyme family are essential to DNA damage repair. The most studied member of the PARP family, PARP-1, acts as a sensor of DNA damage, being involved in DNA repair initiation of both single strand break (base excision repair – BER) and double strand break (HRR and NHEJ – non-homologous end joining) repair pathways (34). Nevertheless, its involvement in the SSB/BER pathway is more established. PARP-1 is recruited to the DNA damage site by SSB intermediates. PARP-1 function is to PARylate itself and other substrates to recruit repair proteins to the damage site and, then, correct the damage (Figure 6) (35). With the inhibition of PARP, SSBs may be converted to double strand breaks (DSB) during replication and then collapse

the replication fork. Cells that have lack of function of BRCA proteins are sensitive to PARP inhibition, probably caused by multiple mechanisms, including the synthetic lethality, as a consequence of failure in DNA repair and replication arrest caused by lack of chromatin remodeling (35–37).

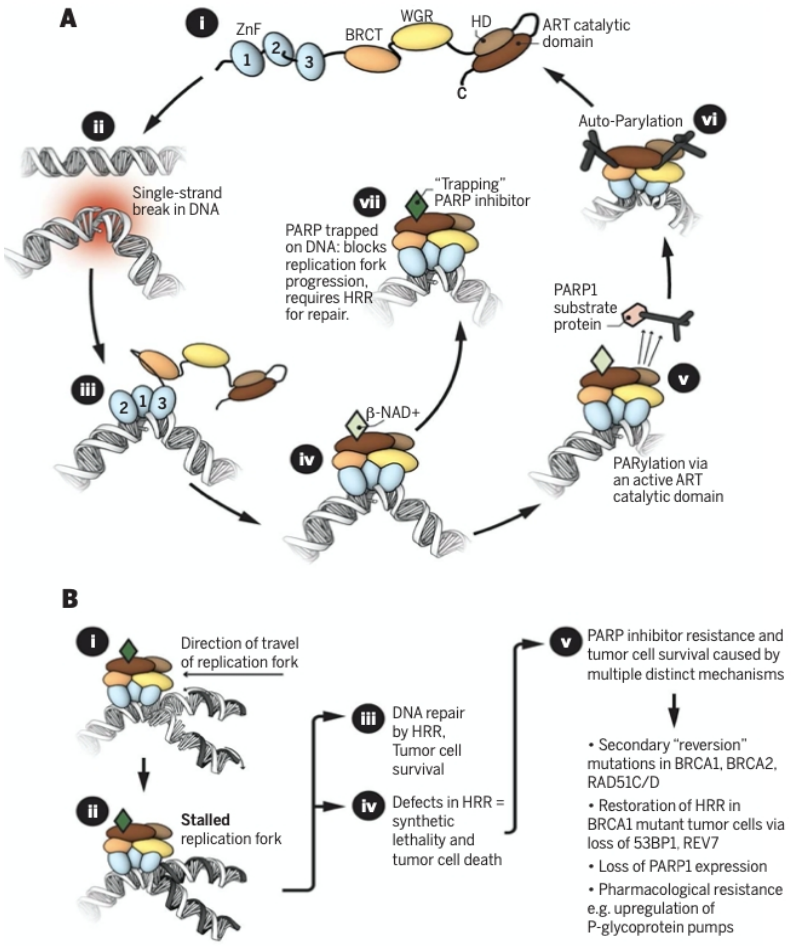


Figure 6. Schematic representation of normal PARP activity (A) and PARP inhibition (B). (A) SSB repair mechanism in normal conditions. (i) PARP in its no-DNA bound form has a conformation called "beads on a string". (ii) In the presence of a single strand break (SSB), the orientation of the double helix change, (iii) providing a bind site for PARP1, (iv) which recruits the other PARP1 proteins. (v) The PARylation of PARP 1 substrate proteins is initiated, mediating the recruitment of DNA repair effectors, chromatin remodeling and eventually DNA repair proteins. (vi) After the correction of SSB, PARP1 autoPARylation causes the release of DNA repair complex from DNA and PARP returns to its inactive state. (vii) PARP inhibitors (PARPi) usually cause PARP1 "trapping" at the site of DNA damage. (B) Synthetic lethality caused by PARP inhibition: (i) PARP is trapped at the site of DNA damage. (ii) Trapped PARP prevents the progression of the replication fork, causing the collapse of the replication fork and a DNA damage response to double strand breaks (DSB). (iii) The main process to repair DSB is Homologous Recombination Repair (HRR), with BRCA1 and BRCA2 tumor-suppression proteins. This process also involves other "BRCAness" proteins. (iv) In the absence of effective HRR, cells can use other DNA repair processes that can lead to large genomic rearrangements, leading to cell death and synthetic lethality. (v) Even when HRR is defective, PARPi resistance can occur. The schema shows the main mechanisms of PARPi resistance. (35)

Clinical evidence

Several clinical trials are being carried out with prostate cancer patients and Olaparib, a PARP inhibitor, as a treatment, in an effort to improve treatment options. An example of this effort is the TOPARP (Trial of PARP Inhibition in Prostate Cancer) phase II clinical trial (NCT01682772), divided in two phases. The first phase was called TOPARP-A, an open-label, single-group, two-stage, multi-site study where 50 patients with metastatic, castration-resistant prostate cancer (mCRPC) were recruited, regardless tumor mutational status, and treated with Olaparib until the occurrence of radiologic progression, unequivocal clinical progression, unacceptable side effects, withdraw of consent or death. Forty-nine patients were followed up and 16 showed sustained response to PARP inhibition. The analysis of tumor biopsies with a panel containing DNA-damage repair (DDR) genes showed that 16 patients (33%) had aberrations in those genes, seven of them with mutations in *BRCA2*. In the group of patients that sustained response to Olaparib, 14 cases had mutations in DNA-damage repair genes, both of germline and somatic origin (29).

In the second phase of TOPARP clinical trial, TOPARP-B, 98 patients with mCRPC carrying mutations in DNA-damage repair genes were randomized in two groups with different Olaparib treatment doses (38). In each treatment group the most altered gene was *BRCA2*, with mutations or homozygous deletions in 31% of the patients, followed by *ATM* and *CDK12*, each mutated in ~21% of the cases. Overall, 43 (46.7%) of 92 eligible patients achieved composite response, defined by radiological objective response, decrease in PSA of 50% or more, or conversion of circulating tumour cell count. The *BRCA1/2* group, comprising 32 patients, had the highest number of responses and the longest median radiographic progression-free survival, with 25 patients (83.3%) achieving composite response. In the *ATM* group, 7 of 19 (36.8%) patients achieved composite overall response, whereas in the *CDK12* group and in *PALB2* group, the composite response was 5 of 20 (25%) patients and 4 of 7 (57.1%) patients, respectively (39).

This trial confirmed the antitumoral activity of Olaparib in mCRPC with DDR gene aberrations, with greatest efficacy in *BRCA1/2* mutated tumors. The response of *ATM*-altered mCRPC patients to Olaparib was less observed, but patients with alterations in *ATM* may benefit from this treatment. However, since *ATM* is a frequently mutated gene in metastatic prostate cancer and found in both patients that respond and do not respond to treatment, detection of *ATM* alterations alone may not be sufficient to identify sensitive tumors (39).

The findings of TOPARP clinical trial lead to the design of the PROfound phase III clinical trial (NCT02987543), a randomized open-label study evaluating efficacy and safety of Olaparib versus enzalutamide or abiraterone in mCRPC patients that have failed treatment with a hormonal agent (40). Patients carrying mutations in one of 15 HRR genes (*BRCA1, BRCA2, ATM, BRIP1, BARD1, CDK12, CHEK1, CHEK2, FANCL, PALB2, PPP2R2A, RAD51B, RAD51C, RAD51D,* and *RAD54L*) were randomized in treatment with Olaparib or enzalutamide or abiterone (AR-targeted therapy). The first results were presented at the ESMO 2019 congress and showed improvement of radiographic progression-free survival (rPFS) and overall survival (OS), well tolerated by patients (40). The original article, published in April 2020, supported previous observations that patients harboring *BRCA1/2* mutations seem to be more sensitive to treatment than patients harboring other HRR gene mutations (40,41).

On May 19, 2020, Olaparib (Lynparza, AstraZeneca AB) was approved by the FDA for the treatment of mCRPC patients with deleterious or suspected deleterious germline or somatic homologous recombination repair genes, as a result of the PROfound clinical trial (42) and on September 17, 2020, the European Medicines Agency (EMA) recommended Olaparib (Lynparza, AstraZeneca AB) treatment in Europe for mCRPC patients that harbors germline or somatic mutation in *BRCA1* or *BRCA2* genes as a second line treatment (43).

Other studies involving another PARPi drug, Rucaparib, have been conducted or are ongoing. The TRITON2 study (NCT02952534) is the most important one, in which patients with mCRPC and alterations in 15 DNA damage repair genes (*BRCA1, BRCA2, ATM, BARD1, RIP1, CDK12, CHEK2, FANCA, NBN, PALB2, RAD51, RAD51B, RAD51C, RAD51D* and *RAD54L*) were treated with Rucaparib. The study concluded that, as in the TOPARP clinical trial, patients harboring *ATM* alterations showed no response to PARP inhibition, suggesting that alterations in this gene only are not sufficient to lead to synthetic lethality (44,45). The strongest response to PARP inhibition was seen in patients harboring *BRCA1/2* alterations, leading to a recent FDA approval of Rucaparib to patients with mCRPC harboring somatic or germline deleterious *BRCA* mutations (46). The TRITON study is now on phase III (TRITON3).

Liquid biopsy

General considerations

The concept of “liquid biopsies” has been used for years. The presence of fragments of DNA in the blood was first reported by Mandel and Métais and the origin and characteristics of these cell-free DNA (cfDNA) fragments were studied for the subsequent years. Raised levels of cfDNA were first reported in 1977, in the serum of cancer patients. In 1997, fetal cfDNA was identified in the serum of pregnant women, leading to various clinical applications in prenatal medicine, such as sex determination, identification of monogenic disorders and noninvasive prenatal testing (NIPT) for aneuploidies, first introduced in 2007 (47).

In the field of oncology, the identification that a portion of the cfDNA was originated from cancer cells was made in 1989, and in 1991 mutations in *TP53* gene were found in urine sediments of patients with invasive bladder cancer, leading to the use of liquid biopsies in the oncology field. In 1994, a correlation between the mutations found in plasma was made with the ones found in tumors, confirming the origin of a part of cfDNA. The term circulating-tumor DNA (ctDNA) was, then, incorporated as a portion of cfDNA, with mutations found in ctDNA being highly specific tumor biomarkers. In the following years, ctDNA was explored as prognostic or predictive biomarker, as well as in cancer detection, giving rise to the term liquid biopsy, which refers to the analysis of tumor material (nucleic acids, for example) obtained in a minimally or non-invasive form, such as blood or other body fluids collection (47).

The mechanism of cfDNA release by non-cancerous cells remains unclear. In tumor cells, the ctDNA may derive from apoptotic or necrotic cells and/or active secretion from extracellular vesicles and circulating tumor cells (CTCs) (Figure 7). Theoretically, all alterations found in cancer cells can be detected in ctDNA, in any genomic region, including coding and noncoding regions, and either microsatellite *loci*, loss of heterozygosity (LOH), mutations, polymorphisms, methylation and copy number variations (CNV) can be assessed (48,49).

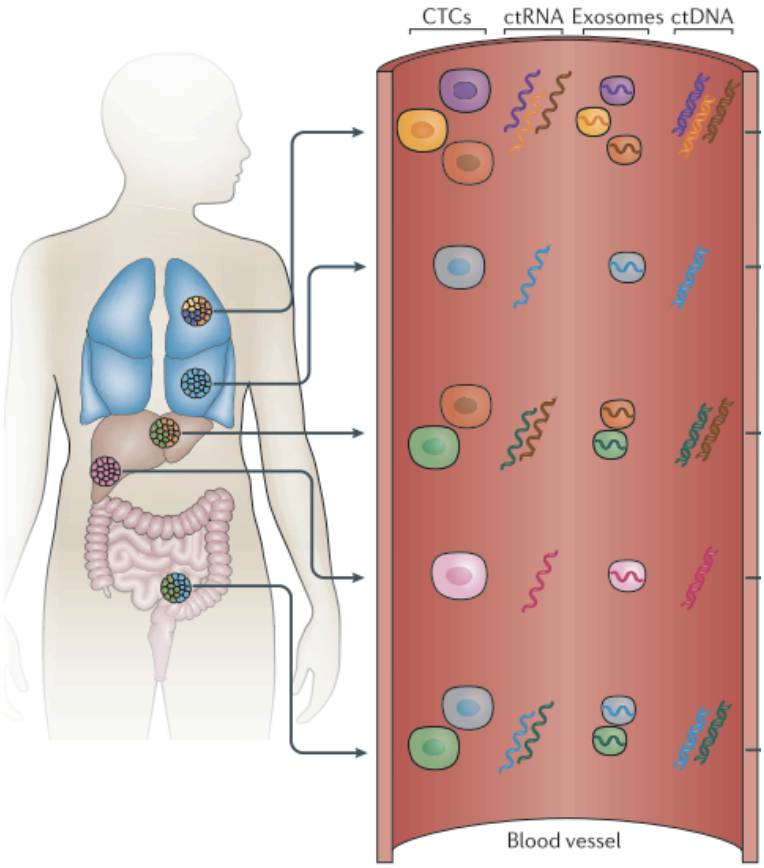


Figure 7. Representation of liquid biopsy components released by primary tumor and metastasis in blood stream. Apoptotic cells, necrotic cells or cell secretion releases tumor components to the blood stream. These components are Circulating Tumor Cells (CTCs), ctDNA, exosomes and ctRNA. All these components represent tumor and metastasis molecular profiles and can be analyzed through liquid biopsies (48)

The levels of ctDNA in cfDNA are dependent of several factors, such as tumor volume, location and cellular turnover. Since the proportion of ctDNA is small (in the order of 0.1% to over 10% depending on disease burden, stage, cellular turnover and treatment response), the use of ctDNA as a biomarker is challenging. With the advance of molecular technologies, now we can detect more easily this small part of cfDNA, first with PCR based techniques, such as digital droplet PCR (ddPCR) and, more recently, with next generation sequencing (NGS), and use ctDNA as a biomarker for cancer diagnosis, prognosis and therapeutic monitoring. The size of a ctDNA fragment is dependent of its biology and has a distribution of 130-170bp, with the peak at 160bp. This size distribution suggests that a high proportion of ctDNA may be originated from the apoptotic process (48,49).

Clinical applications of ctDNA

In the clinical setting, liquid biopsies can be applied in multiple fields, from cancer detection until treatment monitoring and drug resistance. These applications can be used throughout the course of the disease. The major applications are listed below and schematized in Figure 8 (48,49):

Genotyping for somatic alterations: liquid biopsy virtually represents the landscape of alterations of tumor tissue, that is, somatic alterations found in primary and metastatic tumor. Although specificity can approach 100%, the sensitivity is lower and depends of the type of alterations.

Detecting minimal residual disease (MRD): it consists in detecting the presence of a molecular metastatic status in a phenotypically local disease. The current limitation of this approach is the need of highly sensitive assays, which has been the subject of several studies. Proof of concept studies demonstrate that ctDNA can be used to monitor MRD after surgery or other curative treatment in colorectal, breast, lung and other cancer types.

Monitoring treatment response: the short half-life of ctDNA, its minimally invasive sampling and the possibility of multiple sample collection allows disease monitoring during treatment, measuring the level of ctDNA through, previously identified, mutations.

Understanding acquired drug resistance: ctDNA can be used to monitor clonal progression and to identify mechanisms of resistance due to drug exposure. In cancers where know mutations can cause resistance, it is possible to monitor these mutations/subclones, predict drug resistance and guide the next steps of treatment.

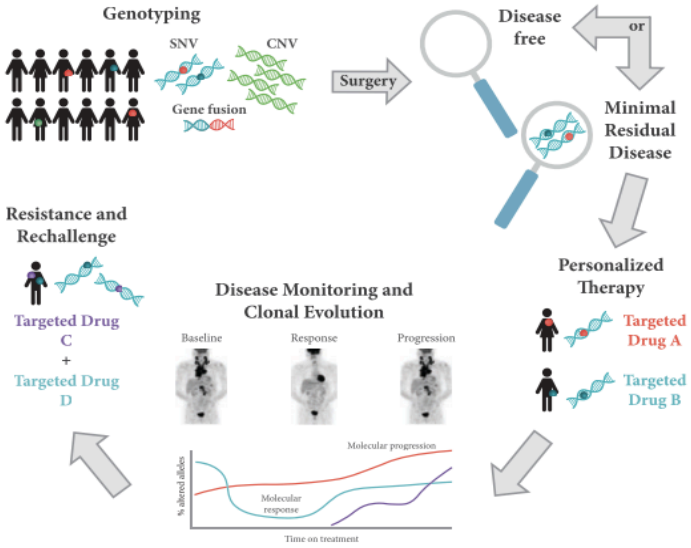


Figure 8. Clinical applications of ctDNA analysis (49)

Landscape of liquid biopsy applications in prostate cancer

In the diagnostic field of prostate cancer, different forms of liquid biopsies are already being used, either evaluating CTCs, nucleic acids (DNA or RNA) or extracellular vesicles, which, combined with PSA levels, Gleason score, histopathologic analysis and DRE, guide clinical management.

Focusing on the field of prognostic tests, the Onco*type* DX AR-V7 Nucleus Detect® test (Onco*type* IQ, powered by Epic Sciences) was design to mCRPC patients. The test detects the AR-V7 mRNA isoform in CTCs nucleus and can predict if a mCRPC patient has developed resistance to abiraterone and enzalutamide treatment, guiding the decision to switch treatment to chemotherapy if the resistance was confirmed (50,51).

The majority of prognostic tests, such as Decipher Prostate® (Decipher Biosciences, Inc) (52), Onco*type* DX® Genomic Prostate Score (Onco*type* IQ) (53), Prolaris® (Myriad Genetics, Inc) (54) and ProMark® Proteomic Prognostic Test (Metamark Genetics, Inc) (55), are processed in tumor FFPE samples, originating from biopsy or RP, all recommended by NCCN guidelines. These tests are recommended to men with low or favorable intermediate risk, to help clinicians in treatment decision on the time of diagnosis, predicting disease aggressiveness (16).

In the research field, regarding ctDNA, Goodall and colleagues (2017) analyzed cfDNA of TOPARP-A clinical trial samples to evaluate if it could be used as a guide to PARPi treatment (56). They found that cfDNA somatic mutation allele frequencies decreased in a sustained way in patients responding to Olaparib treatment and this sustained decreased was not observed in nonresponding patients. The same was observed in patients that had somatic mutations in HRR genes, with five of the six carrier patients showing decreased allele frequencies during treatment. The same study found that ctDNA monitoring can reveal emerging *de novo* mutations throughout therapy, likely of tumor subclones, as a result of therapeutic selective pressure, which can also guide treatment as drug resistance markers.

