DEPARTMENT OF BIOMEDICAL SCIENCES

PhD COURSE IN LIFE SCIENCES ANDBIOTECHNOLOGIES

UNIVERSITY OF SASSARI

Coordinator: Prof. Sechi Leonardo Antonio

Genomic Landscape of Local Prostate Cancer in Sardinia Population

by

Tiansheng Zeng

Supervisors: Prof. Leonardo.A.Sechi

Department of Biomedical Sciences University of Sassari

Prof. David J. Kelvin

Division of Immunology, International Institute of Infection and Immunity

Shantou University Medical College

Index

Abstract..1

Chapter Ⅰ.Introduction 2

1. Incidence, mortality of prostate cancer worldwide 2

2. Risk factors of prostate cancer 3

2.1 Race and ethnicity 3

2.2 PSA Screening 4

2.3Environmental factors 5

2.4 Age 7

2.5Infectious diseases 7

2.6 Family history 9

3. Heritability of prostate cancer 10

3.1Loci with low penetration explain 19% of the hereditary of prostate cancer 10

3.2 Rare mutations on susceptible genes explain 15% of the hereditary of prostate cancer 11

3.2.1 HOXB13 12

3.2.2 BRCA1/BRCA2..13

3.2.3 DNA mismatch repair genes 15

3.2.4 Other prostate cancer germline risk variants 15

4. Somatic Genomic alterations in prostate cancer 16

5. Molecular evidence for racial disparities of prostate cancer 18

6. Oncogenesis and development of tumors 19

7. Prostate cancer diagnosis 22

7.1 PSA screening...22

7.2 Biopsy 23

7.3 Risk stratification...24

7.4 Development of prostate cancer diagnosis and risk stratification 25

7.4.1 PSA derived 25

7.4.1.1PHI index 26

7.4.1.2 The 4Kscore..27

7.4.2 Urine-derived biomarkers 28

7.4.2.1 SelectMDx 28

7.4.2.2 ExoDx 28

7.4.3 Genic Tests...29

8. Treatment of prostate cancer...30

8.1 Active monitoring for very-low risk prostate cancer.............................30

8.2 Activemonitoringfor low risk prostate cancer 31

8.3 Treatment for patients with low risk cancer 31

8.4 Treatment for locally advanced prostate cancer 31

8.5CRPC treatment 32

8.5.1 Second Generation Androgen Receptor Antagonists 32

8.5.2 PAPR inhibitor 33

8.5.3 Siupleucel-T 33

8.5.4 Chemotherapy 33

8.5.5 Ra223mCRPC 34

9. Translate genetics to biology and therapeutics 34

9.1 AR signaling 36

9.1.1 GnRH 36

9.1.2 AR 37

9.1.2.1 AR-LBD 37

9.1.2.2 AR-nonLBD 38

9.1.3 AR-binding protein 39

9.1.4 CYP17A1 40

9.1.5 5alpha reductase 40

9.2 DNA repair Defects 42

9.3 PIK3-AKT pathway 43

9.4 ETS gene rearrangements 45

9.5 TP53 47

9.6 WNT signaling 48

9.7 The RAS-RAF-MEK signaling 49

10. Objectives 50

ChapterⅡ. Material and methods 52

1. Patients and Samples..52

2. DNA extraction from FFPET tissue 53

3. Whole Exome Sequencing Library Preparation 53

4. Sequence data Quality control. 54

5. Read mapping and processing. 54

6. Somatic SNP and INDEL calling and annotation..................................55

7. Identification of mutation drivers 55

8. Germline SNVs and INDELs calling and candidate germline risk identification 56

9. Fusion calling, filtering, ORF prediction and visualization....................56

10. Copy number variation calling, filtering and driver copy number variation identification 57

11. Integrative analysis of SNP, INDEL, fusion and copy number variation 58

12. Differential gene expression and genomic alterations in prostate cancer. 58

13. PCR-based ERG fusion detection..58

Chapter Ⅲ. Results: Genomic landscape of Localized Prostate Cancer in a Sardinian cohort...60

1. General summary of clinical and sequencing parameters.......................60

2. Somatic driver mutations in Sardinia prostate cancer.............................62

3. Somatic copy number variation in Sardinia prostate cancer....................71

4. Comparison of gene mutations between prostate cancer in the TCGAdatabase and Sardinia cohort..78

5. Novel BTBD7-SLC2A5 fusions and ETS family status in prostate cancer of Sardinia..85

6. Germline risk mutations in the patients 88

ChapterⅤ. Discussion 93

Conclusion and prospect 109

A brief summary of the Chinese prostate cancer study 110

Data and material accession numbers..113

Abbreviations 113

Acknowledgment...115

Reference 116

Abstract

Race and ethnicity are risk factors for prostate cancer. In the United States, African American men have the highest rate of mortality followed by Caucasians, and Asian Americans. The effects of race and ethnicity on prostate cancer are also reflected in different frequencies of ETS family fusion in different groups. ETS family fusions is the most common alteration in prostate cancer of Caucasian men at a frequency of ~50%, however, they are lower in African Americans and Chinese at 20-30%. Most of the genomic prostate cancer studies are focused on cohorts of European ancestry, leaving minority groups underrepresentation. Furthermore, in racial mixing, the ethnic contribution to risk is unclear. Sardinia population is an isolated Mediterranean population, and a purported refuge population of Neolithic ancestry with much lower incidence of prostate cancer than that in mainland Europe. Here, we conducted a genomic prostate cancer genomic study on a Sardinia cohort diagnosed with local prostate cancer. We identified a novel germline risk mutation ARSD-G320D occurring in 53 percent of the patients, somatic UGT family amplifications which occurred in 20% the patients, a novel in-frame fusion BTBD7-SLC2A5 occurred in 12 % of the patients. In addition, we pointed out that IRF8 deletion at 16q24.2 is a candidate driver in prostate cancer and patients with IRF8 deletion have worse prognosis. Our data revealed similarities and disparities in genomic alterations of prostate cancer between Sardinians and other ethnic groups. As well we have conducted a study based

on Chinese prostate cancer cohort and have seen greater molecular disparities from TCGA cohort than in the Sardinian prostate cancer cohort. In Chinese cohort we have identified 37 genes significantly mutated and 20 of them have not implicated in prostate cancer in Caucasian and reveals a set of genomic markers that may inform the ethnic disparities.

Chapter Ⅰ.Introduction

1. Incidence, mortality of prostate cancer worldwide

Prostate cancer is the most common and fifth fatal cancer in men worldwide[1]. In 2012, an estimated of 1.1 million people were diagnosed with prostate cancer, and 307,000 deaths[2]. The incidence and mortality of prostate cancer vary greatly in the world [2-4].