Next Generation Sequencing technology

Since the remarkable advent of Sanger chain terminator sequencing, in 1977, the world gained the ability to sequence DNA in a reliable, reproducible and accessible manner (57). In 1987, Applied Biosystems introduced the first automated capillary electrophoresis (CE) equipment, the AB370, and, in 1990, this CE and AB3730xl, launched in 1998, became the equipment used in the Human Genome Project, led by NIH and Celera (58,59).

Sequencing technologies evolving through time, passing through pyrosequencing, a luminescent technique that measures pyrophosphate synthesis in an indirect manner at each dNTP incorporation, until today's NGS technologies. Nowadays, sequencing technology allows multiple target sequencing and multiplexing samples at the same run. Library preparation chemistries and equipment evolved in a manner that allows high throughput, more sensible, accurate and faster sequencing (59,60).

Sequencing by synthesis (SBS) chemistry

Illumina's sequencers technology is similar to capillary electrophoresis (CE) sequencing, or Sanger sequencing. DNA polymerase enzyme catalyzes the incorporation of fluorescently labeled deoxyribonucleotide triphosphates (dNTPs) into a DNA template, creating a new strand. dNTPs are made available in sequential cycles and, at each incorporation, the fluorophore is excited and the nucleotide is identified by a CCD camera coupled in the sequencer. Unlike CE sequencing, the dNTP incorporated in SBS is not irreversible-terminal, that is, the synthesis of new DNA strand continues even after the incorporation. Another major difference is that, instead of sequencing single DNA fragments, NGS extends this process across millions of fragments at the same time, and that is why it is also called massively parellel sequencing (59,61).

NGS run parameters

The total number of samples per run is dependent of the flowcell size, panel size, percentage of PCR duplicates and desired depth of coverage per sample, which is variable depending on applicability/variants origin. Several factors influence the depth of coverage, including sequence platform, sequence complexity of targeted region, library preparation chemistry and types of variants being evaluated. In general terms, germline variants require lower depth of coverage than somatic variants. For germline tests, a minimum of 30x coverage with balanced reads (similar quantity of forward and reverse reads) is usually sufficient (62), while for somatic tests, an average coverage of at least 1000x may be necessary to identify variants at lower frequencies.

According to Illumina NextSeq 550 specifications, a run in a mid-output flowcell with 150 cycles (2 x 75bp) has the duration of approximately 15 hours and the flowcell output is within 16.25-19.5 Gb and up to 260 million reads (63).

NGS run quality control

The first parameter calculated by the NGS instrument is cluster density and the percentage of clusters passing filter. Illumina's sequencers have a cluster quality filter that "chooses" the clusters suitable to collect data. There is not an exact number for raw cluster density, but, according to Illumina's PhiX run, the passing filter cluster density for NextSeq instruments is between 129 and 165K/mm^3 (63).

After the run ends, the instrument calculates other parameters, such as real flowcell output and Phred score and, if possible, performes sample demultiplexing. The Phred score for a mid-output flowcell in a 2 x 75bp run is higher than Q30 in more than 80% of bases (63). When demultiplexing is performed by the instrument, it is possible to infer immediately the homogeneity of the run in terms of sample representativeness in the flowcell.

PCR duplicates

PCR duplicates should be excluded from bioinformatic analysis in order to minimize the call for variants that are, in fact, PCR errors. Reads that have the same start and stop coordinates are likely to be from the same original DNA molecule, so they are considered to be non-independent observations. All of them, except for one pair, are marked and then ignored in the following analysis steps (64). There isn't an exact number recommended for PCR duplicates, as those need to be as lowest as possible. Nevertheless, it is expected that poor-quality DNA, as that obtained from FFPE samples, has higher percentage of PCR duplicates.

Depth of coverage

Because bioinformatic tools are dependent on adequate depth of coverage for sensitive and specific variant detection, guidelines recommend a minimum depth of coverage per target. The Association for Molecular Pathology (AMP) and College of American Pathologists (CAP) strongly recommend, for clinical oncology NGS panels, a minimum of 250x of coverage per amplicon or target, ensuring the detection of variants with 5% of allele frequency with a probability lower than 0.5% of false negative (65). The International Quality Network for Pathology (IQN Path) recommends a minimum of 300-500X per target (66).

Variant interpretation and classification

Variant interpretation and classification follows a series of standards criteria, described in guidelines and recommended by a joint consensus of two or more of the following: American College of Medical Genetics and Genomics (ACMG), Association for Molecular Pathology (AMP), American Society of Clinical Oncology (ASCO), College of American Pathologists (CAP) and European Society for Medical Oncology (ESMO), among others. Germline variants are classified by their pathogenicity, while somatic variants are more complex, classified by their clinical and/or experimental significance in tumorigenesis and the consequences regardless cancer diagnosis, prognosis and/or treatment. These guidelines also recommend databases relevant to interpretation of germline and somatic variants and best practices in variant detection (67–69).

Somatic variant interpretation

The significance in tumorigenesis of variants found in tumor sequencing depends on the type of genetic alteration, location of the variant and the normal function of the protein. Tumor profile can provide information to guide clinical management of cancer patients regardless diagnostic or prognostic information, identify a potential treatment regimen or targeted therapy and determine eligibility for FDA approved medication, medication available as off-label treatment for that specific alteration in a non-approved tumor type or targeted therapy available in clinical trials (68).

In 2015, the AMP, ACMG, ASCO and CAP members formed a workgroup to define standards and guidelines to unify the classification and report of somatic variants (68). Figure 9 shows the schema of AMP/ASCO/CAP variant classification in tiers.

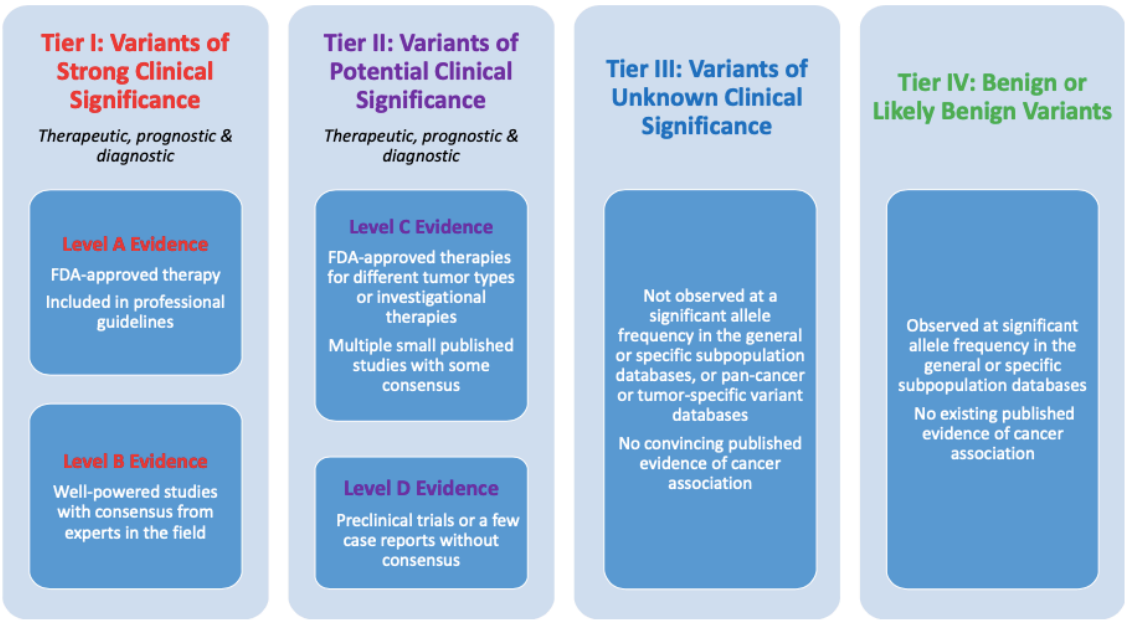


Figure 9. AMP/ASCO/CAP Levels of Evidence

OncoKB database also developed a classification system based on level of evidence that has an interchange with AMP/ASCO/CAP levels of evidence. The OncoKB levels of evidence are divided in 4 levels of predictive drug response and 2 levels of drug resistance (Figure 10) (70).

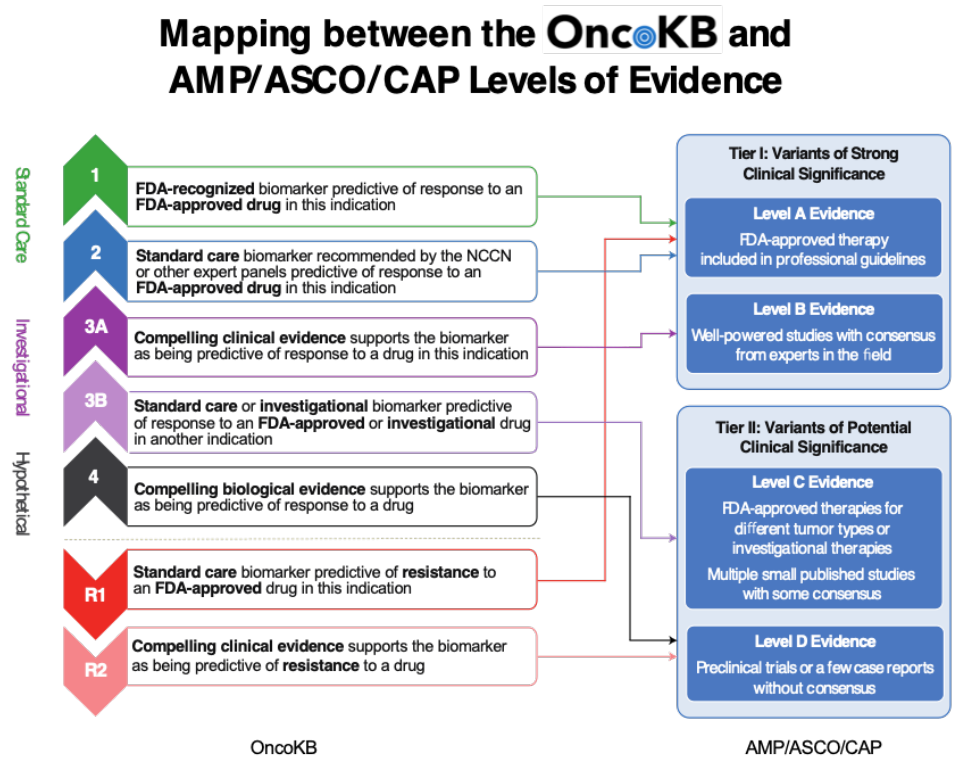


Figure 10. OncoKB Levels of Evidence and their mapping with AMP/ASCO/CAP Levels of Evidence

The ESMO Scale of Clinical Actionability for molecular Targets (ESCAT)

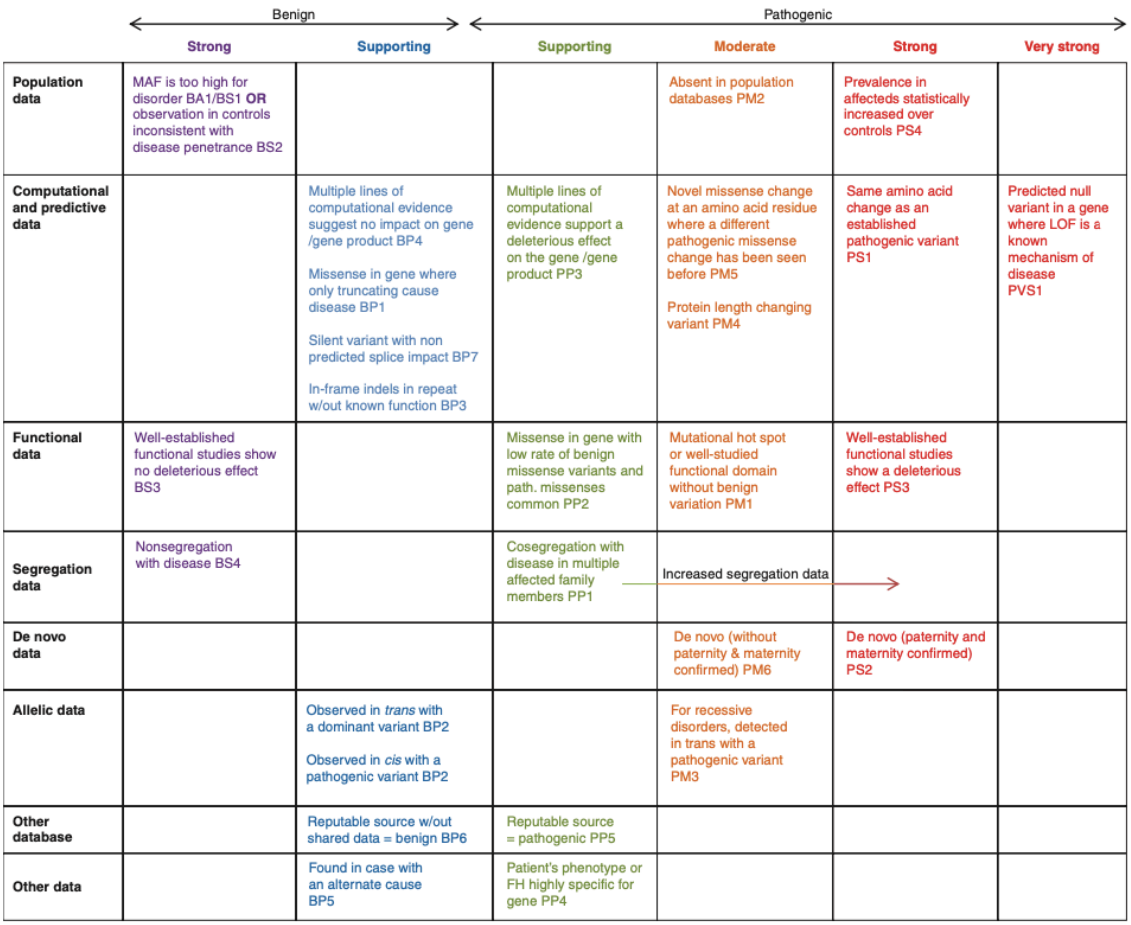
In 2018, the European Society for Molecular Oncology (ESMO), in order to unify and harmonize the classification of targets for precision medicine, proposes a tier-based classification system. The ESCAT (ESMO Scale of Clinical Actionability for Molecular Targets) is a five-tier based system, with one extra tier for variants with lack of evidence for actionability and is based on previous classification schemas. The intent of ESCAT is to introduce a classification system that can be used globally and to make this possible, the ESCAT system does not take into account the American Food and Drug Administration (FDA) or another country-specific regulatory agency. The classification is based on scientific and clinical evidence, taking into account the type of clinical trial, the type of evidence generated by clinical trials and if the evidence is supported by one or more clinical trials (69). Figure 11 summarizes the ESCAT system.

	ESCAT evidence tier	Required level of evidence	Clinical value class	Clinical implication
Ready for routine use	I: Alteration-drug match is associated with improved outcome in clinical trials	I-A: prospective, randomised clinical trials show the alteration-drug match in a specific tumour type results in a clinically meaningful improvement of a survival end point I-B: prospective, non-randomised clinical trials show that the alteration-drug match in a specific tumour type, results in clinically meaningful benefit as defined by ESMO MCBS 1.1 I-C: clinical trials across tumour types or basket clinical trials show clinical benefit associated with the alteration-drug match, with similar benefit observed across tumour types	Drug administered to patients with the specific molecular alteration has led to improved clinical outcome in prospective clinical trial(s)	Access to the treatment should be considered standard of care
Investigational	II: alteration-drug match is associated with antitumour activity, but magnitude of benefit is unknown	II-A: retrospective studies show patients with the specific alteration in a specific tumour type experience clinically meaningful benefit with matched drug compared with alteration-negative patients II-B: prospective clinical trial(s) show the alteration-drug match in a specific tumour type results in increased responsiveness when treated with a matched drug, however, no data currently available on survival end points	Drug administered to a molecularly defined patient population is likely to result in clinical benefit in a given tumour type, but additional data are needed	Treatment to be considered 'preferable' in the context of evidence collection either as a prospective registry or as a prospective clinical trial
Hypothetical target	III: alteration-drug match suspected to improve outcome based on clinical trial data in other tumour type(s) or with similar molecular alteration	III-A: clinical benefit demonstrated in patients with the specific alteration (as tiers I and II above) but in a different tumour type. Limited/absence of clinical evidence available for the patient-specific cancer type or broadly across cancer types III-B: an alteration that has a similar predicted functional impact as an already studied tier I abnormality in the same gene or pathway, but does not have associated supportive clinical data	Drug previously shown to benefit the molecularly defined subset in another tumour type (or with a different mutation in the same gene), efficacy therefore is anticipated for but not proved	Clinical trials to be discussed with patients
	IV: pre-clinical evidence of actionability	IV-A: evidence that the alteration or a functionally similar alteration influences drug sensitivity in preclinical *in vitro* or *in vivo* models IV-B: actionability predicted *in silico*	Actionability is predicted based on preclinical studies, no conclusive clinical data available	Treatment should 'only be considered' in the context of early clinical trials. Lack of clinical data should be stressed to patients
Combination development	V: alteration-drug match is associated with objective response, but without clinically meaningful benefit	Prospective studies show that targeted therapy is associated with objective responses, but this does not lead to improved outcome	Drug is active but does not prolong PFS or OS, probably in part due to mechanisms of adaptation	Clinical trials assessing drug combination strategies could be considered
	X: lack of evidence for actionability	No evidence that the genomic alteration is therapeutically actionable	There is no evidence, clinical or preclinical, that a genomic alteration is a potential therapeutic target	The finding should not be taken into account for clinical decision

Figure 11. The ESCAT - ESMO Scale of Clinical Actionability for molecular Targets

Germline variant interpretation

Guideline recommendations were elaborated by a joint consensus of American College of Medical Genetics and Genomics (ACMG) and Association for Molecular Pathology (AMP) in 2015, in its latest version (71). Figure 12 shows a schema of criteria for germline variant classification.



	Benign		Pathogenic			
	Strong	**Supporting**	**Supporting**	**Moderate**	**Strong**	**Very strong**
Population data	MAF is too high for disorder BA1/BS1 **OR** observation in controls inconsistent with disease penetrance BS2			Absent in population databases PM2	Prevalence in affecteds statistically increased over controls PS4	
Computational and predictive data		Multiple lines of computational evidence suggest no impact on gene /gene product BP4 Missense in gene where only truncating cause disease BP1 Silent variant with non predicted splice impact BP7 In-frame indels in repeat w/out known function BP3	Multiple lines of computational evidence support a deleterious effect on the gene /gene product PP3	Novel missense change at an amino acid residue where a different pathogenic missense change has been seen before PM5 Protein length changing variant PM4	Same amino acid change as an established pathogenic variant PS1	Predicted null variant in a gene where LOF is a known mechanism of disease PVS1
Functional data	Well-established functional studies show no deleterious effect BS3		Missense in gene with low rate of benign missense variants and path. missenses common PP2	Mutational hot spot or well-studied functional domain without benign variation PM1	Well-established functional studies show a deleterious effect PS3	
Segregation data	Nonsegregation with disease BS4		Cosegregation with disease in multiple affected family members PP1	Increased segregation data →		
De novo data				De novo (without paternity & maternity confirmed) PM6	De novo (paternity and maternity confirmed) PS2	
Allelic data		Observed in *trans* with a dominant variant BP2 Observed in *cis* with a pathogenic variant BP2		For recessive disorders, detected in trans with a pathogenic variant PM3		
Other database		Reputable source w/out shared data = benign BP6	Reputable source = pathogenic PP5			
Other data		Found in case with an alternate cause BP5	Patient's phenotype or FH highly specific for gene PP4			

Figure 12. Evidence framework of ACMG classification, showing the criteria and the strength of each one

AIMS AND OBJECTIVES

The main aim of this study is to evaluate the percentage of patients referenced to the Portuguese Oncology Institute of Porto (IPO Porto), between September 2019 and July 2020, with metastatic prostate cancer that have somatic or germline mutations in known DNA repair genes, detectable in formalin-fixed paraffin-embedded (FFPE) tumor tissue through Next Generation Sequencing (NGS).