The highest incidence of prostate cancer is observed in Oceania (111.6/100,000), followed by North America (97.2/100,000), Western Europe (94.9/100,000), Nordic Europe (85.0/100,000) and the Caribbean (79.8/100,000), while the incidence is much lower in Southeast Asia (11.2/100,000), North Africa (10.6/100,000), East Asia (10.5/100,000) and South-Central Asia (4/100,000) [2].

The highest prostate mortality is observed in the Caribbean (29.3/100,000), followed by South Africa and Central Africa (24.4/100,000, 24.2/100,000). As the incidence in Caribbean as high as 79.8/100,000, the ratio of mortality/incidence is 37%, much lower than that in South and Central Africa which is up to 90% [2, 4]. Although both the incidence and mortality in Southeast Asia, Central and South Asia are low, the ratio of mortality/incidence is as high as 64%[4]. Nevertheless, almost 70% of the newly diagnosed cases with prostate cancer in the world are in more developed regions such as North America, Oceania and Northern and Western Europe, the ratio of mortality/incidence in those area is only between 10-18% [2, 4] and the prostate cancer mortality rate has been decreasing over time[5, 6].

2. Risk factors of prostate cancer

2.1 Race and ethnicity

The incidence and mortality of prostate cancer vary greatly by geographic regions, strikingly, in the United States, the incidence and mortality of prostate cancer varies considerably by race and ethnicity [7, 8]. The prostate cancer incidence of the North America Africans is up to 208.7 per 100,000 and the mortality is up to 47.2 per 100,000, while prostate cancer incidence in the Asian American, Native Hawaiians and Pacific Islander (AANHIP) is only 67.8/100000 which is almost one third of North America Africans, and one half of Non-Hispanic Whites (123/100000) [8]. The

incidence, the mortality, and the ratio of mortality to incidence is lowest in AANHIP rather than North America Africans and Non-Hispanic Whites [8]. Even though, disparities in the diagnosis, treatment, and survival of prostate cancer patients of different races are often attributed to socio-economic status and access to healthcare [9, 10], after adjusting for those effects, racial disparities in prostate cancer incidence and mortality rates in United State remain significant [11]. In the United States, a white person has a 16% lifetime risk of prostate cancer and a 2.5% chance of death from prostate cancer, while black people are 70% more likely to develop prostate cancer and 40% more likely to die [12].

2.2 PSA Screening

In 1970 Wang and Valenzuela found that PSA was a highly sensitive marker of prostate cancer [13]. A longitudinal research project in Baltimore, revealed the relationship between serum PSA and prostate cancer [14]. In the late 1980s, PSA screening for prostate cancer diagnosis was adopted in the United States and subsequently in Europe, thus making the incidence of prostate cancer in the United States and Europe increase rapidly in the 1990s [15, 16]. However, due to the low specificity of PSA screening for high-grade, clinically significant disease, a considerable proportion of people were subjected to unnecessary prostate biopsy or diagnosed with indolent cancer resulting in overtreatment [17]. In Europe and the

United States, an estimated 23-42% of prostate cancer cases were over-diagnosed due to PSA screening (2002) [15]. Based on a large clinical study in the United States, USPSTF (US Preventive Services Task Force) recommended against PSA-based screening for prostate cancer for men of any age in 2012 [18], led to 18% relative decreasing in PSA screening rates for men aged over 50 between 2010 and 2013 [19]. However, among men greater than 75 years old recorded in the SEER database, the proportion of men presenting with metastatic disease increased from 2011 (7.8%) to 2013 (12.0%). A study conducted by ERSPC found that for every 1,000 men screened for PSA, three patients were prevented from metastasis and one was prevented from cancer-specific death. Based on these observations, in 2017, the USPSTF revised its recommendations and suggested that the decision to perform a PSA test for men aged 55 to 69 should be individualized by consulting their doctors [20]. Currently PSA screening rates in the United States and Europe are higher than that in other regions in the world. The latest data show that the screening rate for prostate cancer is 30.7% for American whites, 28.1% for American blacks and 25% for Asian Americans [21].

2.3Environmental factors

The incidence of prostate cancer in East Asia (~10/100,000) is much lower than that in North American Asians (~67/100,000) indicating that environment may be important risk factors for is prostate cancer. With the growth of economic and the

adoption of Western lifestyles during the last decades, the incidence of prostate cancer in East Asia has increased significantly. Singapore, Japan and Taiwan have experienced a sharp increase in the incidence of prostate cancer, from a low level (~5/100,000) to 30/100,000, 30/100,000 and 40/100,000 respectively [22-24]. The incidence of prostate cancer in China increased from 5/100,000 (2000) to 10/100,000 (2011) in ten years, and it is still growing. Current evidences have shown that obesity, high-fat diet, smoking and sunshine exposure may be risk factors for prostate cancer. Epidemiological studies on whether obesity increases the incidence of prostate cancer are very inconsistent, instead obesity increases the risk of prostate cancer progression [25, 26]. Animal model studies have shown that high-fat diets may induce prostate cancer progression through affecting growth factor signaling, lipid accumulation, inflammation and endocrine regulation. In addition, epidemiological evidence reveals that smoking is associated with the prognosis of prostate cancer[27], but its molecular mechanism is not well understood. There are studies indicating the short-day in high latitudes regions resulting high incidence of prostate cancer due to the insufficient sunshine exposure affecting vitamin D synthesis[28]. Nevertheless, low fat, vegetables, tomatoes, olive oil rich diets and exercise are protective for prostate cancer[29, 30].

2.4 Age

Age has a great impact on prostate cancer incidence and mortality. The risk of prostate cancer begins to increase in men with age over 40 years old and increases sharply when over 50 years old [31]. Over two-thirds of prostate cancer cases are older than 65 years old [31, 32]. However, there is evidence that the older you are, the more likely you are to develop prostate cancer, and it has been hypothesized that if men do not die of other causes, prostate cancer is inevitable. Men under 35 have almost no risk of prostate cancer. Therefore, the lower incidence of prostate cancer in Africans than that in North America Africans is largely due to the lower life expectancy of African males. With the aging of the global population, we need to pay attention to the increasing burden of prostate cancer.

2.5Infectious diseases

Over the years, several studies have focused on the association between prostate cancer and viral infections including human papillomavirus (HPV), herpesviruses including cytomegalovirus (CMV), human herpes simplex virus type 2 (HSV2), human herpesvirus type 8, (HHV8) and Epstein-Barr virus (EBV), polyomavirus BKV and xenotropic murine leukemia virus-related virus[33]. But a systematic review in 2013 indicated that there was insufficient epidemiological evidence showing that a single infectious pathogen was associated with prostate cancer, which

may due to the limited sample sizes in the study, resulting in the impossibility to assess non-persistent infection, and on the other hand, prostate cancer may be associated by multiple infectious pathogens [34]. A recent analysis of 5000 prostate cancer patients and 6000 healthy people has found that HPV16 infection increased the risk of prostate cancer [35].

Due to the development of second-generation sequencing technology, the doctrine of sterility in urine has been overthrown [36, 37]. Urinary microbial homeostasis and pathogenic microorganisms may cause prostate cancer inflammation and thus promote tumorigenesis [38]. Several bacteria have been proven to cause bacterial inflammation of the prostate, including E.coli. and other species of Enterobacteriaceae [39]. Propionibacterium acnes has also been demonstrated to cause prostate cancer inflammation and promote tumorigenesis [40-43]. Evidence also shows that some sexually transmitted pathogens such as Chlamydia trachomatis, Neisseria gonorrhoeae, or Trichomonas vaginalis are associated with prostate cancer [44-47]. Although it is hard to find a direct link between a single pathogen and the occurrence and development of prostate cancer, the relationship between various infections promoting inflammation of prostate cancer and the occurrence of prostate cancer has been gradually elucidated.

2.6 Family history

The genetic risk of familial prostate cancer was first proposed in the 1950s. Johns et al. carried out statistical analysis of 13 case-control studies, showing that men with a family history of prostate cancer had a 2.5-fold increased risk for prostate cancer. In a subsequent large number of studies, it was confirmed that men with a family history of prostate cancer had an increased risk of prostate cancer, especially with the presence of early onset, a relative, or multiple cases in the family. In 2015, Albright et al. conducted a retrospective analysis of 443 men with complete family history and assessed the relative risk of PC [48]. All these men had complete ancestral genealogy data. It was found that the risk of prostate cancer was increased by 2.46 if there was only one male immediate relative suffered from prostate cancer, while the risk reached by 7.65 when at least four male immediate male relatives had been diagnosed with prostate cancer [48]. In cases who have at least one immediate relative was diagnosed with prostate cancer before the age of 50, the risk was 5.54 [48]. Current studies have shown that the risk of prostate cancer in men is also increased due to rare germline risk variants such as BRCA1/BRCA2 or DNA mismatch, and the presence of breast cancer, ovarian cancer, rectal cancer and Lynch syndrome in families [49]. It is generally believed that the history of prostate cancer, breast cancer, ovarian cancer, rectal cancer and Lynch syndrome in relatives within three generations increased the risk of prostate cancer in men [50].

3. **Heritability of prostate cancer**

Early genetic quantitative studies of identical and fraternal twins have showed that the heritability of prostate cancer is 42% - 58%, higher than that of any other malignant tumors [51-53]. Linkage analysis based on pedigree prostate cancer has identified several chromosomal loci related to prostate cancer genetics [54-58]. In 1996, Jeffery et al analyzed 66 families at high risk for prostate cancer and identified for the first time the susceptibility locus of chromosome 1q24-25, which was named Hereditary prostate cancer 1(HPC1) [54]. Through this method, several chromosomes with different prostate susceptibility loci were also identified, including chromosomes 2, 3, 5, 6, 8, 10, 11, 13, 15, 17, 19, 20 and 22 [59, 60]. However, the chromosome regions identified by linkage analysis are too large to accurately identify the gene variant affecting the disease. Subsequent genome-wide association analysis (GWAS) and genome-wide exome sequencing further revealed susceptible variants in these regions, such as rs1006908 on chromosome 8q24 [61] and HOXB13 (G84E) on chromosome 17q21.[62]

3.1Loci with low penetration explain 19% of the hereditary of prostate cancer

Similar to other complex genetic diseases, the inheritance of prostate cancer is affected by variants with high frequency but low penetration and low frequency but high penetration in the populations [63, 64]. Genome-wide association analysis can

identify SNPs with high frequency but low risk. In 2016, Gudmundsson et al. reported the first genome-wide association analysis of prostate cancer, which confirmed that chromosome 8q24 carries prostate cancer susceptibility genes [65]. With the increase of GWAS research sample size and the development of chip imputing, the efficacy of GWAS in identifying prostate cancer susceptible SNPs has been greatly optimized, and more susceptible sites have been identified [66-69]. Scyumacher et al. recently have identified 63 new prostate cancer sites by analyzing data from 140,000 men[70]. Todate, 167 prostate cancer susceptibility loci have been identified by genome-wide association analysis [70]. These loci with high frequency and low penetration explain about 19% of the familial risk of prostate cancer [70]. However, most of the GWAS associated sites are located in the non-coding regions, and the molecular mechanisms of their effects on prostate cancer is still unresolved [71]. One hypothesis is that variants in these loci are associated coding regions affect prostate cancer and the other one is that these loci affect the transcriptional regulation region of coding genes.

3.2 Rare mutations on susceptible genes explain 15% of the hereditary of prostate cancer

In recent years, the effect of variants with low frequency and high penetration on the heredity of prostate cancer has become clear, such as HOXB13, BRCA1/BRCA2, DNA mismatch and DNA damage repair pathway [72]. These variants account

forabout 5% of the family risk of prostate cancer [73, 74]. In the following we will introduce the mechanism and clinical manifestations of the population frequency of these gene mutations on prostate cancer.

3.2.1 HOXB13

In 2012, Ewing et al. scanned 200 genes in 17-21-17q22 of 96 patients in different prostate cancer family clusters, and found that 18 patients had HOXB13 G84E rare mutations [62]. Furthermore, they verified the mutation in 5083 unrelated prostate cancer cases of European descent and 1401 control subjects. The mutation rate was approximately twenty-fold higher in the prostate cancer cases (1.4% or 72 in 5083) than in the controls (0.07% or1 in 1401)[62]. Odds ratio of HOXB13 G84E for the development of prostate cancer was 5.1 among men with positive family history and early onset and 1.7 among men with no family history and late onset [62]. Subsequently, in a large international family risk study of prostate cancer, 5% of prostate cancer families had the HOXB13 G84E variant [75], with the highest variant frequency around 20% in Finland, followed by Sweden around 8.2% [76]. Evidence suggests that HOXB13 G84E appeared most prevalent in the Nordic population, with a moderate penetration rate [77]. Different HOXB13 mutations have also been detected in prostate cancer cases in other racial or ethnic groups, including in African [62] (G216C and R229G), Asians (G135E) [78] and Portuguese (A128D,F240L)[79],

but the frequency and impact of these variants on the risk of prostate cancer remains to be further confirmed.

The HOXB13 gene encodes for a homeobox related transcription factor regulating a gene expression cascade which is critical for prostate development [80]. Other than that, HOXB13 protein has been shown to interact with Androgen Receptor signaling [81, 82], but its role in prostate tumorigenicity remains unclear and needs to be further investigation.