Additionally, while collecting both peripheral blood leucocytes from the same patients, we aim to assess whether the identified mutations have germline or somatic origin.

Objectives

1. Establish an NGS analysis pipeline to test a panel of 30 DNA damage repair (DDR) genes in FFPE tissue.
2. Identify potential clinically actionable variants in tumor samples of metastatic prostate cancer patients.
3. Clarify the germline/somatic nature of the clinically actionable variants.

MATERIAL AND METHODS

Patient recruitment and sample collection

Metastatic PrCa patients' samples

Patients with metastatic prostate cancer were recruited at IPO Porto between September 2019 and August 2020, after referral for genetic counselling and testing by their treating physician. Matched peripheral blood leucocytes, plasma and tumor samples were gathered, whenever possible, upon genetic counselling consultation and informed consent obtained from all patients.

Peripheral blood was collected in a Streck® tube, suitable to appropriate stabilization of cell-free nucleic acids. Tumor samples were obtained from archived FFPE tissue.

Control samples

Control samples were chosen considering the presence of known variants, including nonsense, frameshift or splicing mutations, and large insertions/duplications (Table 3). Ten samples had previously been analyzed at the Department of Genetics of IPO Porto, eight from FFPE tumors and two from peripheral blood. Overall, the samples chosen for this first run were of various tumor types, with different DNA quality and quantity, in order no mimic the real scenario of a diagnostic routine (Table 3).

Table 3. Control samples information

Sample number	Sample type	[ng/ul]	Gene	Trancript	Pathogenic Variant
S1	FFPE	22.4	*BRCA2*	NM_000059.3	c.6351_6377del p.V2118_C2126del
S2	FFPE	28	*PALB2*	NM_024675.3	c.2401G>A p.D801N
S3	FFPE	37	*BRCA2*	NM_000059.3	c.156_157insAlu
S4	FFPE	39	*RAD51C*	NM_058216.1	c.709C>T p.R237X
S5	FFPE	40	*MSH6*	NM_000179.2	c.3848_3862del p.I1283_Y1287del c.4001+1G>T
S6	FFPE	37	*ARID1A*	NM_006015.4	c.4003C>T c.6704_6729del
S7	FFPE	48	*PTEN*	NM_000314.4	c.15_16insCTTC
S8	Blood	42	*PTEN*	NM_000314.4	c.493-1884_634+1938dup
S9	FFPE	37	*RECQL4* *TP53*	NM_004260.3 NM_001126112.2	c.2297delC p.P766Rfs c.673-1G>A
S10	Blood	64	*MSH2*	NM_000251.2	c.942+3A>T

DNA extraction and storage

Tumor (tDNA) and germline (gDNA) DNA was extracted using standard methods currently used at the Department of Genetics of IPO Porto. Plasma fractions were obtained from peripheral blood (for future studies) and stored at -80ºC for subsequent extraction of cell-free DNA (cfDNA) using the protocol established at the Department of Genetics of IPO Porto for diagnostic purposes. All DNA samples were stored at -20ºC until use.

Blood samples processing (gDNA and ctDNA)

Up to 4 hours after collection, blood samples were centrifuged for 10 minutes at room temperature and 1,500rcf. Immediately after centrifugation, 1.5mL aliquots of the upper fraction were transferred to Eppendorf tubes, without touching the buffy coat, and centrifuged for 10 minutes at 6,000rcf. Plasma was, then, transferred to new Eppendorf tubes in 1.1mL aliquots and stored at -80ºC until cfDNA extraction. The buffy coat was transferred to another Eppendorf tube and the pellet stored at -80ºC until gDNA extraction.

For gDNA extraction, pellet was resuspended in 1mL of phosphate-buffered saline (PBS) and automatic DNA extraction was performed by the MagNA Pure LC DNA Isolation Kit and the DNA LV (Large Volume) Cells protocol in MagNA Pure (Roche).

Tumor samples processing and tDNA extraction

Slides of 5μm thickness were prepared, followed by hematoxylin-eosin (HE) staining in one slide, at the Pathology Department of IPO Porto. Tumor area was delimited by an experienced pathologist and the slides were sent to the Department of Genetics to proceed to deparaffinization and tDNA extraction.

The deparaffinization process consists in placing the slides in a staining jar with xylene for 5 minutes, after which slides are transferred to a new staining jar with absolute ethanol for additional 5 minutes.

Manual microdissection of the tumor was made using a scalpel and the fragments placed in an Eppendorf tube. DNA extraction was performed using the HighPure FFPE DNA Isolation Kit (Roche), following manufacturer protocol with minimal modifications in the tissue digestion step – achieved overnight at 56ºC, with the addition of proteinase K and DNA tissue lysis buffer. All tDNA samples were eluted in 50μl and quantified by Qubit dsDNA HS Assay (ThermoFisher).

Customized NGS gene panel

The customized gene panel (SureSelect XT HS - Agilent Technologies) was designed specifically for this purpose and includes the DNA repair genes described in clinical trials phase II and III (28,39,40). The panel was synthesized by Agilent Technologies, with SureSelect XT chemistry. All genes covered by customized panel are listed in Table 4, in red the genes of the PROfound clinical trial and in blue the DNA repair genes included in the other clinical trials phase II. The *PPP2R2A* gene, described in the PROfound clinical trial, was not included in the analysis because it did not show predictive value for Olaparib treatment, and therefore was excluded of the NCCN Guidelines with indication for PARPi treatment (16).

Table 4. Customized panel gene list

Customized panel gene list		
ATM NM_000051.3	***ERCC3*** NM_000122.1	***MSH6*** NM_000179.2
ATR NM_001184.3	***FAM175A*** NM_139076.2	***NBN*** NM_002485.4
BAP1 NM_004656.3	***FANCA*** NM_000135.2	***PALB2*** NM_024675.3
BARD1 NM_000465.2	***FANCL*** NM_018062.3	***PMS2*** NM_000535.5
BRCA1 NM_007294.3	***GEN1*** NM_001130009.3	***RAD51*** NM_002875.4
BRCA2 NM_000059.3	***HDAC2*** NM_001527.4	***RAD51B*** NM_133509.3
BRIP1 NM_032043.2	***MLH1*** NM_000249.3	***RAD51C*** NM_058216.2
CDK12 NM_016507.2	***MLH3*** NM_001040108.2	***RAD51D*** NM_133629.2
CHEK1 NM_001274.5	***MRE11A*** NM_005591.3	***RAD54L*** NM_001142548.1
CHEK2 NM_007194.3	***MSH2*** NM_000251.2	***XRCC2*** NM_005431.1

Next Generation Sequencing

The SureSelect library preparation protocol was performed under some modifications in order to improve the quality of the DNA library from FFPE samples. These modifications included one additional step of hybridization and, in consequence, alterations in the number of cycles of the hybridization process.

In summary, the SureSelect library preparation involved three major steps: 1) sample preparation, 2) hybridization and capture, and 3) post-capture sample processing for multiplexed sequencing. DNA and library quality control were assessed using Agilent 2200 TapeStation System (Agilent Technologies, Inc), in multiple points of library preparation, described in detail below. The NGS run was performed in a NextSeq 550 (Illumina, Inc) sequencer, using a mid-output flowcell.

1. Sample preparation

Involves all the preparation of each DNA sample to initiate the capture process. First, the quality and quantity of each genomic DNA sample were assessed, in order to adjust each sample input according to its quality. Then, a chemical fragmentation was performed, followed by repair and dA-tail of the DNA ends, preparing DNA fragments to receive molecular-barcoded adapters. Next, molecular-barcoded adapters were ligated to DNA ends, preparing them to be able to be amplified in the surface of the flowcell. Molecular-barcoded adapters identified each sample as unique. To finish this step, the adapter-ligated library was purified and amplified.

2. Hybridization and capture

In the hybridization process, the probes of customized panel annealed with target DNA fragments, previously adapter-ligated, and, to ensure that only target fragments remain in the process and with sufficient quantity, an amplification step was performed. Targeted fragments were size selected through streptavidin-coated magnetic beads. The concentration ratio between sample and streptavidin-coated magnetic beads defined the size of selected fragment, that is, which fragments stayed bound to the bead.

To increase FFPE library quality, this step was performed twice, reducing the number of cycles to 8 in the amplification step.

3. Post-capture sample processing for multiplexed sequencing

In this final step, the libraries were amplified, purified and its quality and quantity were assessed. Until this point, all samples were treated individually. A pool containing all samples equally represented was prepared and diluted according to Illumina's NextSeq Denature and Dilute Libraries Guide standard normalization method and loaded to the equipment cartridge.

Quality control during library preparation

For a successful NGS run, several parameters have to be obeyed regarding DNA quality and quantity. As FFPE DNA is degraded due to the tissue fixation and embedding process, more strict protocols and quality checkpoints are needed. All parameters

presented below were determined following the manufacturer recommendations (Agilent Technologies, Inc).

DNA quality

First, the quality of all DNA samples was assessed through Agilent's Tapestation Genomic ScreenTape assay, in the TapeStation equipment. This assay analyzes DNA fragmentation and outputs the DNA Integrity Number (DIN) score, used to correct the input of each sample. The samples suitable for library preparation were selected based on the DIN score, DNA quantification and available DNA volume. These parameters are demonstrated in Table 5.

Table 5. Recommended DNA input for FFPE samples based on DIN score

FFPE samples library input according to DNA quality		
DIN>8	**DIN 3-8**	**DIN<3**
10ng to 200ng DNA, quantified by Qubit Assay	Use at least 15 ng for more intact samples and at least 40ng for less intact samples. Use the maximum amount of DNA available, up to 200ng, for all samples. Quantify by Qubit Assay	Use at least 50 ng for more intact samples and at least 100ng for less intact samples. Use the maximum amount of DNA available, up to 200ng, for all samples. Quantify by Qubit Assay

Library fragmentation

The second quality checkpoint assessed library fragmentation/ligation of adapters and quantity and it was measured via a TapeStation run, using a D1000 ScreenTape assay. Libraries suitable to continue the protocol showed an electropherogram with a peak of fragment size positioned between 300 to 400bp, for high quality DNA, and approximately 200 to 400bp for FFPE DNA (Figure 13). The concentration of the library was determined by integrating under the peak, which is given by the TapeStation software.

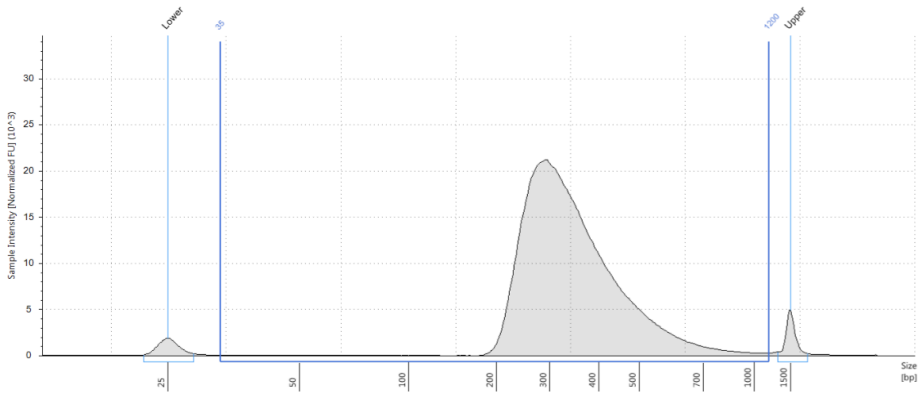


Figure 13. Typical FFPE fragmented library electropherogram using the D1000 ScreenTape assay

Sequencing library

The last quality checkpoint assessed quality and quantity of the final library and it was performed after the hybridization step using the High Sensitivity D1000 ScreenTape assay. Libraries suitable for sequencing showed an electropherogram with a peak of fragment size positioned between 200 to 400bp (Figure 14). The concentration of the library was determined by integrating under the peak, given by the TapeStation software.

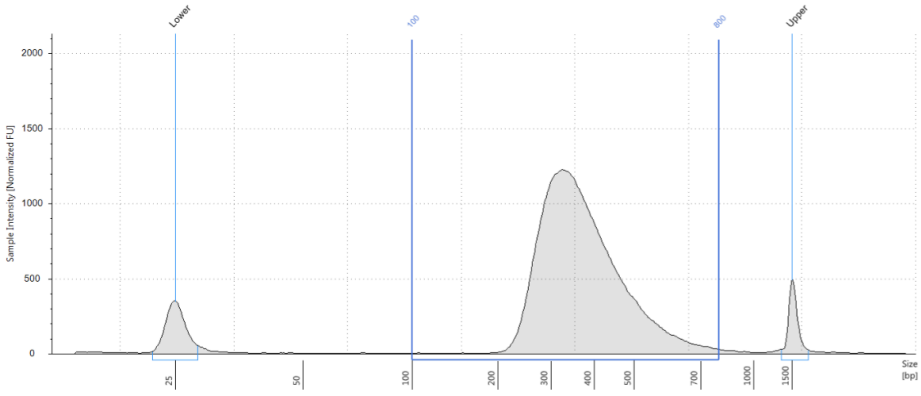


Figure 14. Typical FFPE post-capture library electropherogram using the High Sensitivity D1000 ScreenTape

Next Generation Sequencing

Control samples run

Since a customized panel was used, a pilot run with 10 control samples was performed in order to validate the design, determine probe design efficiency and robustness, and define the bioinformatic pipeline for data analysis. The run was also

necessary to validate library preparation protocol modifications, made to increase the performance of poor quality FFPE samples.

The NGS run was performed in a NextSeq 550 (Illumina, Inc) sequencer, using a mid-output flowcell with read-length of 2 x 75bp in paired-end mode.

Prostate tumor samples run

After the pilot run with control samples, one run with tumor samples from metastatic prostate patients was performed. Library preparation underwent modifications (two hybridization steps, decrease of PCR cycles in hybridization step), in order to minimize the percentage of duplicates, and run parameters were the same used in the pilot run.

Bioinformatic analysis

The bioinformatic analysis was performed using three different pipelines/software for somatic variants, namely, NextGENe version 2.4.2 (SoftGenetics), Basespace DRAGEN version 3.4.5 (Illumina, Inc) and Agilent SureCall version 4.1.1.5 (Agilent Technologies). NextGENe settings were the same as currently used at the Department of Genetics of IPO Porto, while for Basespace DRAGEN and Agilent SureCall the software default settings were used ("Somatic Pipeline" and "SureSelect Method TL", respectively). All samples from all analyses were aligned against UCSC hg19 Human Reference Genome.

The Basespace DRAGEN output was a *.vcf* file, whose was subsequently annotated by VarAFT (Genetics and Bioinformatics team at INSERM UMR 1251 in the Medical School of la Timone Aix-Marseille University in Marseille, France), using software default settings (72).

Coverage assessment per base

To assess the average depth of coverage *per base* of the 14 genes described in PROfound clinical trial (40), a workflow was developed by the group's bioinformatician (Figure 15). Briefly, after the alignment step, the duplicate reads previously marked were removed from the *.bam* files, the coverage *per base* was calculated with the *bedtools* function "coverage" and, finally, the coverage metrics were calculated by an *in house* python script.

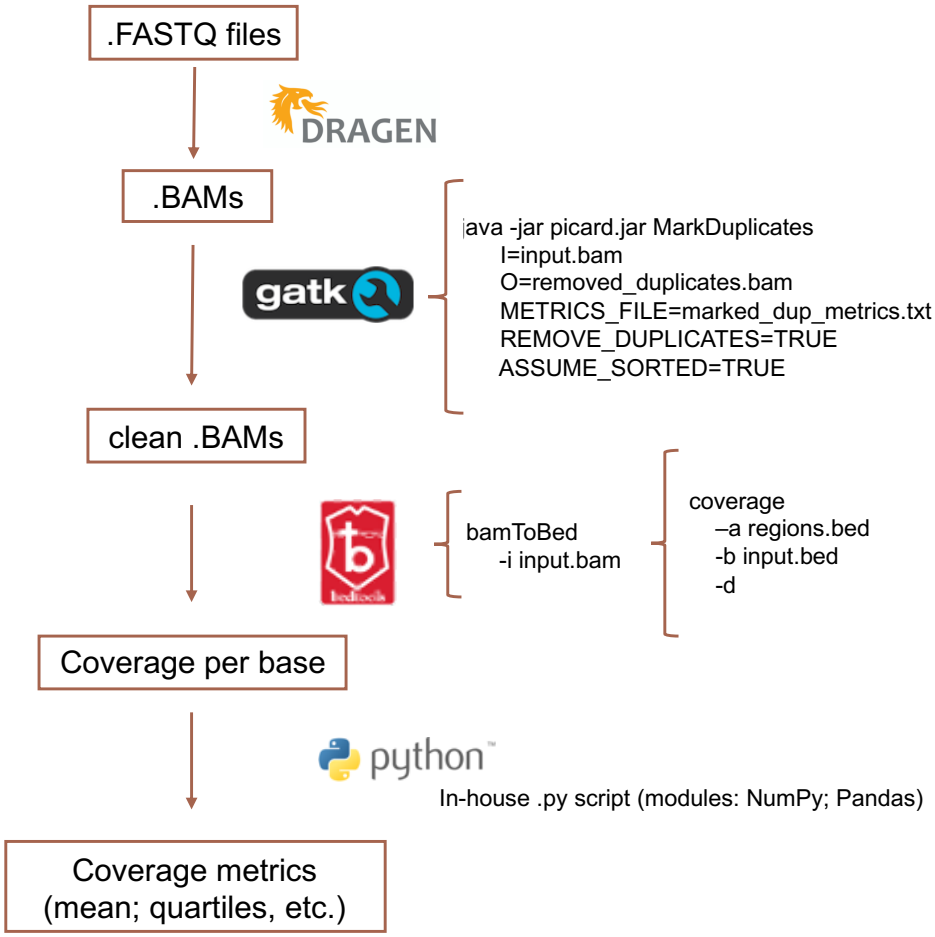


Figure 15. Workflow of coverage assessment for the 14 genes in the PROfound trial

Somatic variant filtering and classification

Before annotation with VarAFT, variants were filtered out following the workflow shown in Figure 16. Briefly, first the *.vcf* files were merged and this unique file was submitted to *vcftools* "counts" function. The generated file contains the number of times that each position/variant appears in the unique *.vcf* file. Using the Excel filters function, variants that appeared in more than 10 samples were filtered out, giving rise to a *.txt* file with only the positions/variants that appears in less than 10 samples. This file was used as a filter for the function "filter" of *bcftools*, in which individual *.vcf* files were submitted. The filtered *.vcf* files were then annotated with VarAFT software.

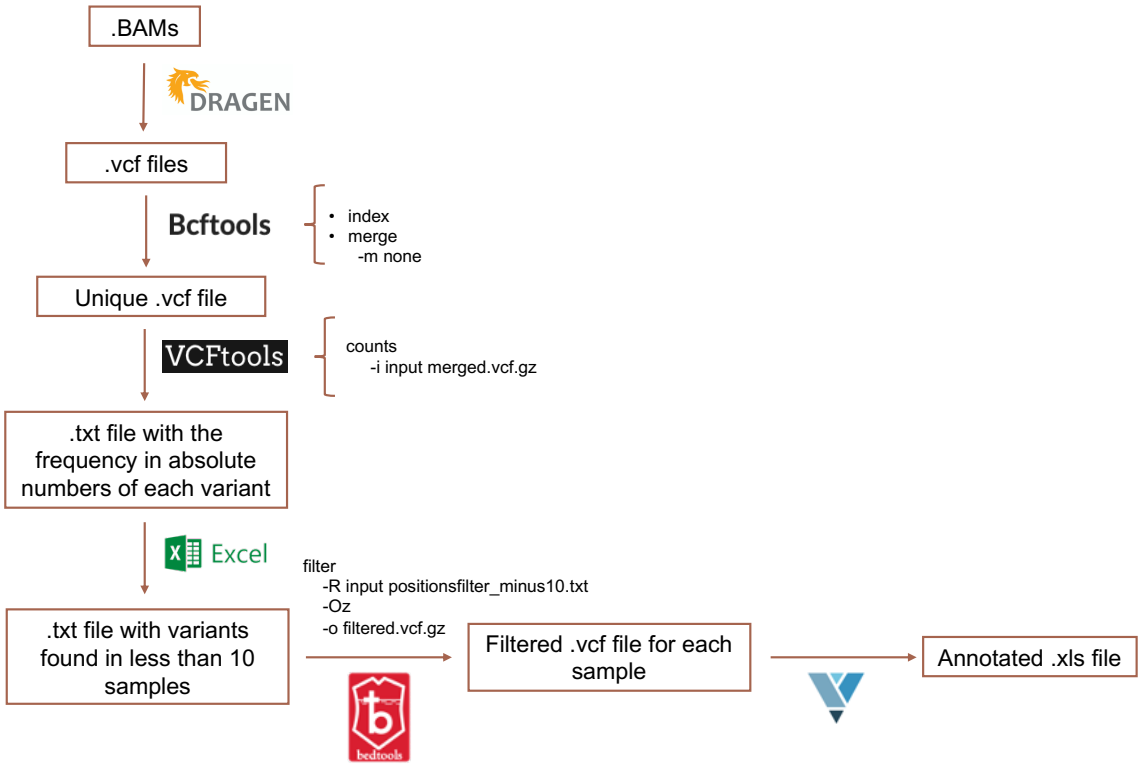


Figure 16. Workflow used to obtain filtered .vcf files from DRAGEN

Following the recommendations of the Standards and Guidelines for the Interpretation and Reporting of Sequence Variants in Cancer (68) and focusing in our goal, variant filtering was performed in the annotated *.vcf* files as showed in Figure 17.

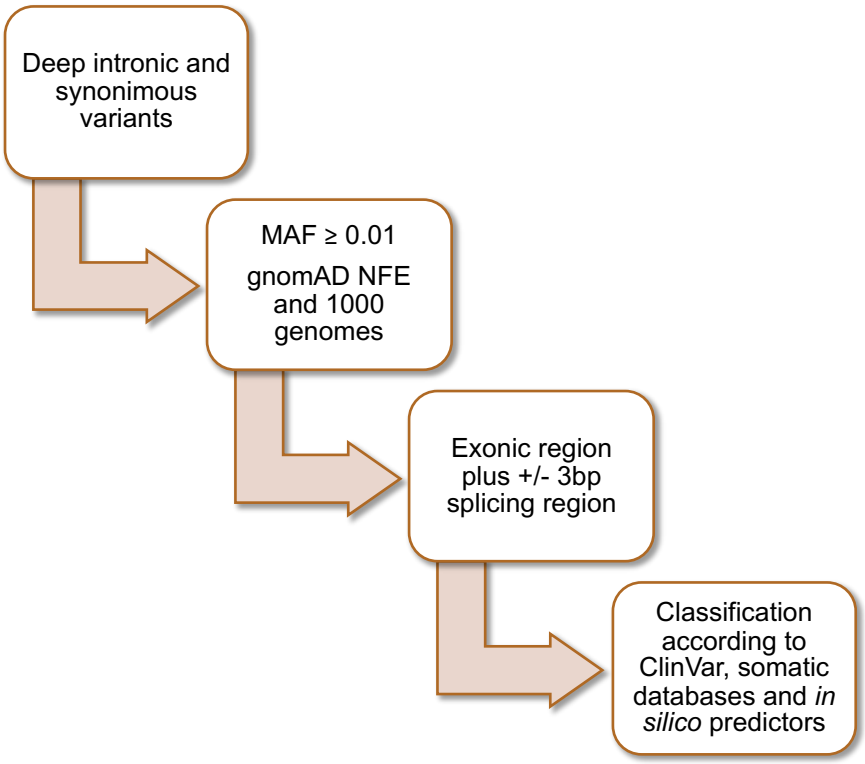


Figure 17. Variant filtering workflow. MAF, minor allele frequency described in population databases; NFE, Non-Finish Europeans

Databases

The following databases were used for analysis and interpretation of germline and somatic variants, according to Standards and Guidelines for Interpretation of Sequence Variants: a joint consensus recommendation of the American College of Medical Genetics and Genomics and the Association for Molecular Pathology (71); Standards and Guidelines for the Interpretation and Reporting of Sequence Variants in Cancer: a joint consensus recommendation of the Association for Molecular Pathology, American Society of Clinical Oncology, and College of American Pathologists (68); and A framework to rank genomic alterations as targets for cancer precision medicine: the ESMO Scale for Clinical Actionability of molecular Targets (ESCAT) (69).

Population databases

Genome Aggregation Database (gnomAD): provides information about exomes and genomes from a variety of large-scale sequencing projects. For these analyses, the considered dataset was of the Non-Finnish Europeans (NFE) (73,74).

The 1000 Genomes Project: provides comprehensive description of common human variants (75–77). There are no populational distinction and the cutoff used was of 1% frequency, that is, excluding only polymorphisms.

Disease databases

ClinVar: provides information about clinical significance and phenotype relationship among human variations, with supporting evidence (78,79). Variants classified as “benign” and “likely benign” were excluded. Variants classified as “uncertain significance” were submitted to *in silico* analysis and variants classified as “likely pathogenic” and “pathogenic” were considered.

OMIM: the Online Mendelian Inheritance in Man database is a comprehensive compendium of human genes and phenotypes (80). The OMIM database was consulted since hereditary cancer syndromes are cataloged, adding information to the variant.

Cancer specific databases

COSMIC: The Catalogue of Somatic Mutations in Cancer is an expert manually curated database of somatic mutation information about human cancers (81,82). COSMIC database has many somatic variants cataloged, linked with cancer types, being a source of information about a given variant.

cBioPortal for Cancer Genomics: stores data sets of various large-scale cancer genomic projects (83–85). This database, besides gathering information about large-scale studies, also annotates variants using the information of many databases, including OncoKB, My Cancer Genome, ClinicalTrials.gov and *in silico* predictors.

OncoKB: contains information about effects and treatment implications of somatic variants, related to specific types of cancer. Treatment information is classified according to the levels of evidence, which have correspondence with the AMP/ASCO/CAP levels of evidence (70,86).

ClinicalTrials.gov: provides information of public and privately supported clinical trials of various diseases and conditions in human volunteers at global level (87).

My Cancer Genome: provides information about driver mutations, their related therapeutic implications, clinical trials, biomarkers, disease and drug ontologies,

therapeutic, prognostic and diagnostic information, pathways and molecular testing techniques (88,89).

Cancer Genome Interpreter (CGI): dedicated to annotating and interpret tumor variants (90,91). Focused in gene actionability.

In silico pathogenicity predictors

These algorithms can predict the damage of missense variants and are recommended in AMP/ASCO/CAP guidelines (68,71). For classification of missense variants of uncertain significance, we performed *in silico* analysis with data gathered from 11 pathogenicity protein predictors and 4 conservative predictors, retrieved from the dbNSFP database (v3.0) by VarAFT. Variants of uncertain significance (VUS) were considered more likely to be pathogenic/deleterious if showing evolutionary conservation by at least 3 conservation predictors and, in total, at least 12 predictors are indicative of pathogenicity (92). The *in silico* protein damage predictors used were SIFT, Polyphen HDIV, LRT, Mutation Taster, Mutation Assessor, FATHMM, PROVEAN, VEST3, Meta SVM, Meta LR and CADD. The *in silico* evolutionary conservation predictors used were GERP, phyloP, SiPhy and PhastCons.

Next Generation Sequencing of gDNA samples

Genomic DNA was sequenced as part of the diagnostic routine of the Department of Genetics, following the department's Standard Operating Procedures. The panel used was Illumina TruSight Hereditary Cancer Panel (v2, Illumina, Inc), a cancer comprehensive panel with 117 genes associated with cancer predisposition (93), which covers 13 of the 14 genes of the PROfound trial (only *CDK12* is not covered, and it only has somatic variants).

Bioinformatic analysis was made in NextGENe software in a routine manner, following the Standard Operating Procedures of the Department of Genetics. Germline variants were classified according to the ACMG-AMP guidelines. For diagnostic purposes, only pathogenic and likely pathogenic variants are reported.

Sanger sequencing

Samples whose tumor had *CDK12* variants of possible somatic origin (high MAF), were submitted to Sanger sequencing.

Variant-specific primers were designed using Primer Designing Tool (NCBI – NIH) (94), with an M13 tail (sequences below). Samples were amplified and the amplification of the PCR product was confirmed by agarose gel. After confirmation, the product was purified with ExoSAP-IT^{TM} PCR product cleanup reagent (Applied Biosystems, ThermoFisher Scientific) and forwarded to sequencing PCR. The sequencing product was purified using Sephadex G-50 (Sigma Aldrich) and then forwarded to capillary electrophoresis in a 3500 Genetic Analyzer (Applied Biosystems, ThermoFisher Scientific). Electropherograms were inspected in 4Peaks software (Nucleobytes B.V. – The Netherlands).

Primer sequences:

Forward primer: TGTAAAACGACGGCCAGTGGAAGTACCTCGAGAAGGCAG

Reverse primer: CAGGAAACAGCTATGACCGACCATAGGGACTGGGCGAT

RESULTS

Patient samples

A total of 47 patients have been recruited at the time of run. Clinico-pathological data and sample info are summarized in Appendix 1. The ages at diagnosis ranged from 47 to 84 years, median of 68 years. The lowest Gleason score was 7a (3+4) and the highest was 10, in two patients each, with a prevalence of the scores 8 and 9, in 15 and 13 patients, respectively.

The majority of tumor samples were obtained from prostate biopsies (n=40), and the remaining seven were obtained from prostatectomies. Of all patients, four had no tumor available for analysis and 14 had two separated biopsies of the same procedure, totalizing 57 tumor samples from 43 patients.

Regarding other biological samples, 46 patients collected blood and one collected buccal swab. The buccal swab was not submitted to analysis due to lack of quantity of DNA, and from one blood sample *buffy coat* pellet was mistakenly discarded, totalizing 45 patients with gDNA available.

Tumor DNA analysis by next generation sequencing

All 57 tumor samples were submitted to DNA quality control using the Tapestation Genomic Screentape assay. Samples with concentration warning in the Tapestation Genomic assay output were quantified by Qubit dsDNA HS Assay (ThermoFisher).

Seventeen samples were excluded from the run, 14 were patients' duplicates, from which the best DNA sample (in quality and quantity) was chosen, and three (patients 16, 19 and 37) had null DIN and/or low yield (Table 6), and, therefore, tumor DNA was not sequenced.

Table 6. Tumor DNA samples quality and quantity. Samples in grey were not included in the NGS run.

Sample number	DIN	[Tapestation]	[Qubit]
1	5.1	9.3	
2	4.5	35.2	
3_B1	2.6	7.17	
3_B2	2.5	4.1*	
4	5.2	6.49	6.88
5	4.5	29.3	
6_A	3.4	44.3	
6_B***	3.5	19.6	
7	4.8	57.7	
8	4.2	17.3	
9	5.2	15.4	
10	5.4	63,00	
13	2.9	54.9	
14	2.3	14.8	
15	3.3	93.2	61.2
16	3.2	2.86*	
17	5.5	9.58	
18	2	53.2	
19	0**	2.39*	
20_E2	5.1	10.9	
20_E1***	4.4	6.82	
21	3.6	5.98	
22_2	2.9	17.5	
22_1***	2.5	5.96	
23	3.2	95.2	
25_2	2.2	23.9	
25_1***	2.1	13.6	
26	3.7	14.3	

Sample number	DIN	[Tapestation]	[Qubit]
27_2	5.1	15.2	
27_1***	4.9	16.5	
28_1	2.1	53.1	
28_2***	2	38.1	
29_E1	2.9	10.7	
29_E2***	2.5	5.98	
30_1	3.9	13.2	
30_2***	3.9	9.35	
31	5.3	111,00	
32_2	5.1	7.84	
32_1***	5.2	5.26	
33_1	2.2	25.4	
33_2***	2.1	16.6	
35	5.5	24.9	
36_2	3.4	27.2	13.3
36_1***	3.8	21.8	17.1
37_D	0**	1.2	0*
37_E	0**	1.41	0.224*
38	4.7	5.11	4.74
39	3.6	115,00	66.6
40	2.9	58.4	41.2
41	5.2	10.1	8.84
42	5	51.2	49.8
43	5.5	27.4	23.6
44	3	101,00	53.4
45_E	2.1	11.9	4.5
45_D	1.8	3.35	1.03*
46	3.3	21.9	14.9
47	5.3	9.69	9.88

** low DNA yield samples; ** null DIN samples; *** duplicate samples.*

Germline DNA analysis by next generation sequencing

Germline samples (DNA extracted from *buffy coat* pellets) were submitted to NGS using TruSight Hereditary Cancer panel. Of the 44 patients with gDNA available, five were not screened because of low DNA yield. The DNA extraction of these four samples was already repeated and the NGS test will be performed in the next run. Table 9 shows the patients that had tumor and germline tests performed. One patient had no tumor or germline test available (Table 7).

Table 7. Tumor and germline NGS tests performed for each patient.

Patient number	Tumor NGS	Germline NGS	Patient number	Tumor NGS	Germline NGS
1	Yes	Yes	25	Yes	Yes
2	Yes	No***	26	Yes	Yes
3	Yes	Yes	27	Yes	Yes
4	Yes	Yes	28	Yes	Yes
5	Yes	Yes	29	Yes	No*
6	Yes	Yes	30	Yes	Yes
7	Yes	No*	31	Yes	No**
8	Yes	Yes	32	Yes	Yes
9	Yes	Yes	33	Yes	Yes
10	Yes	Yes	34	No***	Yes
11	No***	Yes	35	Yes	Yes
12	No***	Yes	36	Yes	Yes
13	Yes	Yes	37	No*	Yes
14	Yes	Yes	38	Yes	Yes
15	Yes	Yes	39	Yes	Yes
16	No*	Yes	40	Yes	No***
17	Yes	Yes	41	Yes	Yes
18	Yes	Yes	42	Yes	No*
19	No*	No*	43	Yes	Yes
20	Yes	Yes	44	Yes	No*
21	Yes	Yes	45	Yes	Yes
22	Yes	Yes	46	Yes	Yes
23	Yes	Yes	47	Yes	Yes
24	No***	Yes			

**Low DNA yield. **Buccal swab sample. ***No sample*

Next Generation Sequencing

Control samples run

The DIN score of the 10 control samples ranged from 2.7 to 8.7, with FFPE samples having lowest DIN, as shown in Table 8. Quantity of DNA was measured in Tapestation Genomic run. Run metrics are summarized in Table 8.

The alignment metrics showed an extremely high percentage of PCR duplicates, leading to low sequence efficiency and, in consequence, low average depth of coverage for this test (number of samples *versus* panel size *versus* flowcell output).

Table 8. Control samples quality and alignment metrics

Sample number	Sample type	DIN	[ng/ul]	Input [ng]	% PCR duplicates	Average depth of coverage
S1	FFPE	2.7	22.4	157	94.03	107.22
S2	FFPE	5.5	28	200	85.01	351.63
S3	FFPE	5.8	37	200	76.18	676.85
S4	FFPE	5.1	39	200	70.82	737.27
S5	FFPE	4.0	40	200	79.39	497.75
S6	FFPE	4.1	37	200	74.68	646.44
S7	FFPE	2.5	48	200	87.8	254.46
S8	Blood	8.7	42	200	69.15	745.69
S9	FFPE	5.9	37	200	67.82	633.86
S10	Blood	8.5	64	200	58.45	626.73

To improve sequencing efficiency, the “two hybridizations” strategy was maintained, but with less PCR cycles in post-hybridization steps in order to reduce PCR duplicates. In addition, although the minimum DNA concentration recommended by Agilent’s SureSelect library preparation protocol is 10ng, we observed that samples with low DIN (below 3) and low DNA concentration were poorly sequenced, so we defined a minimum DNA input of 50ng.

Regarding bioinformatic analysis, almost all variants were identified in all samples, with equivalent quality throughout the three software. The exceptions were large deletions, which SureCall did not detect and Basespace Dragen identified, namely, the *BRCA2* Alu insertion. Because of these results, the software chosen for the analysis of tumor samples was Basespace Dragen – Somatic Pipeline. Control samples results are shown in Appendix 2.

Prostate tumor samples run

Quality control

Cluster density was under Illumina’s specified parameters, with raw cluster density of 198K/mm^3 and 91.28% of the clusters passing filter, that is, approximately 180K/mm^3 of clusters passing filter. The percentage of bases with quality higher than Q30 was of 92.91%. Sample representativeness in this run was relative homogeneous, with only two samples more represented and that did not affect the overall results.

PCR duplicates and depth of coverage

The alignment metrics showed a decrease in PCR duplicates. Some samples still had high percentage of duplicates and this was probably due to low DNA quality and/or quantity. The average depth of coverage is proportionally better than in the control samples run, with values compatible with number of samples *versus* panel size *versus* flowcell output, despite being below of what is recommended in guidelines. Mapped reads on target is variable according to sample quality and/or average depth of coverage (Table 9).