3.2.2 BRCA1/BRCA2

The relationship between BRCA1/BRCA2 and hereditary breast and ovarian cancer first revealed that BRCA1/BRCA2 mutations increase the risk of breast and ovarian cancer[83]. Subsequent studies have found that rare mutations in BRCA1/BRCA2 also increase the risk of prostate cancer [84]. Many studies have shown that BRCA1 mutations increase the risk of prostate cancer between 1.07 and 3.75 (odds ratio), while BRCA2 mutations increase the risk of prostate cancer more significantly between 4.65 and 8.6 (odds ratio). In a BCLC study involving 173 BRCA2 mutated families, the overall risk associated with prostate cancer in BRCA2 was 4.65, but in men who developed the disease before age 65, the risk rose to 7.33. In the study of prognostic analysis of patients with BRCA1/BRCA2 mutation, PSA-free survival and

overall five-year survival were significantly reduced in patients with BRCA2 mutations. Strikingly, in a recent genomic study that including 150 metastatic castration resistance prostate cancers (mCRPC), 8% of the patients had germline variants of BRCA2. BRCA1 and BRCA2 occur in about 0.87% and 5% of familial prostate cancers. Both BRCA1/BRCA2 mutations increases the risk of prostate cancer and BRCA2 mutations are associated with early onset of disease and prognosis. Due to relatively low numbers of cases carrying BRCA1/BRCA2 and the variability of specific mutations, the differences in mutation frequencies among different races is still unknown.

BRCA1/BRCA2 is a tumor suppressor gene encoding proteins that repair damaged double-stranded DNA by high-fidelity replication using the undamaged sister chromatid as a template [85]. BRCA1 has a broad range of functions, including recruiting effector factors to double strand break sites, regulating end resection of DSBs, activating G1/S, S-phase, and G2/M checkpoints, and mediating non-homologous end joining, and single-stranded DNA annealing repair pathways [85, 86]. The BRCA2 protein mainly recruits RAD51 and HR60 to DSB at the beginning of homologous recombination repair [87]. Deleterious mutations in BRAC1 and BRCA2 impair BRCA1/BRCA2 function and lead to instability of DNA, thus promoting carcinogenesis [88].

3.2.3 DNA mismatch repair genes

DNA mismatch repair gene mutations was first identified as the causative gene of Lynch syndrome[89]. Researchers subsequently observed that Lynch syndrome patients had a higher incidence of cancer than that of the normal population, especially in rectal cancer. In 2004, Harakdsdottir et al. analyzed the SEER database and observed 11 out of 188 men with Lynch syndrome developed prostate cancer (relative risk 4.87). In 2009, Grindedal et at. confirmed that DNA MMR mutations increased the risk of prostate cancer by analyzing 106 men with DNA MMR mutations in the Norwegian Cancer Registry [90]. Ryan et al. assessed the relative risk of prostate cancer associated with DNA MMR mutation to be 3.36 [91]. Then Rosty et al. estimated the relative risk of prostate cancer associated with MSH2 (5.8), MLH1(1.1), and MSH6(1.3)respectively[92]. Inactivation of MMR proteins, result in a high rate of microsatellite instability (MSI) in their tumors [93, 94].

3.2.4 Other prostate cancer germline risk variants

Other germline mutations are also associated with the risk of prostate cancer. In a study investigating germline risk mutations in metastatic prostate cancer, researchers found that 1.87% of metastatic tumors had CHEK2 germline mutations, 1.6% of ATM, 0.43% of PALB2 and 0.43% of RAD51D germline mutations [95]. A Polish study has found that NBN657del5 was present in 9% of pedigree prostate cancer

patients, and 2% of sporadic prostate cancer patients, while the mutation frequency in the control group was only 0.6%. In addition, in the recent TCGA pan-cancer study for screening germline risk variants, it suggested that germline variants on BRIP1, DKC1, EPCAM, GALNT3, MTAP, PMS2, POT1, RAD51C, RECQL, SEPRINA1, TSC1, TSC2, UROD and some other genes might also affect prostate cancer.

4. Somatic Genomic alterations in prostate cancer

Normal cells need to acquire multiple abilities to transform into tumor, including sustained growth signals, desensitization of growth inhibition signals, resistance to death, unlimited replication and proliferation, sustained angiogenesis, tissue infiltration and metastasis, avoidance of immune surveillance, and abnormal energy metabolism. Inflammation and genomic instability generate somatic mutations that expedites the acquisition of tumor promoting abilities. Before innovation of next generation sequencing technology, researchers identified a number of somatic mutation events in prostate cancer through fluorescent in-situ hybridization and comparative genomic hybridization, including loss of NKX3-1 [96, 97] and PTEN[98], amplification of MYC [99, 100] and AR [101]. Subsequent advances in gene expression profiling technologies led to the discovery of ETS family transcript factor over expression[102], and the discovery of recurrent TMPRSS22-ETS gene fusions in prostate cancer [103]. With the development of next generation sequencing

and bioinformatic analysis technologies comprehensive examination of prostate cancer genomes became an efficient tool for identification of cancer related mutations. In early prostate cancer genomic studies using next-generation sequencing, a large number of novel prostate cancer genomic alterations were identified [104-111]. However, due to the limitation of the sample sizes and early generation analytical methods, these studies failed to uncover the full spectrum of the complex heterogeneity of prostate cancer genomes. In 2015, the Cancer Genome Atlas （TCGA）published a multi-omic genomic study which included 333 cases of primary prostate cancer [112]. Researchers found that 74% of prostate cancers could fall into seven molecular subtypes: 1) ERG fusion (46%), 2) ETV1 fusion (8%), 3) ETV4 fusion (4%), 4) FL1 fusion (1%), 5) SPOP mutation (11%), 6) FOXA1 mutation (3%) and 7) IDH1 mutation (1%) [112]. Besides these reported mutations, TP53, PTEN, PIK3CA, RB1 and other gene alterations were also altered. In 2018, Joshua Armenia et al. combined genomic data from 1013 prostate cancer samples and identified that 97 genes were significantly mutated in prostate cancer, 70 of which had not been reported before, such as the ubiquitin ligase CUL3 and the transcription factor SPEN [113]. In addition, they defined a new molecular prostate cancer subtype that has mutations on epigenetic related genes [113]. Other related studies have also revealed that local indolent prostate cancer had lower numbers of molecular mutations [114], instead metastatic prostate cancer had higher mutation burden rates [115].

The unveiling of prostate cancer genome provided insights into prostate cancer biology and have enabled the identification of novel drug targets, diagnosis and risk stratification biomarkers for this disease. The molecular biology, diagnosis, risk stratification and therapeutic development of prostate cancer are presented in Chapter 1.7，1.8and 1.9.

5. Molecular evidence for racial disparities of prostate cancer

In recent years, the molecular basis for racial disparities of prostate cancer has accumulated. Early GWAS studies revealed the 8q24 locus [58, 61, 116] and some other prostate cancer risk alleles [117, 118] were associated with African Americans. GWAS studies based on men from different populations, such as Latino [119], South Asian [120], Japanese [68, 121], and Chinese [122, 123] ancestries have further identified potential race-specific prostate cancer risk alleles. Later whole exome sequencing revealed that the rare mutation HOXB13 G84E occurred more frequently in the family prostate cancer of the Nordic populations [62, 76, 77, 124] and the rare variation in Tet2 is associated with clinically relevant prostate carcinoma in African Americans [125]. Most recently, prostate cancer genomic studies revealed ETS family fusions was the most common alteration in prostate cancer of Caucasian men at a frequency of ~50%, however, they are much lower in African American and Chinese [103, 126, 127]. Similar to ETS fusion, PTEN loss is more frequently found in

prostate cancer of Caucasian men than other ethnicities [128]. There are also other race-specific prostate cancer genomic alterations such as LSAMP loss [129], ERF loss-of-function mutations [130, 131] and CDC27-OAT [132] fusion in prostate cancer of African Americans. However, most of the genomic prostate cancer studies are focused on cohorts of European ancestry, leaving minority groups under represented [133]. Furthermore,in racial mixing, the ethnic contribution to riskis not fully understood.

6. Oncogenesis and development of tumors

The size of the prostate is slightly larger than the walnut. It is located at the bottom of the pelvic cavity, under the neck of the bladder, on the urethra; behind the pubis, before the rectum, and surrounds the junction of the bladder mouth and the urethra, with a seminal vesicle gland attached, and urethra and the vas deferens wrapped [134]. **(Figure 1.1)** The main function of the prostate is to secrete and store prostatic fluid. Prostatic fluid can be mixed with sperm to form semen which contains sperm and semen. About 10% to 30% of the semen is made from the prostate [134, 135]. The prostate also contains smooth muscle tissue, which helps ejaculation[135]. As prostate is located upstream of most male genital organs, including the vas deferens, epididymis and testis, the prostate is also considered the first line of defense for the male reproductive system against foreign antigens or pathogens from the bladder and

lower urethra [136]. The prostate is divided into transitional zone, central zone and peripheral zone. It consisted of 30 to 50 glands, with each one consisting of three basic types of cells, basal cells, luminal cells and neuroendocrine cells [137]. Most prostate cancers originate from luminal cells in the peripheral area of prostate cancer [138, 139]. The normal epidermal cells of the prostate develop into prostatic intraepithelial neoplasia, then local prostate cancer, then locally advanced prostate cancer, and eventually developed into metastatic prostate cancer [140-144]. With increasing age, inflammation prostatic atrophy is common [145, 146]. Inflammation atrophy has a larger impact area, especially in the peripheral areas with high incidence of prostate cancer [147]. Inflammation can be caused by infection and disruption of the epithelial barrier [33, 148]. Evidence suggests that inflammation-induced oxidative stress or reactive oxygen species (ROS) cause mutations in cell genes, leading to the transformation of normal cells to prostatic intraepithelial neoplasia[149-151]. In 2016, Mani, R.S. et.al found that inflammatory cytokine signals such as tumor necrosis factor (TNF) signals in epithelial cells result in DNA breaks which eventually lead to the fusion of TMPRSS2-ERG [152]. Other studies indicate bacterial infection leads to low expression of NKX3-1 in prostate epidermal cells [149]. Proliferative luminal epithelial cells of intermediate phenotype which abnormally express gene such as CDKN1B, GSTA1, COX2, MYC, PSA, AR, NKX3-1, MET, GSTP1 are enriched in prostatic proliferative inflammatory atrophy (PIA) [148, 153]. These studies suggested that inflammation can transform normal

cells to prostatic intraepithelial neoplasia. Besides, environmental toxins may cause DNA damage in normal epidermal cells of the prostate by triggering oxidative stress, which may lead to the carcinogenesis of epidermal cells, such asheterocyclic amines ingested from burnt diets through blood circulation to the prostate or by the urine reflux into the prostate also trigger oxidative stress [154-156].

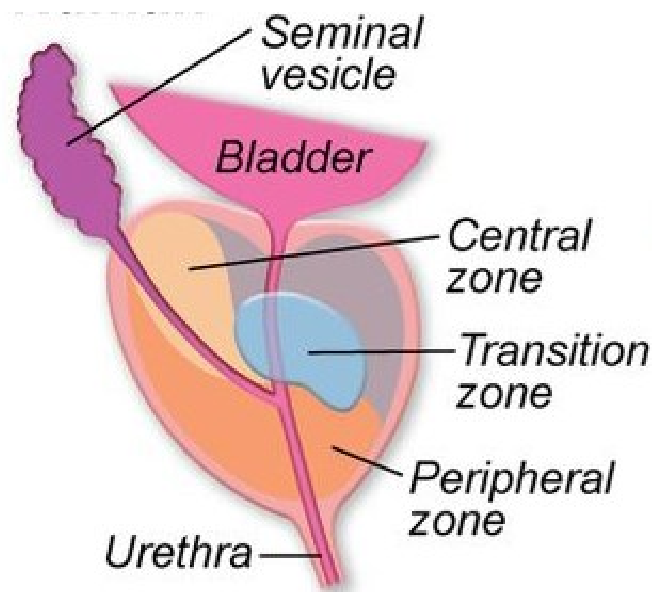

Figure1.1 The size of the prostate is slightly larger than the walnut. It is located at the bottom of the pelvic cavity, under the neck of the bladder, on the urethra; behind the pubis, before the rectum, and surrounds the junction of the bladder mouth and the urethra, with a seminal vesicle gland attached, and urethra and the vas deferens wrapped.

Most prostate cancers are indolent, and only a small proportion of prostate cancers progress aggressively to metastatic tumors. Metastasis prostate cancer undergo

Epithelial to mesenchymal transition (EMT) to migrate to the adjacent lymph nodes and then to the lungs, liver or bone[157, 158]. Recent studies have suggested that circulating tumor cells [159, 160] and exosomes[161] may play an important role in tumor metastasis. Despite the efforts to reveal the mechanism of prostate cancer metastasis and to develop drugs for the treatment of metastatic prostate cancer, metastatic prostate cancer has not yet been cured so far.

7. Prostate cancer diagnosis

Current clinical diagnosis of prostate cancer is based on the framework established in the 1990s. It mainly includes three main indicators: serum PSA level, rectal digital examination and tissue biopsy. Patients with serum PSA greater than 4.0 or rectal finger positive need to be further confirmed by biopsy [162, 163].