Table 9. Prostate tumor samples (n=40) quality and alignment metrics

Sample number	DIN	Input [ng]	% PCR Duplicates	Average depth of coverage	Mapped reads on target	Sample number	DIN	Input [ng]	% PCR Duplicates	Average depth of coverage	Mapped reads on target
1	5.1	93	50.99	317.84	72.56	26	3.7	143	59.9	296.71	64.84
2	4.5	200	41.34	424.83	74.23	27	5.1	152	63.19	256.36	70.03
3	2.6	71.7	78.73	101.28	64.53	28	2.1	200	63.86	199.25	61.33
4	5.2	68.8	51.36	327.04	72.11	29	2.9	107	70.56	173.36	58.54
5	4.5	200	54.44	271.62	71.28	30	3.9	132	60.81	306.64	65.98
6	3.4	200	48.54	318.85	71.74	31	5.3	200	29.32	488.89	76.67
7	4.8	200	31.94	457.32	76.82	32	5.1	78.4	55.28	247.36	71.57
8	4.2	173	35.2	452.58	75.25	33	2.2	200	52.58	316.57	63.93
9	5.2	154	43.85	402.15	70.07	35	5.5	200	36.05	435.97	74.27
10	5.4	200	43.32	464.95	75.76	36	3.4	200	46.94	308.22	72.5
13	2.9	200	63.11	226.03	69.27	38	4.7	47.4	57.11	230.9	67.42
14	2.3	148	76.83	101.04	60.03	39	3.6	200	48.5	362.68	73.48
15	3.3	200	49.57	325.42	72.36	40	2.9	200	25.54	256.66	72.48
17	5.5	95.8	42.15	381.48	72.34	41	5.2	101	47,00	316.31	71.72
18	2	200	58.36	207.12	58.86	42	5	200	23.87	306.77	76.7
20	5.1	109	64.17	241.22	68.47	43	5.5	200	33.47	458.78	74.74
21	3.6	59.8	87.99	158.75	58.06	44	3	200	50.51	232.54	73.95
22	2.9	175	86.53	66.92	40.91	45	2.1	119	77.81	101.25	53.51
23	3.2	200	39.27	484.57	74.3	46	3.3	200	45.33	419.41	71.31
25	2.2	200	57.41	472.19	59.46	47	5.3	96.9	50.75	374.38	71.08

Panel average coverage was calculated taking into consideration either the whole panel (30 DDR genes plus other genes that were not part of this study) or the 14 DDR actionable genes, that is, the genes of homologous recombination that have indication for therapy with Olaparib (Table 10). The coverage for the whole panel was calculated in multiple ranges of depth, with the percentage of analyzable regions that have that depth.

For diagnostic purposes, the Department of Genetics established a metric of at least 95% of analyzable regions with 500X depth of coverage. With this run configuration (40 samples in a mid-output flowcell), none of samples fulfills this metric, but a decrease in the number of samples per run should solve this issue (Table 10).

During the coverage analysis, we identified a serious problem: during panel improvements, accidentally, probes for the *BRCA1* gene were not fully synthetized in the

custom panel, with only six exons being covered. This issue is being solved, but, for this study, analysis of *BRCA1* is concluded for the germline only.

Table 10. Coverage metrics per sample (whole panel and 14 genes) and percentage of analyzable regions in different ranges of depth

Sample number	Average depth of coverage	Average depth of coverage 14 genes	Percentage of analyzable target regions covered by at least 50 reads	Percentage of analyzable target regions covered by at least 100 reads	Percentage of analyzable target regions covered by at least 200 reads	Percentage of analyzable target regions covered by at least 500 reads
1	317.84	430.0	99.58	98.96	93.36	66.51
2	424.83	565.0	99.69	98.90	94.88	75.55
3	101.28	115.0	99.37	94.88	65.67	10.50
4	327.04	788.0	99.63	98.38	91.80	67.08
5	271.62	326.0	99.53	98.43	92.95	64.37
6	318.85	345.0	99.63	98.59	93.68	66.56
7	457.32	328.0	99.58	98.59	94.83	73.20
8	452.58	415.4	99.69	98.85	94.78	74.92
9	402.15	443.0	99.63	98.54	93.78	73.04
10	464.95	192.0	99.69	99.27	96.92	78.94
13	226.03	383.0	99.58	98.28	90.70	54.65
14	101.04	120.0	97.65	89.29	61.55	9.77
15	325.42	370.0	99.58	97.91	93.05	68.13
17	381.48	367.0	99.58	98.75	95.04	70.95
18	207.12	226.0	98.75	95.35	83.12	45.61
20	241.22	277.0	99.74	98.85	92.89	61.39
21	158.75	178.0	99.48	97.60	85.27	29.99
22	66.92	92.0	93.31	76.49	38.40	2.61
23	484.57	399.0	99.69	98.69	95.09	72.62
25	472.19	495.0	99.74	99.37	97.96	82.39
26	296.71	432.0	93.31	87.51	80.72	61.49
27	256.36	296.0	99.69	98.69	93.36	62.43
28	199.25	224.0	99.58	97.60	87.46	44.36
29	173.36	232.0	96.50	90.13	76.59	37.04
30	306.64	401.0	97.81	94.36	86.21	66.77
31	488.89	476.0	99.58	98.69	94.57	73.98
32	247.36	309.0	98.96	95.66	85.68	57.63
33	316.57	344.0	99.69	98.85	94.57	68.23
35	435.97	473.0	99.63	98.59	94.04	72.10
36	308.22	348.0	99.63	98.28	91.90	65.83
38	230.90	285.0	97.75	93.68	83.44	53.92
39	362.68	410.0	99.79	99.01	95.25	72.36
40	256.66	250.0	98.75	95.25	81.87	53.13
41	316.31	307.0	99.69	98.43	92.74	64.79
42	306.77	282.0	99.22	96.13	84.01	56.11
43	458.78	441.0	99.69	98.17	93.68	72.94
44	232.54	225.0	98.54	95.30	84.95	50.89
45	101.25	120.0	98.80	92.48	63.53	10.66
46	419.41	425.0	99.79	99.32	96.66	75.24
47	374.38	428.0	99.74	98.64	93.78	72.20

According to the AMP/CAP guidelines for Validation of Next Generation Sequencing-based oncology panels, the depth of coverage necessary to detect a given allele frequency can be estimated using the binomial distribution equation. Excel allows to calculate the

binomial probability using the BINOM.DIST function, based on the binomial distribution equation (Figure18), where “number_s” is the number of variant reads that contains the mutation (*x* in binomial distribution equation), “trials” is the total number of reads (depth of coverage, *n* in the binomial equation), “probability_s” is the probability of detecting a variant allele (VAF, *p* in binomial equation) and “cumulative” should be set as TRUE, once this probability is cumulative for a given *x* and *n* (65).

$$P(x) = \frac{n!}{x!(n-x)!} p^x (1-p)^{n-x}$$

Figure 18. Binomial distribution equation

Using the thresholds determined in the guideline, that is, at least 5 reads with the mutation in a sample with 15% of variant allele frequency, the true positive probability of detecting this variant in at least 5 reads is 99.957% and the probability of false negative is 0.043%, in a minimum depth of coverage of 100X (Table 11). This shows that our test, without any improvement, has a theoretical limit of detection of 15% and this sensibility can be improved increasing the average depth of coverage, which will be discussed.

Table 11. Binomial distribution for different depths of coverage and variant allele frequencies

True Positive binomial distribution Probability of detetct four or fewer mutated reads - threshold in 5 mutated reads										
Allele burden	15%	15%	10%	10%	5%	5%	3%	3%	3%	3%
Depth	60X	100X	100X	150X	200X	250X	300x	350X	400X	450X
Binomial distribution	0.04237	0.00043	0.02371	0.00055	0.02645	0.00457	0.05243	0.01975	0.00698	0.00234
Prob. of true positive	95.763%	99.957%	97.629%	99.945%	97.355%	99.543%	94.757%	98.025%	99.302%	99.766%
Prob. of false negative	4.237%	0.043%	2.371%	0.055%	2.645%	0.457%	5.243%	1.975%	0.698%	0.234%

Variant filtering and interpretation

Variant filtering was performed as described in methodology. Figure 19 shows the number of remaining variants at each filtering step. The main somatic annotator used was OncoKB, a manually curated database and annotator, recommended by ESMO (69).

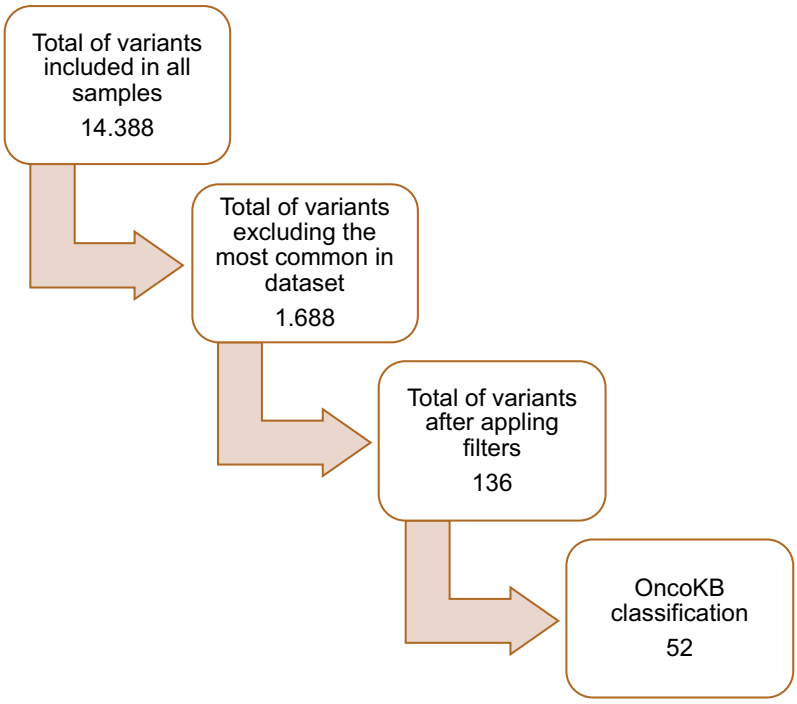


Figure 19. Variant filtering - number of variants remaining at each step of filtering, considering the 30 genes

Clinically actionable variants

Eleven variants had OncoKB classification of “Oncogenic” or “Likely Oncogenic”, with indication for Olaparib treatment (Table 12). Variants were also classified following the AMP/ASCO/CAP guidelines and the ESCAT system. Information of Cancer Genome Interpreter (CGI) was aggregated to OncoKB information, as well as clinical trial information.

Overall, these 11 variants are present in 20% (n=8) of the 40 patients tested. The most frequently altered gene is *CDK12*, with five variants in three patients, followed by variants in *ATM*, *BRCA2* and *CHEK2* variants, found in two patients each.

All altered genes, except *CDK12*, are part of the TruSigh Hereditary Cancer Panel, therefore germline variants in *BRCA2* and *CHEK2* were confirmed by NGS. *ATM* variants were not found in the NGS analysis of the matched gDNA, being considered somatic. The *CDK12* variant c.929C>A, with an allele frequency with potential to be germline, was confirmed somatic by Sanger sequencing (Appendix 3).

Table 12. Somatic variant classification according to somatic databases and interpretation according to AMP/ASCO/CAP and ESCAT systems

Sample number	Gene	Transcript	Variant		Consequence	ClinVar	Dragen		SureCall		OncoKB				CGI				Clinical Trial	Level of Evidence	
			cDNA	Protein			Depth	VAF	Depth	VAF	Clinical Implications	Biological Effect	Level	Drug	Gene role	Drug	Biomarker	Effect		AMP/ASCO/ CAP	ESCAT
4	*ATM*	NM_000051	c.1607+1G>T		splicing;intronic	Pathogenic/Likely pathogenic	230	0,23	227	0,24	Likely Oncogenic	Likely Loss of function	1	Olaparib	Loss of Function	Olaparib	ATM oncogenic mutation	Responsive	NCT02987543 NCT01682772	Tier 1 Level A	Tier I-A
9	*CDK12*	NM_016507	c.1046_1046+1delinsTT		exonic;splicing	.	529	0,26	477	0,25	Likely Oncogenic	Likely Loss of function	1	Olaparib	Loss of Function	Olaparib	.	.	NCT02987543 NCT01682772	Tier 1 Level A	Tier I-A
9	*CDK12*	NM_016507	c.2772_2785del	p.C924fs	frameshift deletion	.	291	0,10	NA	NA	Likely Oncogenic	Loss of Function	1	Olaparib	Loss of Function	Olaparib	.	.	NCT02987543 NCT01682772	Tier 1 Level A	Tier I-A
21	*ATM*	NM_000051	c.216_217del	p.T72fs	frameshift deletion	Pathogenic	104	0,19	99	0,17	Likely Oncogenic	Likely Loss of function	1	Olaparib	Loss of Function	Olaparib	ATM oncogenic mutation	Responsive	NCT02987543 NCT01682772	Tier 1 Level A	Tier I-A
27	***CHEK2***	**NM_007194**	**c.349A>G**	**p.R117G**	**nonsynonymous SNV**	**Pathogenic/Likely pathogenic**	**256**	**0,43**	**264**	**0,41**	**Likely Oncogenic**	**Loss of Function**	**1**	**Olaparib**	**Loss of Function**	**Olaparib**	**CHEK2 oncogenic mutation**	**Responsive**	**NCT02987543 NCT01682772**	**Tier 1 Level A**	**Tier I-A**
27	*CDK12*	NM_016507	c.2907delC	p.F969fs	frameshift deletion	.	290	0,29	296	0,29	Likely Oncogenic	Likely Loss of function	1	Olaparib	.	.	.	.	NCT02987543 NCT01682772	Tier 1 Level A	Tier I-A
27	*CDK12*	NM_016507	c.1931+1insT		splicing;intronic	.	111	0,25	120	0,22	Likely Oncogenic	Likely Loss of function	1	Olaparib	.	.	.	.	NCT02987543 NCT01682772	Tier 1 Level A	Tier I-A
29	*CDK12*	NM_015083	c.929C>A	p.S310*	stopgain	.	196	0,37	200	0,37	Likely Oncogenic	Likely Loss of function	1	Olaparib and Rucaparib	.	.	.	.	NCT02952534 NCT01682772 NCT02987543	Tier 1 Level A	Tier I-A
30	***BRCA2***	**NM_000059**	**c.4814delT**	**p.V1605fs**	**frameshift deletion**	**.**	**220**	**0,36**	**223**	**0,36**	**Likely Oncogenic**	**Likely Loss of function**	**1**	**Olaparib and Rucaparib**	**Loss of Function**	**Olaparib**	**BRCA2 oncogenic mutation**	**Responsive**	**NCT02952534 NCT01682772 NCT02987543**	**Tier 1 Level A**	**Tier I-A**
43	***BRCA2***	**NM_000059**	**c.3680_3681del**	**p.L1227fs**	**frameshift deletion**	**Pathogenic**	**335**	**0,69**	**305**	**0,68**	**Likely Oncogenic**	**Likely Loss of function**	**1**	**Olaparib and Rucaparib**	**Loss of Function**	**Olaparib**	**BRCA2 oncogenic mutation**	**Responsive**	**NCT02952534 NCT01682772 NCT02987543**	**Tier 1 Level A**	**Tier I-A**
45	***CHEK2***	**NM_007194**	**c.349A>G**	**p.R117G**	**nonsynonymous SNV**	**Pathogenic/Likely pathogenic**	**68**	**0,53**	**78**	**0,50**	**Likely Oncogenic**	**Likely Loss of function**	**1**	**Olaparib**	**Loss of Function**	**Olaparib**	**CHEK2 oncogenic mutation**	**Responsive**	**NCT02987543 NCT01682772**	**Tier 1 Level A**	**Tier I-A**

Germline variants are highlighted in bold

Not clinically actionable variants

Variants with uncertain significance (VUS) are considered non-actionable, and include both missense variants found in the 14 clinically actionable genes and variants, deleterious and missense, found in the remaining 16 DDR genes covered by the custom panel. As described in methodology, missense variants that are classified as VUS in ClinVar were submitted to *in silico* analysis and only those with higher probability to be deleterious are described (Table 13). These variants may be reclassified as "actionable" in the future, but meanwhile they cannot be used in any clinical decisions.

Table 13. Non-actionable VUS identified in the 47 tumor samples

Non-actionable variants										
Sample number	Gene	Transcript	Variant cDNA	Protein	Consequence	ClinVar	Dragen Depth	Dragen VAF	SureCall Depth	SureCall VAF
2	*PMS2*		c.2186_2187del	p.L729fs	frameshift deletion	Uncertain_significance	94	0,13	-	-
22	***ATM***	**NM_000051**	**c.2932T>C**	**p.S978P**	**nonsynonymous SNV**	**Conflicting_interpretations_of_pathogenicity**	**32**	**0,47**	**41**	**0,41**
23	***MLH1***	**NM_000249**	**c.649C>T**	**p.R217C**	**nonsynonymous SNV**	**Uncertain_significance**	**262**	**0,53**	**272**	**0,53**
30	*PMS2*		c.2186_2187del	p.L729fs	frameshift deletion	Uncertain_significance	64	0,17	91	0,12
10	***FANCA***	**NM_000135**	**c.1874G>C**	**p.C625S**	**nonsynonymous SNV**	**Conflicting_interpretations_of_pathogenicity**	**358**	**0,45**	**191**	**0,44**
28	***FANCA***	**NM_000135**	**c.1874G>C**	**p.C625S**	**nonsynonymous SNV**	**Conflicting_interpretations_of_pathogenicity**	**108**	**0,17**	**187**	**0,15**
22	***ERCC3***	**NM_000122**	**c.2111C>T**	**p.S704L**	**nonsynonymous SNV**	**.**	**57**	**0,39**	**63**	**0,43**
36	***MRE11A****	**NM_005591**	**c.1378G>T**	**p.E460X**	**stopgain**	**.**	**470**	**0,46**	**528**	**0,47**

Germline variants are highlighted in bold

Germline findings and family history

Of the two patients that have germline variants in *BRCA2*. One has a family history with one paternal uncle with brain cancer, the mother and one maternal aunt with brain cancer, the maternal grandfather with brain cancer and the maternal grandmother with gynecological cancer (Figure 20A). The second family has no family history of cancer (Figure 20B).

Of the patients carrying the *CHEK2* germline variant c.349A>G, one has limited family history of cancer, with only the mother diagnosed with breast and colon cancers (Figure 20C), while the other has a sister affected with melanoma, a brother with cancer (unknown origin) and a paternal cousin with prostate cancer (Figure 20D).