7.1 PSA screening

In 1970s, Wang and Valenzuela found that PSA was a highly sensitive biomarker of prostate cancer[13]. Then, in a longitudinal study in Baltimore, the relationship between PSA in serum and prostate was further elucidated [14]. In the late 1990s, PSA was used to screen prostate cancer because of its high sensitivity to detect prostate cancer, replacing the previously used prostatic acid phosphatase (PAP)[164].

PSA is a glycoprotein and composed of 237 amino acids, secreted by prostate epithelial cells. Normally, PSA is not released into the blood, because of natural blood-epithelial barrier between the prostatic duct system and the peripheral circulation system, thus maintaining low concentration in the blood. The invading and migration of prostate cancer cells disrupt blood-epithelial barrier, resulting in increasing of PSA concentration in the blood. Unfortunately, blood PSA elevation is also observed in benign prostatic hyperplasia (BPH) and prostatitis [165]. High levels cannot distinguish cancer from inflammation especially in patients with serum PSA levels <10 ng/ml[166]. Instead of only serum PSA level, the ratio of free PSA to total PSA, PSA density, PSA ROS curve have been shown to increase the specificity of PSA, but with limited effect [167].

7.2 Biopsy

The mode of prostate biopsies has evolved considerably over the years. Open or finger-guided transperineal biopsy with low accuracy was the standard method for biopsy sampling in the early 1920s. This method is very invasive, leading to high incidence of urinary incontinence and ED, and patients need to spend a long time in hospital to recover. The emergence of transrectal ultrasound (TRUS) guided prostate biopsy greatly improves the accuracy of biopsy and reduces the side effects of tissue biopsy [168]. In 1998, Levine and colleagues found that 12-core prostate biopsy

increased the detection rate of prostate cancer by 30% compared with 6 cores[169]. In 2003, the guideline for prostate cancer biopsy was revised to 12 needle biopsy and enhanced sampling of the anterior and lateral regions [170]. Nevertheless, needle biopsy has a significant chance of missing the prostate cancer. Thirty percent of patients with negative biopsy were identified as prostate cancer positive by repeat biopsy [169, 170]. In addition, transrectal ultrasound guided prostate cancer puncture also increases the risk of bacterial infection of the prostate [164].

7.3 Risk stratification

The risk stratification system for prostate cancer was originally derived from the D'Amico classification [171] of low-risk, medium-risk, high-risk disease or the the Epstein criteria [172] for low-risk (clinically unimportant) cancer mainly based on Gleason score of tissue puncture biopsy, serum PSA level and clinical classification. However, the inaccuracy of the risk stratification resulted in a high proportion of prostate cancers were inaccurately assessed, which lead to over-treatment or inadequate treatment.

7.4 Development of prostate cancer diagnosis and risk stratification

7.4.1 PSA derived

Fortunately, the understanding of the biology of PSA secreting and maturation helps to the identification of greater specificity of PSA isoform biomarkers. In particular, the detection level of the [-2] pro-PSA ('p2PSA') in serum can improve the diagnostic specificity of prostate cancer compared with free PSA or total PSA. PSA is initially translated as inactive pre-pro PSA, which carries a 17-amino acid signal peptide. During the secretion of PSA, pre-pro PSA is cleaved to pro-PSA by removing the signal peptide. Mature PSA, pro-PSA still requires cleavage by human kallikrein 2 (hK2) to remove the Pro precursor peptide of 7 amino acids at N terminal. In this step, alterative splicing isoforms have been found including [-1], [-2], [-4], [-5] and [-7] pro-PSA [173].

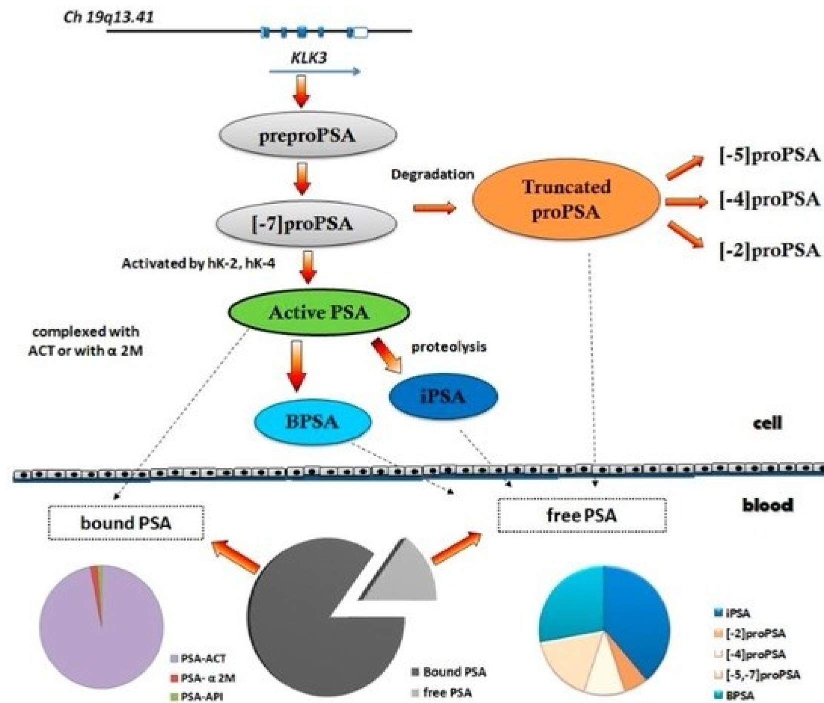

Figure1.2 Molecular forms of PSA. The arrows with dashed line mean the forms of PSA that go from the cell to the blood. PSA: prostate specific antigen, BPSA: benign PSA, iPSA: intact PSA, PSA-ACT: Alpha 1-antichymotrypsin-PSA, PSA-API: alpha1-trypsin inhibitor PSA, PSA-A2M: alpha 2macroglobulin, hK-2: human kallicrein 2, hk-4: human kallicrein 4.

7.4.1.1PHI index

In 2012, the Prostate health index (PHI)was approved by the U.S. Food and Drug Administration (FDA) for aiding doctors to determine whether a patient with PSA level 4-10ng/ml requires biopsy[174]. Prostate health index (PHI) is a diagnostic index for prostate cancer that combines serum total PSA, free PSA and [-2] pro-PSA

levels. It is calculated by the formula ([-2] pro-PSA/free PSA) *√tPSA. Patients with high serum level of total PSA and [-2] pro-PSA, and with low level of free PSA have higher PHI scores and are more likely to suffer from prostate cancer [174]. Fossati and colleagues evaluated PHI index in a total of 2034 patients with PSA at 2.5-10 ng/ml. In the ROC analysis, the AUC (Area under the curve) of PHI was 0.77, while the AUC of -2] pro-PSA was 0.76, % fPSA was 0.68, PSA was only 0.5[175].

7.4.1.2 The 4Kscore

The 4Kscore combines four prostate-specific biomarkers including total PSA, free PSA, intact PSA, and human kallikrein 2 [hK2] with clinical information to provide men with an accurate and personalized measure of their risk for aggressive prostate cancer [176, 177]. The 4Kscore Test was evaluated in multiple cohorts in Europe and subsequently validated in a US multicenter prospective study. Overall, the 4Kscore Test reduced prostate biopsy rates by 94% in men and there is evidence that the 4Kscore can predict the likelihood of cancer spreading to other parts of the body in the next 20 years [178-180].

7.4.2 Urine-derived biomarkers

7.4.2.1 SelectMDx

SelectMDx is a urine-based test based on prostate cancer genomic biomarkers HOXC6 and DLX1, which together have a 76% sensitivity (reliability) for detecting a prostate cancer that is aggressive enough (Gleason score of 7 or greater) to require treatment [181]. When additional risk factors (namely age, PSA, PSA density, family history and rectal examination) are added to the calculation, the accuracy of the test (negative predictive value) for excluding prostate cancer rises to 98%, with an expected total reduction of biopsies of 42%. So clearly there is great potential to avoid unnecessary prostate biopsies. A multi-centre scientific study made in 2016 and published in European Urology confirms the accuracy of the test[181]. while a 2018 study in the Journal of Urology has concluded that ‘routine use of the SelectMDx urinary biomarker panel to guide biopsy decision making improved health outcomes and lowered costs in American men at risk for prostate cancer [182]. This strategy may optimize the value of prostate cancer risk assessment in an era of increasing financial accountability.

7.4.2.2 ExoDx

ExoDx Prostate(IntelliScore) is a clinically validated, non-digital rectal exam (DRE) urine-based liquid biopsy test that predicts the presence of high-grade (Gleason score

≥7) prostate cancer for men 50 years of age and older with a PSA 2 – 10 mg/mL presenting for an initial biopsy. A "rule out" test, ExoDx Prostate(IntelliScore) is designed to more accurately predict whether a patient presenting for an initial biopsy does not have high-grade prostate cancer and, thus, could potentially avoid an initial biopsy and, instead, they continue to be monitored [183]. ExoDx Prostateanalyzes the urine for three biomarkers on exoRNA that are expressed in men with high-grade prostate cancer. Using a proprietary algorithm that combines the relative weighted expression of the three-gene signature, the test assigns an individual risk score for patients ranging from 0 to 100. A score >15.6 is associated with an increased likelihood of high-grade prostate cancer on a subsequent biopsy [184].

Additional molecular markers of prostate cancer based on different urine are being studied, such as MI-Prostate Score based on urinary secretory microRNA, Progensa and Protarix based on urinary proteomics[185].

7.4.3 Genic Tests

In 2017, the Philadelphia Consensus recommended that prostate cancer patients with a family history of HBOC and prostate cancer, as well as metastatic prostate cancer, need to undergo genetic screening for germline risk variants on genes includingBRCA1/2, HOXB13, and DNA mismatch repair genes [50]. Risk variants in

these genes not only increase the susceptibility to prostate cancer, but are also partly associated with the early onset and prognosis of prostate cancer or treatment selection [50].The risk genes and evidence discussed in the Philadelphia Consensus are shown in Table 3.

8. Treatment of prostate cancer

Alterative treatment selection for prostate cancer patients are mainly based on risk classification. In recent years, with the available of new targeted therapies and immunotherapy, genic testing become a supplemental option to guide the treatment selection [49].

8.1 Active monitoring for very-low risk prostate cancer

Recent studies have shown that very low-risk prostate cancer tends to remain indolent. The probability of their progression to metastatic prostate cancer or cause death is very low. To avoid side effects caused by treatment, in 2017, the American Urological Association announced that it is best for men who are diagnosed with very low-risk prostate cancer to actively monitor disease progression rather than receive treatment.

8.2 Activemonitoringfor low risk prostate cancer

Men who are diagnosed with low-risk prostate cancer should also be given priority in actively monitoring disease progression as well.Active monitoring is usually performed on patients with low-risk prostate cancer undergoing serum PSA monitoring, repeated prostate biopsy and MRI. Datahave shown that patients with low-risk prostate cancer under active monitoring have a probability of dying from prostate cancer less than 1% within10 years.[186]

8.3 Treatment for patients with low risk cancer

Because some lesions affect the daily life of patients, for patients with low-risk local prostate cancer, radical prostatectomy, external beam radiotherapy and brachytherapy are often used clinically. Othertreatments, such as cryotherapy, high-intensity focal ultrasound and photodynamic therapy, are also used [186].

8.4 Treatment for locally advanced prostate cancer

For patients with locally advanced or high-risk prostate cancer, recent studies have shown that radical prostatectomy plus ADT or radiotherapy plus ADT can significantly reduce the 10-year mortality risk compared with ADT alone. However, 10% to 20% of patients havebiochemical recurrence with elevated serum PSA in 2~3

years after ADT treatment and develop into castration resistant prostate cancer (CRPC) followed by metastasis to bone, lung and other sites (mCRPC). The median survival time of patients diagnosed with castration resistance prostate cancer is 15 to 36 months[187].

8.5CRPC treatment

8.5.1 Second Generation Androgen Receptor Antagonists

Although many patients develop resistance after 2~3 years of ADT treatment, mCRPC is AR signaling dependant. In 2012, the second generation AR antagonist Enzalutamide [188]was approved by FDA in the United States for the treatment of mCRPC. In clinical trials, Enzalutamide significantly improved overall survival (OS) and progression-free survival (PFS) in patients with prostate cancer after chemotherapy.

In 2012, Abiraterone, an inhibitor of CYP17A1, was also approved by FDA to treat mCRPC. CPY17A1 is an important enzyme in androgen synthesis pathway. It can effectively inhibit androgen synthesis by inhibiting CYP17A1. Clinical data showed that abiraterone combined with low-dose synthetic glucocorticoids effectively prolonged progression-free survival of the patients (5.6 vs 3.6 months, $P < 0.001$)[189].

8.5.2 PAPR inhibitor

A recently completed phaseclinical trial showed that PARP inhibitor (Lynparza/a/olaparib) significantly improved disease-free progression survival in mCRPC patients with BRCA1/2, and ATM mutations, compared with Enzaluramide and Abiraterone [190].

8.5.3 Siupleucel-T

Sipuleucel-T is an autologous dendritic cell vaccine targeting Prostatic acid phosphatase (PAP). In 2010, the FDA approved sipuleucel-T as the first and only immunotherapy for mCRPC[191]. In clinical trials, the 36-month survival rate in the Sipuleucel-T group was 31.7% while 23.0% in the placebo group[192].