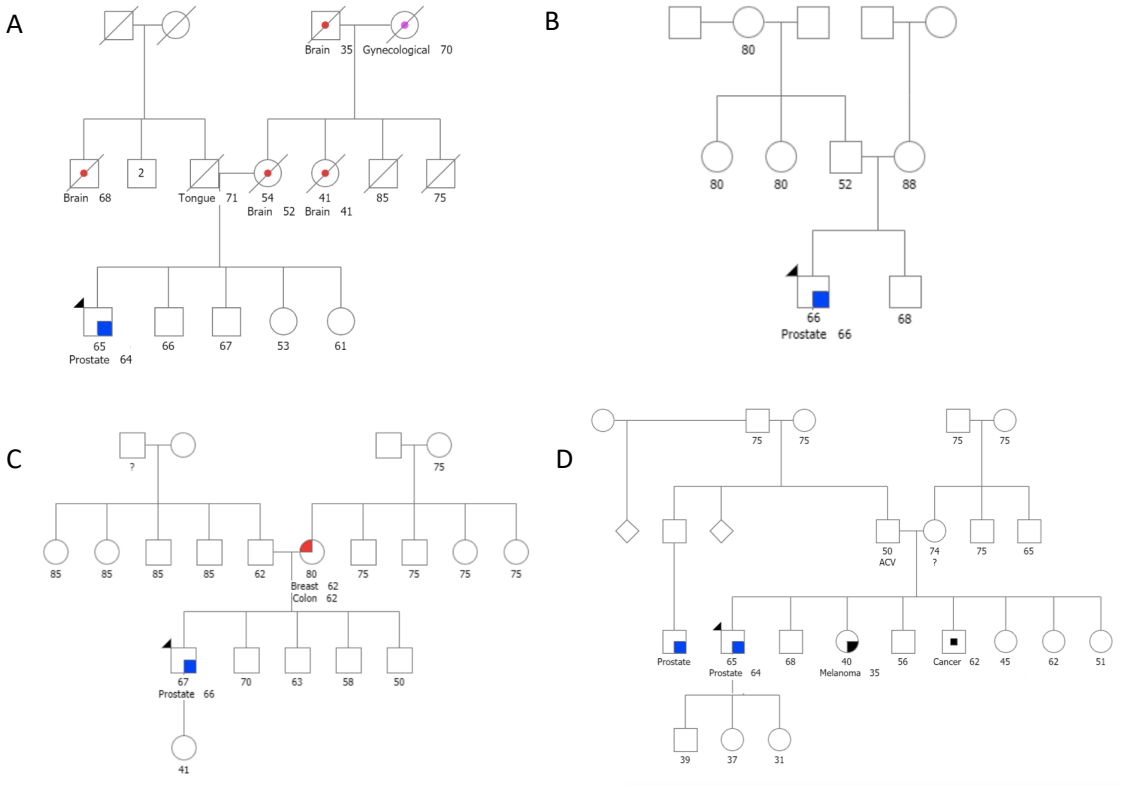


Figure 20. Pedigree of carriers of actionable germline variants. A, patient MetPC43 carrying the BRCA2 variant c.3680_3681del; B, patient MetPC30 carrying the BRCA2 variant c.4814delT; C and D, patients MetPC27 and MetPC45, respectively, carrying the CHEK2 variant c.349A>G;

DISCUSSION

Panel optimization

Formalin-fixed paraffin-embedded samples are known as containing poor DNA quality. The fixation process leads to DNA damages, turning this type of sample difficult to amplify and with artefacts from the fixation process (95). Library preparation protocol was modified in order to maximize the performance of highly degraded samples, but optimization was needed.

The first run, with positive controls, showed extremely high percentage of PCR duplicates, some samples reaching the rate of 90% or more. During bioinformatic analysis of enrichment-based library preparation, PCR duplicates are marked and removed (64), remaining only fragments that are different from each other, that is, fragments that, in theory, are not originated from PCR clonal amplification of the same DNA molecule. This step minimizes the variant calling of PCR artefacts. The level of PCR duplicates decreased with less PCR cycles in post-hybridization step on the second run, but remained elevated in samples with poor quality and/or quantity of DNA. In general, samples with DIN below 3 and low DNA yield had higher rates of PCR duplicates (above 50%).

The analysis of positive controls with three software allowed us to compare different pipelines. NextGENe software recognized almost all known variants, except for the Alu insertion in *BRCA2* gene and one *MSH2* splice region. These variants are especially difficult to recognize for being in repetitive regions of the genome or, in the case of the Alu insertion, being a large insertion of a repetitive element on the genome. SureCall had the worst performance, failing to recognize large deletions and to identify the Portuguese founder *BRCA2* Alu insertion. The best software performance was achieved with Basespace DRAGEN – Somatic Pipeline, identifying all known SNVs and indels. Some variants had a low-quality flag and these regions were improved in the panel design subsequently used in the second run with patients tumor samples. This run was closer to the reality of a daily diagnostic routine, covering the diverse range of DNA quality that characterizes FFPE samples. Based on the first run, we could set parameters to DNA quality and quantity for each sample, with samples with DNA concentrations less than 5ng/ul and DIN score below 3 not being suitable for achieving the level of accuracy needed for this NGS approach.

Regarding depth of coverage, the majority of samples in the tumors run did not reach the AMP/CAP guideline metrics (65) for a diagnostic routine, that is, at least 250X of depth of coverage per target, with a limit of detection of 5%, and none of them reached the minimum coverage validated at the Department of Genetics routine and recommended by the IQN Path guideline (66), that is, 500X in at least 95% of targets. Considering the high heterogeneity of DNA quality among tumor DNA samples from FFPE tissue, with samples with worse quality and/or quantity having more PCR duplicates, a new run will be performed with less samples and a high output flowcell, to increase the depth of coverage to the needed diagnosis level and overcome the high percentage of duplicates common to low quality FFPE samples. Attending to the theoretical limit of detection of this run (set at 15%) in a minimum coverage of 100X, ideally, the next run should include a set of samples with different allele frequencies, some closer to the theoretical limit of detection and some below it, in order to validate the sensibility and the power of the test, adding some statistical data, such as positive predicted value, positive percentage agreement, reproducibility and real limit of detection (65).

Targeted therapy - PARP inhibitors treatment eligibility

According to the PROfound clinical trial (40), the prevalence of alterations, somatic or germline, in one or more of the 14 homologous recombination genes analyzed (*PPP2R2A* not included) was about 24%, which is similar to our prevalence of 20%. It is important to note that we did not consider the *PPP2R2A* gene as a biomarker, since in this same clinical trial patients harboring alterations in this gene showed no response to Olaparib treatment and, for this reason, *PPP2R2A* was excluded from NCCN Guidelines for Prostate Cancer (16). The main limitation of this study as of today concerns the *BRCA1* gene, which was only analyzed in the germline samples. Attending to a reported frequency about 10-fold lower for somatic *BRCA1* variants comparing with somatic *BRCA2* variants in prostate cancer, it is unlikely that any of the 40 patients analyzed has a somatic *BRCA1* variant, something that we will be able to clarify when the study is concluded.

The final overall survival data of the *BRCA1*, *BRCA2* and *ATM* arm of the PROfound study was recently presented at the ESMO Virtual Congress 2020. The study concluded that patients harboring deleterious somatic or germline mutations in these genes had a meaningful prolongation of overall survival, with a reduction of 31% of risk

of death (96). With four patients harboring variants in *ATM* and *BRCA2*, we can presume that 10% of our patients may benefit from an extremely meaningful decrease of risk of death. After validation of the findings with further clinically validated methods, this information will be made available whenever the treating clinician asks for it. Futhermore, the *CHEK2* and *CDK12* genes are included in the FDA approval for PARPi, but not yet in the EMA recommendations. We must threfore wait for the final decision of EMA and INFARMED about exactly which genes will be approved as predictive markers for PARPi treatment of mCRPC patients.

Regarding germline alterations, Pritchard et al. (28) showed that the prevalence of DNA-repair alterations in men with metastatic prostate cancer at diagnosis, unselected for family history, is 11.8% (*BRCA2* – 5.3%, *ATM* – 1.6%, *CHEK2* – 1.9%, *BRCA1* – 0.9%, *RAD51C* – 0.4% and *PALB2* – 0.4%). Our findings are in accordance with the literature, with 10% of our cohort carrying germline alterations, 5% of the patients with *BRCA2* deleterious alterations and additional 5% with *CHEK2* likely pathogenic alterations. This two-fold higher frequency of *CHEK2* alterations, resumed to the c.349A>G variant, is in line with the founder effect of this variant in the Portuguese population, recently reported (97).

Regarding non-actionable variants found in this study, they are all in genes that do not have clinical implications/treatment indication, except for *PMS2* and *MLH1* variants, which are in actionable genes for Pembrolizumab treatment (16). The *PMS2* frameshift somatic variant, found in two samples, is classified as non-actionable due to its VUS classification in ClinVar. In fact, because of the pseudogene *PMS2CL*, variants found in *PMS2* after exon 11 have to be confirmed with a nested long-range PCR technique, which amplifies only the *PMS2* gene in a first PCR cycling and the product is used as template to a second regular PCR, followed by Sanger sequencing (98). Due to its DNA degraded nature, it was not possible to perform a long-range PCR in DNA from FFPE tissue. The missense *MLH1* germline variant has uncertain significance in ClinVar, so is not considered for clinical/treatment purposes. In any case, we will request immunohistochemistry analysis for the MMR proteins in those cases in order to ascertain if there is an effect of those variants on protein expression.

Germline variants, genetic counseling, and follow-up

Besides PARPi treatment indications, the four families harboring germline variants are recommended for genetic counseling. The patient *MetPC43*, harboring the truncating

BRCA2 c.3680_3681del variant, has family history of brain cancer in both sides of the family. His mother side of the family meets the Chompret criteria for *TP53* mutation testing (*"individual with a tumor of LFS spectrum before age 46 years and at least one first-degree or second-degree relative with any of the LFS spectrum tumors before age 56 years"*) (99,100). Sequencing of gDNA of the *MetPC43* patient with the TruSight Cancer panel, which encompasses the *TP53* gene as well, confirmed tumor findings regarding the germline nature of the *BRCA2* c.3680_3681del variant and did not find any evidence of the involvement of *TP53* in his prostate cancer development. Li-Fraumeni-Like syndrome has a very similar phenotype to Li-Fraumeni syndrome, but with a less restrictive definition. In those cases, a mutation in the *TP53* gene is found in about 40% of the families (101), with *BRCA2* mutation being associated with LFL (102). According to NCCN Guidelines (100), there is no specific management for LFL families, only for *BRCA2* carriers, with special attention to breast cancer early detection and prevention in women, and breast and prostate cancer early detection in men. Besides prostate cancer, *BRCA2* mutations confer an increased risk to breast and ovarian cancer (103), pancreatic cancer (104,105) and melanoma (24). In this particular family, it is not possible to analyze segregation with cancer as all affected relatives are deceased, but pre-symptomatic genetic testing will be offered to all interested relatives. This is also the case for the carrier of the *BRCA2* variant c.4814del (MetPC30), with no affected relatives.

Regarding *CHEK2*, studies have shown that the prevalence of germline *CHEK2* mutations in PrCa families is comprised between 1.6% and 3.3% (28,92). *CHEK2* is also related with increased risk of breast cancer (106) and colon cancer (107,108). In families positive for *CHEK2* mutations, it is recommended that women starts annual screening with mammography at age of 40 years (100) and, in families with colon cancer, screening with colonoscopy starting at age of 50 years (109), or 10 years before the youngest age of diagnosis in the family. The specific variant c.349A>G was associated with an increased risk for breast cancer in a multiconsortia collaboration study (110), and is found both in Portuguese families with PrCa aggregation (92) and in HBOC families without PrCa cases (111). Patient MetPC27 has history of both breast and colon cancers in his mother, in which segregation for the *CHEK2* mutation would be informative for familial risk assessment. Patient MetPC45 has a paternal cousin diagnosed with prostate cancer and two brothers affected with cancer, one with melanoma and the other unspecified. In light of the, now identified, carrier status for germline deleterious variants in *CHEK2*, both families will be offered genetic testing, both for affected and non-affected family members that wish to be screened.

Although variants of unknown significance have, currently, no clinical implications, the NCCN Guidelines for Prostate Cancer recommends that families with germline VUS and with familial history of cancer be forwarded to genetic counseling in order to evaluate the possibility to participate in family and reclassification studies (16). Tumor sequencing may provide evidence for loss of heterozigozity in tumor suppressor genes by showing significantly higher VAFs (112), as can be seen in patient MetPC43 with a germline deleterious *BRCA2* mutation with VAF 0.68, but none of the VUS identified in this study showed VAFs signficantly higher than 50%. Of the five patients carrying germline VUS, only one variant in *MRE11A* is clearly biologically deleterious, leading to a stop codon at aminoacid 460. The *MRE11A* gene encodes a protein that is part of the MRN complex, which plays a key role in DNA double strand break repair, meiotic recombination, cell cycle checkpoints and maintenance of telomeres (113). This gene is associated with a ataxia-telangiectasia-like disorder, being a gene with moderate to low penetrance (114,115), thus not included in the NCCN Guidelines recommendations for management of relatives follow-up due to lack of evidence (100). Family history of this patient (Appendix 4) shows pancreatic, colon and gastric cancers, which are features of Lynch syndrome (LS). The correlation between *MRE11A* mutations and Lynch syndrome remains controversial, with few reports suggesting that mutations in *MRE11A* can cause microsatellite instability (116,117). Nevertheless, evidence is scarce and therefore there is no recommendation for management or follow-up for patients and healthy relatives with *MRE11A* mutations in the NCCN Guidelines for colorectal cancer either (109). Unfortunately, there are no affected relatives alive to perform segregation analysis. In any case, all families with VUS should be followed according to the family history of cancer and not according to the variant.

CONCLUSION AND FUTURE PERSPECTIVES

Conclusion

Next generation sequencing of tumor samples is indeed a good approach for screening patients that can benefit from PARPi treatment. It allows not only to select patients for targeted therapy, but also to identify families that are at increased risk for cancer development and can benefit from genetic counseling and pre-symptomatic germline testing. As NGS is typically a multigene approach, this method could also be used for screening patients that can benefit from other targeted therapies, such as immunotherapy with pembrolizumab for patients with mismatch repair (MMR) defects, as well as to identify carriers of variants in genes predisposing to hereditary prostate cancer, such as *HOXB13* (118,119).

Formalin-fixed paraffin-embedded (FFPE) samples have technical limitations, such as degraded DNA and sequencing artifacts, but with validated and well stablished pre-analytical quality check points and bioinformatics parameters, these difficulties can be overcome.

Despite the technical limitations of the NGS test at the time of writing of this dissertation, we could observe that our data is in accordance with the literature and a meaningful percentage of the patients referenced at IPO Porto can benefit from Olaparib treatment, which, besides having less side effects than conventional chemotherapy, has the potential to increase overall survival with better quality of life.

Future perspectives

To validate for routine diagnosis use, the next version of the NGS panel, which will have full coverage of the *BRCA1* gene, we will perform a new run with less samples in a high output flowcell, aiming to increase the sensitivity and validate the limit of detection. Additionally, we will attempt to perform segregation studies whenever possible in the families with VUS, as well as immunohistochemistry analyses of the MMR proteins PMS2 and MLH1 in the patients with variants in the respective genes, in order to obtain more data for variant classification.

In a near future, we will test NGS of cfDNA from peripheral blood samples for the simultaneous detection of both germline and somatic mutations, and the findings will be compared to those obtained in tumor and leucocyte DNA testing of the same patients.

Cell-free DNA will allow a step further in predictive genetic testing and disease monitoring. Being an easily accessible and minimally invasive source of tumor DNA, it has the potential to, more accurately and in real-time, reflect the genetic landscape of the tumor and their metastases. Targeted therapy is indicated as a second line treatment, after patients' relapse from first line treatment. The analysis of tumor biopsy performed at diagnosis is useful, but it may not reflect the exact genetic landscape at the time of analysis. The systemic treatments that the patients are submitted to induce a selective pressure on tumor subclones and can cause a change in the genetic landscape as compared to the primary tumor. Analysis of cell-free DNA samples throughout the patients' treatment journey holds the promise to allow characterization of the genetic profile over time and to adjust treatment as necessary.

REFERENCES

1. Hall JE, Guyton AC. Reproductive and Hormonal Functions of the Male (and Function of the Pineal Gland). In: Guyton and Hall Textbook of Medical Physiology. 13 ed. Elsevier; 2016. p. 1021–35.
2. Junqueira LC, Carneiro J. Aparelho Reprodutor Masculino. In: Histologia Básica - Texto e Atlas. 12 ed. Guanabara Koogan; 2013. p. 411–26.
3. Epstein JI, Netto GJ. Biopsy Interpretation of the Prostate. 5th Editio. Vol. 207, Biopsy Interpretation Series. Philadelphia, PA: Wolters Kluwer Health; 2015.
4. Epstein JI, Lotan TL. The Lower Urinary Tract and Male Genital System. In: Kumar V, Abbas AK, Aster JC, editors. Robbins and Cotran - Pathologic Basis of Disease. 9th ed. Elsevier Saunders; 2015. p. 959–90.
5. Bray F, Ferlay J, Soerjomataram I, Siegel RL, Torre LA, Jemal A. Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. CA Cancer J Clin. 2018 Nov;68(6):394–424.
6. Rawla P. Epidemiology of Prostate Cancer. World J Oncol. 2019;10(2):63–89.
7. Ferlay J, Colombet M, Soerjomataram I, Mathers C, Parkin DM, Piñeros M, et al. Estimating the global cancer incidence and mortality in 2018: GLOBOCAN sources and methods. Int J Cancer. 2019 Apr 15;144(8):1941–53.
8. Ferlay J, Colombet M, Soerjomataram I, Dyba T, Randi G, Bettio M, et al. Cancer incidence and mortality patterns in Europe: Estimates for 40 countries and 25 major cancers in 2018. Eur J Cancer. 2018;103:356–87.
9. PDQ® Screening and Prevention Editorial Board. PDQ Prostate Cancer Prevention [Internet]. Bethesda, MD: National Cancer Institute. 2020 [cited 2020 Feb 19]. Available from: https://www.cancer.gov/types/prostate/hp/prostate-prevention-pdq
10. Shen MM, Abate-Shen C. Molecular genetics of prostate cancer: New prospects for old challenges. Vol. 24, Genes and Development. Cold Spring Harbor Laboratory Press; 2010. p. 1967–2000.
11. Wang G, Zhao D, Spring DJ, Depinho RA. Genetics and biology of prostate cancer. Vol. 32, Genes and Development. Cold Spring Harbor Laboratory Press; 2018. p. 1105–40.
12. Perdana NR, Mochtar CA, Umbas R, Rizal A, Hamid AH. The Risk Factors of

Prostate Cancer and Its Prevention: A Literature Review The risk factors of prostate cancer and its prevention. 2016.

13. Carter BS, Bova GS, Beaty TH, Steinberg GD, Childs B, Isaacs WB, et al. Hereditary prostate cancer: Epidemiologic and clinical features. J Urol. 1993;150(3):797–802.

14. Brandt A, Bermejo JL, Sundquist J, Hemminki K. Age at Diagnosis and Age at Death in Familial Prostate Cancer. Oncologist. 2009 Dec 25;14(12):1209–17.

15. Bruner DW, Moore D, Parlanti A, Dorgan J, Engstrom P. Relative risk of prostate cancer for men with affected relatives: Systematic review and meta-analysis. Int J Cancer. 2003 Dec 10;107(5):797–803.

16. National Comprehensive Cancer Network. NCCN Clinical Practice Guidelines in Oncology - Prostate Cancer Early Detection, Version 2.2020. 2020.

17. Parker C, Gillessen S, Heidenreich A, Horwich A, Committee on behalf of the EG. Cancer of the prostate: ESMO Clinical Practice Guidelines for diagnosis, treatment and follow-up. Ann Oncol. 2015 Sep;26(suppl 5):v69–77.

18. Litwin MS, Tan HJ. The diagnosis and treatment of prostate cancer: A review. Vol. 317, JAMA - Journal of the American Medical Association. American Medical Association; 2017. p. 2532–42.

19. Epstein JI, Egevad L, Amin MB, Delahunt B, Srigley JR, Humphrey PA. The 2014 International Society of Urological Pathology (ISUP) Consensus Conference on Gleason Grading of Prostatic Carcinoma. Am J Surg Pathol. 2015;40(2):1.

20. Epstein JI, Amin MB, Reuter VE, Humphrey PA. Contemporary gleason grading of prostatic carcinoma. Am J Surg Pathol. 2017 Apr;41(4):e1–7.

21. The Cancer Genome Atlas Research Network T. The Molecular Taxonomy of Primary Prostate Cancer. Cell. 2015;163(4):1011–25.

22. Robinson D, Van Allen EM, Wu YM, Schultz N, Lonigro RJ, Mosquera JM, et al. Integrative clinical genomics of advanced prostate cancer. Cell. 2015 May 30;161(5):1215–28.

23. Mersch J, Jackson MA, Park M, Nebgen D, Peterson SK, Singletary C, et al. Cancers associated with BRCA1 and BRCA2 mutations other than breast and ovarian. Cancer. 2015 Jan 15;121(2):269–75.

24. Moran A, O'Hara C, Khan S, Shack L, Woodward E, Maher ER, et al. Risk of cancer other than breast or ovarian in individuals with BRCA1 and BRCA2

mutations. Fam Cancer. 2012 Jun;11(2):235–42.

25. Castro E, Goh C, Olmos D, Saunders E, Leongamornlert D, Tymrakiewicz M, et al. Germline BRCA mutations are associated with higher risk of nodal involvement, distant metastasis, and poor survival outcomes in prostate cancer. J Clin Oncol. 2013 May 10;31(14):1748–57.

26. Castro E, Goh C, Leongamornlert D, Saunders E, Tymrakiewicz M, Dadaev T, et al. Effect of BRCA Mutations on Metastatic Relapse and Cause-specific Survival after Radical Treatment for Localised Prostate Cancer. Eur Urol. 2015 Aug 1;68(2):186–93.

27. Castro E, Romero-Laorden N, Del Pozo A, Lozano R, Medina A, Puente J, 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. J Clin Oncol. 2019 Feb 20;37(6):490–503.

28. Pritchard CC, Mateo J, Walsh MF, De Sarkar N, Abida W, Beltran H, et al. Inherited DNA-Repair Gene Mutations in Men with Metastatic Prostate Cancer. N Engl J Med. 2016 Aug 4;375(5):443–53.

29. Mateo J, Carreira S, Sandhu S, Miranda S, Mossop H, Perez-Lopez R, et al. DNA-Repair Defects and Olaparib in Metastatic Prostate Cancer. N Engl J Med. 2015 Oct 29;373(18):1697–708.

30. Haraldsdottir S, Hampel H, Wei L, Wu C, Frankel W, Bekaii-Saab T, et al. Prostate cancer incidence in males with Lynch syndrome. Genet Med. 2014;16(7):553–7.

31. Ryan S, Jenkins MA, Win AK. Risk of prostate cancer in lynch syndrome: A systematic review and meta-Analysis. Vol. 23, Cancer Epidemiology Biomarkers and Prevention. American Association for Cancer Research Inc.; 2014. p. 437–49.

32. Helgstrand JT, Røder MA, Klemann N, Toft BG, Lichtensztajn DY, Brooks JD, et al. Trends in incidence and 5-year mortality in men with newly diagnosed, metastatic prostate cancer—A population-based analysis of 2 national cohorts. Cancer. 2018;124(14):2931–8.

33. Al Olama AA, Kote-Jarai Z, Schumacher FR, Wiklund F, Berndt SI, Benlloch S, et al. A meta-analysis of genome-wide association studies to identify prostate cancer susceptibility loci associated with aggressive and non-aggressive disease. Hum Mol Genet. 2013 Jan;22(2):408–15.

34. Davar D, Beumer JH, Hamieh L, Tawbi H. Role of PARP Inhibitors in Cancer Biology and Therapy. Curr Med Chem. 2012;19(23):3907–21.

35. Lord CJ, Ashworth A. PARP inhibitors: Synthetic lethality in the clinic. Vol. 355, Science. American Association for the Advancement of Science; 2017. p. 1152–8.

36. Bryant HE, Schultz N, Thomas HD, Parker KM, Flower D, Lopez E, et al. Specific killing of BRCA2-deficient tumours with inhibitors of poly(ADP-ribose) polymerase. Nature. 2005 Apr 14;434(7035):913–7.

37. Farmer H, McCabe H, Lord CJ, Tutt AHJ, Johnson DA, Richardson TB, et al. Targeting the DNA repair defect in BRCA mutant cells as a therapeutic strategy. Nature. 2005 Apr 14;434(7035):917–21.

38. Robson M, Im S-A, Senkus E, Xu B, Domchek SM, Masuda N, et al. Olaparib for Metastatic Breast Cancer in Patients with a Germline BRCA Mutation. N Engl J Med. 2017 Aug 10;377(6):523–33.

39. Mateo J, Porta N, Bianchini D, McGovern U, Elliott T, Jones R, et al. Olaparib in patients with metastatic castration-resistant prostate cancer with DNA repair gene aberrations (TOPARP-B): a multicentre, open-label, randomised, phase 2 trial. Lancet Oncol. 2020 Jan 1;21(1):162–74.

40. De Bono J, Mateo J, Fizazi K, Saad F, Shore N, Sandhu S, et al. Olaparib for Metastatic Castration-Resistant Prostate Cancer. N Engl J Med. 2020 Apr 28;382(22):NEJMoa1911440.

41. Hussain M, Mateo J, Fizazi K, Saad F, Shore ND, Sandhu S, et al. PROfound: Phase III study of olaparib versus enzalutamide or abiraterone for metastatic castration-resistant prostate cancer (mCRPC) with homologous recombination repair (HRR) gene alterations - ScienceDirect. ESMO 2019 - Genitourinary Tumors, Prostate. 2019.

42. FDA approves olaparib for HRR gene-mutated metastatic castration-resistant prostate cancer | FDA [Internet]. [cited 2020 Sep 29]. Available from: https://www.fda.gov/drugs/drug-approvals-and-databases/fda-approves-olaparib-hrr-gene-mutated-metastatic-castration-resistant-prostate-cancer

43. CHMP. Committee for Medicinal Products for Human Use (CHMP) [Internet]. 2020 [cited 2020 Sep 30]. Available from: www.ema.europa.eu/contact

44. Abida W, Campbell D, Patnaik A, Sautois B, Shapiro J, Vogelzang NJ, et al. Preliminary results from the TRITON2 study of rucaparib in patients (pts) with

DNA damage repair (DDR)-deficient metastatic castration-resistant prostate cancer (mCRPC): updated analyses - OncologyPRO. 2019;

45. Abida W, Campbell D, Patnaik A, Shapiro JD, Sautois B, Vogelzang NJ, 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. Clin Cancer Res. 2020 Jun 1;26(11):2487–96.

46. FDA grants accelerated approval to rucaparib for BRCA-mutated metastatic castration-resistant prostate cancer | FDA [Internet]. [cited 2020 Oct 8]. Available from: https://www.fda.gov/drugs/fda-grants-accelerated-approval-rucaparib-brca-mutated-metastatic-castration-resistant-prostate

47. Wan JCM, Massie C, Garcia-Corbacho J, Mouliere F, Brenton JD, Caldas C, et al. Liquid biopsies come of age: Towards implementation of circulating tumour DNA. Vol. 17, Nature Reviews Cancer. Nature Publishing Group; 2017. p. 223–38.

48. Siravegna G, Marsoni S, Siena S, Bardelli A. Integrating liquid biopsies into the management of cancer. Nat Rev Clin Oncol. 2017 Sep;14(9):531–48.

49. Siravegna G, Mussolin B, Venesio T, Marsoni S, Seoane J, Dive C, et al. How liquid biopsies can change clinical practice in oncology. Ann Oncol Off J Eur Soc Med Oncol. 2019 Aug 2;

50. Antonarakis ES, Lu C, Wang H, Luber B, Nakazawa M, Roeser JC, et al. AR-V7 and Resistance to Enzalutamide and Abiraterone in Prostate Cancer. N Engl J Med. 2014 Sep 11;371(11):1028–38.

51. Scher HI, Graf RP, Schreiber NA, Jayaram A, Winquist E, McLaughlin B, et al. Assessment of the validity of nuclear-localized androgen receptor splice variant 7 in circulating tumor cells as a predictive biomarker for castration-resistant prostate cancer. JAMA Oncol. 2018 Sep 1;4(9):1179–86.

52. Decipher Biosciences I. Decipher Identifies Which Patients are Likely to Recur When Treated With Radiation Alone Decipher Identifies the Patients Who May Benefit from Treatment Intensification. 2019.

53. Brand TC, Zhang N, Crager MR, Maddala T, Dee A, Sesterhenn IA, et al. Patient-specific Meta-analysis of 2 Clinical Validation Studies to Predict Pathologic Outcomes in Prostate Cancer Using the 17-Gene Genomic Prostate Score. Urology. 2016 Mar 1;89:69–75.

54. Blume-Jensen P, Berman DM, Rimm DL, Shipitsin M, Putzi M, Nifong TP, et al. Biology of Human Tumors Development and clinical validation of an in situ biopsy-based multimarker assay for risk stratification in prostate cancer. Clin Cancer Res. 2015 Jun 1;21(11):2591–600.

55. Tosoian JJ, Chappidi MR, Bishoff JT, Freedland SJ, Reid J, Brawer M, et al. Prognostic utility of biopsy-derived cell cycle progression score in patients with National Comprehensive Cancer Network low-risk prostate cancer undergoing radical prostatectomy: implications for treatment guidance. BJU Int. 2017 Dec 1;120(6):808–14.

56. Goodall J, Mateo J, Yuan W, Mossop H, Porta N, Miranda S, et al. Circulating cell-free DNA to guide prostate cancer treatment with PARP inhibition. Cancer Discov. 2017;

57. Sanger F, Nicklen S, Coulson AR. DNA sequencing with chain-terminating inhibitors. Proc Natl Acad Sci U S A. 1977;74(12):5463–7.

58. Collins FS, Morgan M, Patrinos A. The Human Genome Project: Lessons from large-scale biology. Vol. 300, Science. American Association for the Advancement of Science; 2003. p. 286–90.

59. Illumina I. An introduction to Next-Generation Sequencing Technology [Internet]. [cited 2020 Apr 27]. Available from: www.illumina.com/technology/next-generation-sequencing.html

60. Heather JM, Chain B. The sequence of sequencers: The history of sequencing DNA. Vol. 107, Genomics. Academic Press Inc.; 2016. p. 1–8.

61. Bentley DR, Balasubramanian S, Swerdlow HP, Smith GP, Milton J, Brown CG, et al. Accurate whole human genome sequencing using reversible terminator chemistry. Nature. 2008 Nov 6;456(7218):53–9.

62. Aziz N, Zhao Q, Bry L, Driscoll DK, Funke B, Gibson JS, et al. College of American pathologists' laboratory standards for next-generation sequencing clinical tests. Arch Pathol Lab Med. 2015 Apr 1;139(4):481–93.

63. Illumina I. NextSeq Series Specifications | Key performance parameters [Internet]. [cited 2020 Apr 27]. Available from: https://emea.illumina.com/systems/sequencing-platforms/nextseq/specifications.html

64. Data pre-processing for variant discovery – GATK [Internet]. [cited 2020 Sep 6]. Available from: https://gatk.broadinstitute.org/hc/en-us/articles/360035535912

65. Jennings LJ, Arcila ME, Corless C, Kamel-Reid S, Lubin IM, Pfeifer J, et al. Guidelines for Validation of Next-Generation Sequencince Based Oncology Panels A Joint Consensus Recommendation of the Association for Molecular Pathology and College of American Pathologists. J Mol Diagnostics. 2017;19(3).

66. Deans ZC, Costa JL, Cree I, Dequeker E, Edsjö A, Henderson S, et al. Integration of next-generation sequencing in clinical diagnostic molecular pathology laboratories for analysis of solid tumours: an expert opinion on behalf of IQN Path ASBL. Vol. 470, Virchows Archiv. Springer; 2017. p. 5–20.

67. Kalia SS, Adelman K, Bale SJ, Chung WK, Eng C, Evans JP, et al. Recommendations for reporting of secondary findings in clinical exome and genome sequencing, 2016 update (ACMG SF v2.0): A policy statement of the American College of Medical Genetics and Genomics. Genet Med. 2017 Feb 1;19(2):249–55.

68. Li MM, Datto M, Duncavage EJ, Kulkarni S, Lindeman NI, Roy S, et al. Standards and Guidelines for the Interpretation and Reporting of Sequence Variants in Cancer: A Joint Consensus Recommendation of the Association for Molecular Pathology, American Society of Clinical Oncology, and College of American Pathologists. Vol. 19, Journal of Molecular Diagnostics. Elsevier B.V.; 2017. p. 4–23.

69. Mateo J, Chakravarty D, Dienstmann R, Jezdic S, Gonzalez-Perez A, Lopez-Bigas N, et al. A framework to rank genomic alterations as targets for cancer precision medicine: The ESMO Scale for Clinical Actionability of molecular Targets (ESCAT). Ann Oncol. 2018 Sep 1;29(9):1895–902.

70. Chakravarty D, Gao J, Phillips S, Kundra R, Zhang H, Wang J, et al. OncoKB: A Precision Oncology Knowledge Base. JCO Precis Oncol. 2017 Nov 16;(1):1–16.

71. Richards S, Aziz N, Bale S, Bick D, Das S, Gastier-Foster J, et al. Standards and guidelines for the interpretation of sequence variants: A joint consensus recommendation of the American College of Medical Genetics and Genomics and the Association for Molecular Pathology. Genet Med. 2015 May 8;17(5):405–24.

72. Desvignes JP, Bartoli M, Delague V, Krahn M, Miltgen M, Béroud C, et al. VarAFT: A variant annotation and filtration system for human next generation sequencing data. Nucleic Acids Res. 2018 Jul 2;46(W1):W545–53.

73. About gnomAD | gnomAD [Internet]. [cited 2020 Sep 1]. Available from: https://gnomad.broadinstitute.org/about

74. Karczewski KJ, Francioli LC, Tiao G, Cummings BB, Alföldi J, Wang Q, et al. The mutational constraint spectrum quantified from variation in 141,456 humans. Nature. 2020 May 28;581(7809):434–43.

75. Data | 1000 Genomes [Internet]. [cited 2020 Sep 1]. Available from: https://www.internationalgenome.org/data

76. Auton A, Abecasis GR, Altshuler DM, Durbin RM, Bentley DR, Chakravarti A, et al. A global reference for human genetic variation. Vol. 526, Nature. Nature Publishing Group; 2015. p. 68–74.

77. Clarke L, Fairley S, Zheng-Bradley X, Streeter I, Perry E, Lowy E, et al. The international Genome sample resource (IGSR): A worldwide collection of genome variation incorporating the 1000 Genomes Project data. Nucleic Acids Res. 2017 Jan 1;45(D1):D854–9.

78. What is ClinVar? [Internet]. [cited 2020 Sep 1]. Available from: https://www.ncbi.nlm.nih.gov/clinvar/intro/

79. Landrum MJ, Lee JM, Benson M, Brown GR, Chao C, Chitipiralla S, et al. ClinVar: Improving access to variant interpretations and supporting evidence. Nucleic Acids Res. 2018 Jan 1;46(D1):D1062–7.

80. OMIM - Online Mendelian Inheritance in Man [Internet]. [cited 2020 Sep 1]. Available from: https://omim.org/

81. Tate JG, Bamford S, Jubb HC, Sondka Z, Beare DM, Bindal N, et al. COSMIC: The Catalogue Of Somatic Mutations In Cancer. Nucleic Acids Res. 2019 Jan 8;47(D1):D941–7.

82. COSMIC | Catalogue of Somatic Mutations in Cancer [Internet]. [cited 2020 Sep 1]. Available from: https://cancer.sanger.ac.uk/cosmic

83. cBioPortal for Cancer Genomics [Internet]. [cited 2020 Sep 1]. Available from: https://www.cbioportal.org/

84. Cerami E, Gao J, Dogrusoz U, Gross BE, Sumer SO, Aksoy BA, et al. The cBio Cancer Genomics Portal: An open platform for exploring multidimensional cancer genomics data. Cancer Discov. 2012 May;2(5):401–4.

85. Gao J, Aksoy BA, Dogrusoz U, Dresdner G, Gross B, Sumer SO, et al. Integrative analysis of complex cancer genomics and clinical profiles using the cBioPortal. Sci Signal. 2013 Apr 2;6(269):pl1.

86. OncoKB [Internet]. [cited 2020 Sep 1]. Available from: https://www.oncokb.org/

87. Home - ClinicalTrials.gov [Internet]. [cited 2020 Sep 1]. Available from:

https://clinicaltrials.gov/ct2/home

88. Home - My Cancer Genome [Internet]. [cited 2020 Sep 2]. Available from: https://www.mycancergenome.org/

89. Jain N, Mittendorf KF, Holt M, Lenoue-Newton M, Maurer I, Miller C, et al. The My Cancer Genome clinical trial data model and trial curation workflow. J Am Med Inform Assoc. 2020 Jul 1;27(7):1057–66.

90. Cancer Genome Interpreter - Identification of therapeutically actionable genomic alterations in tumors [Internet]. [cited 2020 Sep 1]. Available from: https://www.cancergenomeinterpreter.org/home

91. Tamborero D, Rubio-Perez C, Deu-Pons J, Schroeder MP, Vivancos A, Rovira A, et al. Cancer Genome Interpreter annotates the biological and clinical relevance of tumor alterations. Genome Med. 2018 Mar 28;10(1):25.

92. Paulo P, Maia S, Pinto C, Pinto P, Monteiro A, Peixoto A, et al. Targeted next generation sequencing identifies functionally deleterious germline mutations in novel genes in early-onset/familial prostate cancer. Eng C, editor. PLOS Genet. 2018 Apr 16;14(4):e1007355.

93. TruSight Hereditary Cancer Panel | Used with Nextera Flex for Enrichment [Internet]. [cited 2020 Sep 15]. Available from: https://emea.illumina.com/products/by-type/clinical-research-products/trusight-cancer-hereditary.html

94. Ye J, Coulouris G, Zaretskaya I, Cutcutache I, Rozen S, Madden TL. Primer-BLAST: a tool to design target-specific primers for polymerase chain reaction. BMC Bioinformatics. 2012;13:134.

95. Do H, Dobrovic A. Sequence Artifacts in DNA from Formalin-Fixed Tissues: Causes and Strategies for Minimization. Clin Chem. 2015 Jan 1;61(1):64–71.

96. Bono JS de, Mateo J, Fizazi K, Saad F, Shore N, Sandhu S, et al. Final overall survival (OS) analysis of PROfound: Olaparib vs physician's choice of enzalutamide or abiraterone in patients (pts) with metastatic castration-resistant prostate cancer (mCRPC) and homologous recombination repair (HRR) gene alterations - Abs. ESMO Virtual Congr 2020. 2020;

97. Brandão A, Paulo P, Maia S, Pinheiro M, Peixoto A, Cardoso M, et al. The CHEK2 variant c.349A>G is associated with prostate cancer risk and carriers share a common ancestor. Cancers (Basel). 2020;in press.

98. Clendenning M, Hampel H, LaJeunesse J, Lindblom A, Lockman J, Nilbert M, et

al. Long-range PCR facilitates the identification of PMS2-specific mutations. Hum Mutat. 2006 May;27(5):490–5.

99. Chompret A, Abel A, Stoppa-Lyonnet D, Brugières L, Pagès S, Feunteun J, et al. Sensitivity and predictive value of criteria for p53 germline mutation screening [1]. Vol. 38, Journal of Medical Genetics. BMJ Publishing Group; 2001. p. 43–7.

100. Pilarski R, Berry MP, Jude S, Buys SS, Dickson P, Domchek SM, et al. NCCN Guidelines Version 1.2021 Genetic/Familial High-Risk Assessment: Breast, Ovarian, and Pancreatic. 2020.

101. Bachinski LL, Olufemi S-E, Zhou X, Wu C-C, Yip L, Shete S, et al. Genetic Mapping of a Third Li-Fraumeni Syndrome Predisposition Locus to Human Chromosome 1q23. Cancer Res. 2005;65(2).

102. Evans DGR, Wu CL, Birch JM. BRCA2: A cause of Li-Fraumeni-like syndrome. Vol. 45, Journal of Medical Genetics. J Med Genet; 2008. p. 62–3.

103. Ford D, Easton DF, Stratton M, Narod S, Goldgar D, Devilee P, et al. Genetic heterogeneity and penetrance analysis of the BRCA1 and BRCA2 genes in breast cancer families. Am J Hum Genet. 1998;62(3):676–89.

104. Chaffee KG, Oberg AL, McWilliams RR, Majithia N, Allen BA, Kidd J, et al. Prevalence of germ-line mutations in cancer genes among pancreatic cancer patients with a positive family history. Genet Med. 2018 Jan 1;20(1):119–27.

105. Hu C, Hart SN, Polley EC, Gnanaolivu R, Shimelis H, Lee KY, et al. Association between inherited germline mutations in cancer predisposition genes and risk of pancreatic cancer. In: JAMA - Journal of the American Medical Association. American Medical Association; 2018. p. 2401–9.

106. Lu HM, Li S, Black MH, Lee S, Hoiness R, Wu S, et al. Association of Breast and Ovarian Cancers with Predisposition Genes Identified by Large-Scale Sequencing. JAMA Oncol. 2019 Jan 1;5(1):51–7.

107. Liu C, Wang QS, Wang YJ. The CHEK2 I157T variant and colorectal cancer susceptibility: A systematic review and meta-analysis. Asian Pacific J Cancer Prev. 2012;13(5):2051–5.

108. Xiang HP, Geng XP, Ge WW, Li H. Meta-analysis of CHEK2 1100delC variant and colorectal cancer susceptibility. Eur J Cancer. 2011 Nov;47(17):2546–51.

109. Gupta S, Weiss JM, Burke CA, Chen L-M, Chung DC, Clayback KM, et al. NCCN Guidelines Version 1.2020 Genetic/Familial High-Risk Assessment: Colorectal Continue. 2020.

110. Southey MC, Goldgar DE, Winqvist R, Pylkäs K, Couch F, Tischkowitz M, et al. PALB2, CHEK2 and ATM rare variants and cancer risk: Data from COGS. J Med Genet. 2016 Dec 1;53(12):800–11.

111. Pinto P, Paulo P, Santos C, Rocha P, Pinto C, Veiga I, et al. Implementation of next-generation sequencing for molecular diagnosis of hereditary breast and ovarian cancer highlights its genetic heterogeneity. Breast Cancer Res Treat. 2016 Sep 1;159(2):245–56.

112. Peixoto A, Pinto P, Guerra J, Pinheiro M, Santos C, Pinto C, et al. Tumor Testing for Somatic and Germline BRCA1/BRCA2 Variants in Ovarian Cancer Patients in the Context of Strong Founder Effects. Front Oncol. 2020 Jul 31;10:1318.

113. Williams GJ, Lees-Miller SP, Tainer JA. Mre11-Rad50-Nbs1 conformations and the control of sensing, signaling, and effector responses at DNA double-strand breaks. Vol. 9, DNA Repair. DNA Repair (Amst); 2010. p. 1299–306.

114. Apostolou P, Fostira F. Hereditary breast cancer: The Era of new susceptibility genes. Vol. 2013, BioMed Research International. Hindawi Limited; 2013.

115. Damiola F, Pertesi M, Oliver J, Le Calvez-Kelm F, Voegele C, Young EL, et al. Rare key functional domain missense substitutions in MRE11A, RAD50, and NBN contribute to breast cancer susceptibility: Results from a Breast Cancer Family Registry case-control mutation-screening study. Breast Cancer Res. 2014 Jun 3;16(3):R58.

116. Situ Y, Chung L, Lee CS, Ho V. MRN (MRE11-RAD50-NBS1) complex in human cancer and prognostic implications in colorectal cancer. Vol. 20, International Journal of Molecular Sciences. MDPI AG; 2019.

117. Alemayehu A, Fridrichova I. The MRE11/RAD50/NBS1 complex destabilization in Lynch-syndrome patients. Eur J Hum Genet. 2007 Sep 30;15(9):922–9.

118. Breyer JP, Avritt TG, McReynolds KM, Dupont WD, Smith JR. Confirmation of the HOXB13 G84E germline mutation in familial prostate cancer. Cancer Epidemiol Biomarkers Prev. 2012 Aug;21(8):1348–53.

119. Xu J, Lange EM, Lu L, Zheng SL, Wang Z, Thibodeau SN, et al. HOXB13 is a susceptibility gene for prostate cancer: Results from the International Consortium for Prostate Cancer Genetics (ICPCG). Hum Genet. 2013 Jan 12;132(1):5–14.

APPENDIX

Appendix 1

Table A 1. Tumor and blood samples obtained from each patient

Patient number	Age	Gleason	Tumor Sample	Origin	Blood sample
1	56	4+5	Yes	Biopsy	Yes
2	61	4+3	Yes	Biopsy	Yes
3	70	4+4	B1 B2	Biopsy	Yes
4	56	5+4	Yes	Biopsy	Yes
5	70		Yes	Biopsy	Yes
6	67	4+3	A B	Biopsy	Yes
7	80	4+5	Yes	Biopsy	Yes
8	84	4+5	Yes	Biopsy	Yes
9	64	4+5	Yes	Biopsy	Yes
10	62	3+4	Yes	Biopsy	Yes
11	69	-	No	-	Yes
12	84	4+4	No	-	Yes
13	55	4+3	Yes	Prostatectomy	Yes
14	66	4+4	Yes	Prostatectomy	Yes
15	72	-	Yes	Prostatectomy	Yes
16	71	4+3	Yes	Biopsy	Yes
17	68	4+4	Yes	Biopsy	Yes
18	70	4+4	Yes	Prostatectomy	Yes
19	68	4+4	Yes	Biopsy	Yes
20	47	4+3	E1 E2	Biopsy	Yes
21	64	4+5	Yes	Biopsy	Yes
22	82	4+3	1 2	Biopsy	Yes
23	65	5+4	Yes	Prostatectomy	Yes
24	76	8	No	-	Yes
25	69	-	1 2	Biopsy	Yes
26	61	4+4	Yes	Biopsy	Yes
27	67	4+5	1 2	Biopsy	Yes
28	77	4+5	1 2	Biopsy	Yes
29	68	4+4	E1 E2	Biopsy	Yes
30	67	4+3	1 2	Biopsy	Yes
31	76	5+5	Yes	Biopsy	No*
32	66	8	1 2	Biopsy	Yes
33	77	4+5	1 2	Biopsy	Yes
34	72	3+4	No	-	Yes
35	67	4+4	Yes	Biopsy	Yes
36	73	4+4	1 2	Biopsy	Yes
37	68	3+4	D E	Biopsy	Yes
38	76	4+4	Yes	Biopsy	Yes
39	71	9	Yes	Biopsy	Yes
40	62	3+4	Yes	Prostatectomy	No**
41	59	4+5	Yes	Biopsy	Yes
42	52	4+4	Yes	Biopsy	Yes
43	65	5+5	Yes	Biopsy	Yes
44	67	4+5	Yes	Prostatectomy	Yes
45	65	4+4	D E	Biopsy	Yes
46	74	-	Yes	Biopsy	Yes
47	70	4+5	Yes	Biopsy	Yes

** buccal swab. ** buffy coat pellet discarded*

Appendix 2

Table A 2. Control samples results – comparation of the three software

Sample number	Gene	Transcript	Previous result	Consequence	ClinVar	Position (hg19)	NextGene		Dragen		SureCall	
							Depth	VAF	Depth	VAF	Depth	VAF
S1	*BRCA2*	NM_000059.3	c.6351_6377del p.V2118_C2126del	Frameshift	.	chr13:32914839	89 TGTAAACTCAGAAATGGAAAAAACCTG	0.146	31 ACTGTGTAAACTCAGAAATGGAAAAAAC	0.233	NA	NA
S2	*PALB2*	NM_024675.3	c.2401G>A p.D801N	Missense	VUS	chr16: 23641074	2968	0.424	477	0.424	523	0.421
S3	*BRCA2*	NM_000059	c.156_157insAlu	Coding sequence	Pathogenic	chr13:32893302	NA	NA	175 ATGCCGGGCGCGGTGGCTCACGCCTG	0.252	NA	NA
S4	*RAD51C*	NM_058216.1	c.709C>T p.R237X	Stop gained	Pathogenic	chr17: 56787223	925	0.741	315	0.756	328	0.753
S5	*MSH6*	NM_000179.2	c.3848_3862del p.I1283_Y1287del	Inframe deletion	.	chr2:48033634	1021 CTATTACGTTCCTCTATA	0.515	463 CTATTACGTTCCTCTA	0.387	293 CTATTACGTTCCTCTA	0.307
			c.4001+1G>T	Splice donor	Likely Pathogenic	chr2:48033791	656	0.337	236	0.322	232	0.323
S6	*ARID1A*	NM_006015.4	c.4003C>T	Stop gained	Pathogenic	chr1:27100207	6441	0.279	1347	0.309	1352	0.308
			c.6704_6729del	Frameshift	.	chr1:27107091	9065 CCCGCGCGCTGCTTGCCTTGGCCAAG	16,92	2019 TGCCCGCGCGCTGCTTGCCTTGGCCAA	0.176	NA	NA
S7	*PTEN*	NM_000314.4	c.15_16insCTTC	5'UTR	.	chr10: 89624239	1405	49,61	247 ATCCT	0.615	250 ATCCT	0.468
S8	*PTEN CNV*	NM_000314.4	c.493-1884_634+1938dup	-	-	-	-	-	-	-	-	-
S9	*RECQL4*	NM_004260.3	c.2297delC p.P766Rfs	Frameshift	.	chr8:145738767	5828	0.887	1392	1.000	1378	0.916
	TP53	NM_001126112.2	c.673-1G>A	Splice acceptor	Pathogenic	chr17:7577609	1900	0.637	781	0.638	782	0.651
S10	*MSH2*	NM_000251.2	c.942+3A>T	Splice region	Pathogenic	chr2:47641560	NA	NA	27	0.375	NA	NA

Appendix 3

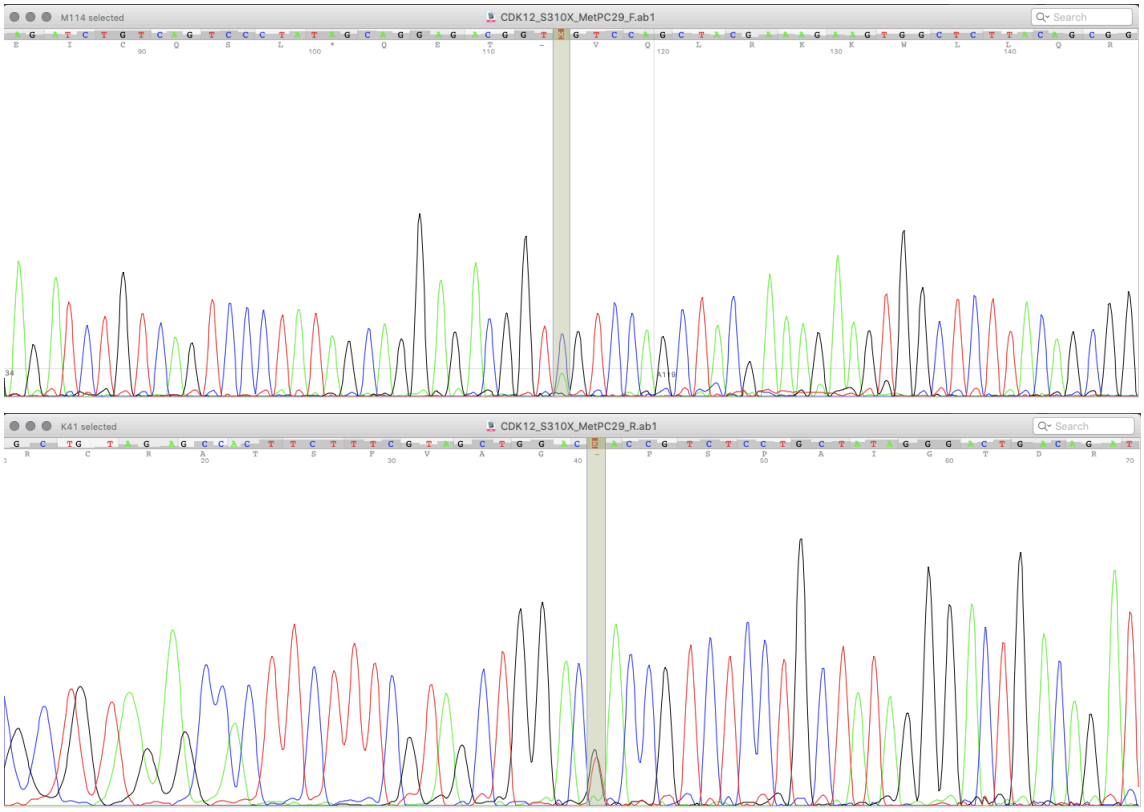


Figure A 1. MetPC29 CDK12 variant confirmation in the tumor sample. Electropherograms showing the forward (above) and reverse (below) sequences, with the alteration marked. This variant was found only in tumor sample (bottom lane), therefore is somatic.

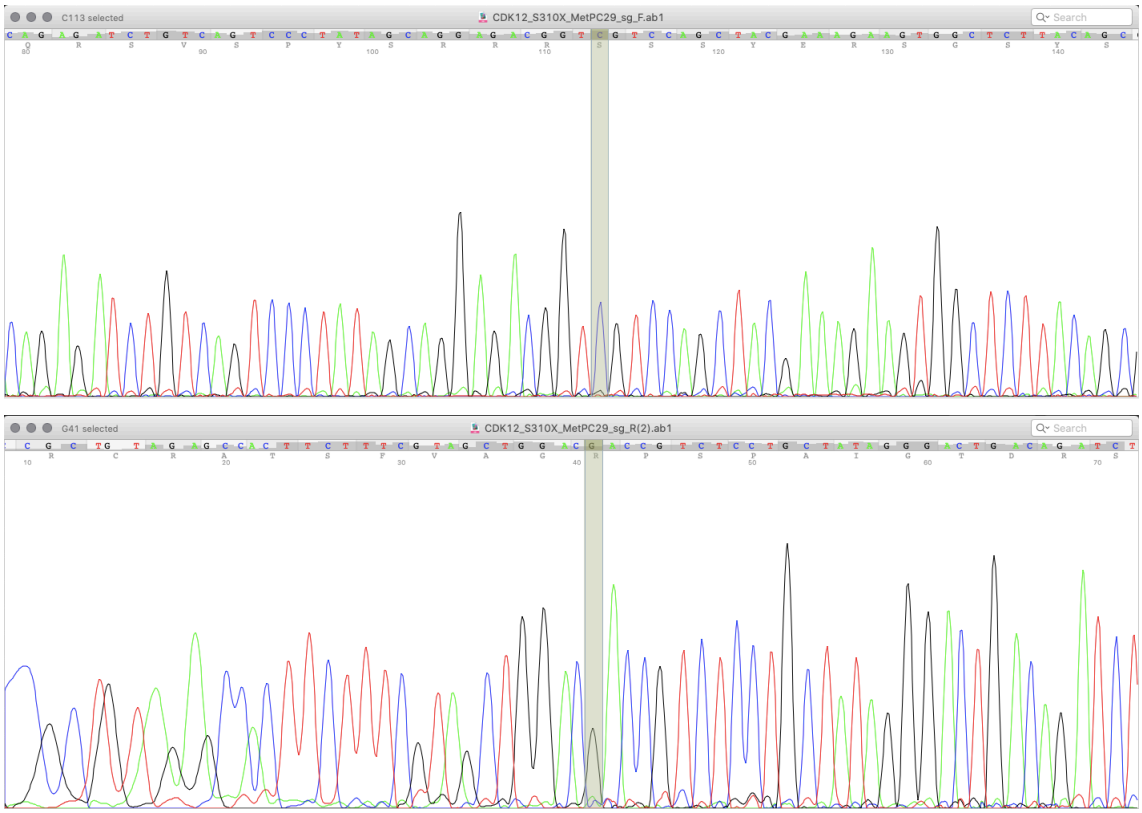


Figure A 2. MetPC29 CDK12 variant confirmation – "buffy coat" pellet sample (germline DNA). Electropherograms showing the forward (above) and reverse (below) sequences, with the alteration site marked, showing no variant. This variant was found only in tumor sample, therefore is somatic.

Appendix 4

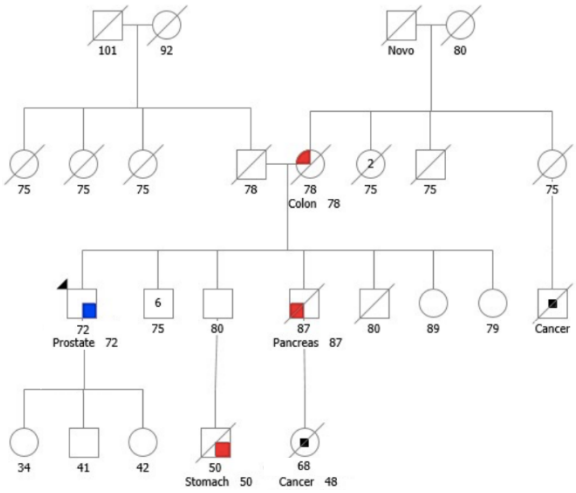


Figure A 3. Pedigree of patient MetPC36, with the MRE11A deleterious variant. Index with an arrow.