8.5.4 Chemotherapy

In 2004, TAX327 reported that chemotherapy drug docetaxel improves outcomes in mCRPC. Tannock et al. demonstrated in a randomized trial of 1,000 men with mCRPC that docetaxel chemotherapy improved patient survival by nearly 3 months, with 45% of patients having a 50% reduction in PSA [193]. In 2015, the taxanes

represented by docetaxel in combination with prednisone have been used as first-line treatments for mCRPC patients [194].

8.5.5 Ra223mCRPC

Ra223 was approved by FAD in 2013 for the treatment of mCRPC with bone metastases. Clinical trial shows that Ra223 can increase the median survival time of mCRPC patients with bone metastases by 4.5 months[195].

9. Translate genetics to biology and therapeutics

The treatment of mCRPC has made great progress, but due to the heterogeneity of tumors and drug resistance during treatment, the fact is that mCRPC still cannot be cured. In recent years, a large number of new prostate cancer driver genes (potential targets) have been identified in prostate cancer genomic studies, and the breakthroughs in immune cell therapy and immune checkpoint therapy in other tumors have encouraged researchers to move forward.

In 1941, Charles Huggins first reported the beneficial effect of androgen ablation on metastatic prostate cancer. This discovery greatly inspired researchers to develop androgen deprivation therapy (ADT) for prostate cancer. ADT therapy is still a very important therapy for prostate cancer. Although 10% to 20% of patients have

biochemical recurrence with elevated serum PSA in 2~3 years after ADT treatment and progresses to CRPC which is associated with a poor prognosis.

Recent advances in prostate cancer genetics and genomics have provided considerable insights into prostate cancer biology and have identified a considerable number of cancer drivers which can be exploited as novel drug targets.

By 2018, More than 97 somatic drivers, 20 susceptibility genes, 167 germline risk alleles, in prostate cancer have been identified. To unveil the targetability of them, Wedge et al. conducted computational chemogenomic analysis of prostate cancer drivers and identified 11 targets of approved drugs, 7 targets of investigational drugs, and 62 targets with compounds that may be active and should be considered candidates for future clinical trials[196].

The following summarizes the most promising targets for prostate cancer discovery and highlight key signaling pathways as potential sources of targets including Androgen receptor signaling, the PIK3-AKT signaling, the WNT signaling, the DNA repair defects, the MAPK signaling.

9.1 AR signaling

9.1.1 GnRH

Gonadotropin-Releasing Hormone (GnRH), also known as Luteinising-hormone releasing hormone (LHRH), plays a crucial role in anti-androgen therapy [197]. Natural GnRH receptor agonists are secreted by hypothalamus and activate the receptors in the pituitary, increasing the release of LH (luteinizing hormone[198]) and ACTH (adreno-cortico-tropic-hormone) from the pituitary[199]. LH and ACTH promote corresponding target organs respectively to increase androgen secretion. In 1971, the chemical structure of GnRH in pigs was obtained [200]. Then a series of active analogue agonists were developed.

The persistent activation of GnRH receptor by GnRH analogue agonists depletes the pituitary of LH and ACTH, and the ultimate result is that androgen levels continue to drop to a very low level. In a phase III trial for patients with locally advanced prostate cancer, the 5-year clinical disease-free survival rate was 40% (95% CI 32% - 48%) in the radiotherapy group compared to 74% (95% CI 67% - 81%, $P < 0.001$) in theGnRH analogues and radiotherapy combined therapy group [201]. Therefore, GnRH analogues can significantly improve the outcome of locally advanced prostate cancer patients.

The persistent activation of GnRH receptor in the early stage causes a temporary rise in testosterone levels and promotes the progression of the disease. To resolve the adverse effects of GnRH receptor agonists, the development of GnRH receptor antagonists has been investigated. Antagonists inactivate the GnRH receptor by competitive binding with receptors. In 1997, Abarelix, a potent GnRH receptor antagonist, was successfully developed. Subsequently, several GnRH receptor antagonists, such as Degarelix, and Relugolix were introduced into the market [202, 203].

9.1.2 AR

9.1.2.1 AR-LBD

In 1990, the first generation of non-steroidal androgen receptor antagonists, flutamide [204] was approved and rapidly used as an important drug in the treatment of advanced prostate cancer, followed by Nelutamide and bicalutamide [205]. Flutamide must be absorbed in the gastrointestinal tract and metabolized in the liver to be activated leading to hepatotoxicity [206]. In addition， Drug resistance to the first generation non-steroidal androgen receptor antagonists caused by T877A[207],W741C[208] and F876L[209] mutations in the ligand binding domain of AR is usually observed within 1 year after first administration. In 2012, the second-generationnon-steroidal androgen receptor antagonist Enzalutamide was

approved by FDA. Enzalutamide is able to overcome the resistance caused by W741C mutation and has 8-fold greater affinity for AR than the first generation of non-steroidal androgen receptor antagonist[188]. Apalutamide with the same mother-ring chemical structure as Enzalutamide has been approved to treat CRPC in 2018 [210]. There was no obvious hepatotoxicity for the two drugs but seizures or rash and hypothyroidism are common side effects [188, 211]. Drug resistance is usually observed about 2-3 year later after administration because of AR F867L mutations, over-expression of AR-V7, AR co-activator, or activation of glucocorticoid receptor signaling [212].

A new non-steroidal androgen receptor antagonist,Darolutamide which is able to overcome the resistance caused by AR mutations including F867L, W741L and T877A has shown stronger antitumor activity and stronger AR affinity than Enzalutamide [213, 214]. It should be approved by the FDA to treat the mCRPC in the near future.

9.1.2.2 AR-nonLBD

Androgen receptor (AR) is a steroid hormone receptor in the nucleus, which contains a central DNA binding domain (DBD), ligand binding domain (LBD), and hinge domain and N-terminal domain(NTD). Mutations in AR-LDB domains and the

expression of AR-V7 without AR-LDB domain can lead to resistance to Enzalutamide and other second generation of AR antagonists [212, 215]. To overcome the resistance, drugs that bind AR non-LBD domain become a new strategy. Niclosamide has long been used as an anthelmintic, but it have been found that it can promote the degradation of AR-V7and effectively inhibit the growth of tumors [216]. Other strategies for degradation of AR have been investigated as well. The clinical trial of AR-110, an effective AR degrading agent developed by PROTAC (Proteolysis Targeting Chimera) technology is ongoing.

9.1.3 AR-binding protein

HSP90 is a chaperone protein that binds AR and maintains full-length AR in a high-affinity ligand-binding conformation[217]. Inhibition of HSP90 results in abnormal AR signaling [218, 219]. Both in vitro and in vivo models, HSP90 inhibitors also result in depletion of AR-V7[220]. However, phase I and phase II studies of HSP90 inhibitors have been generally disappointing because of poor patient tolerability and modest antitumor activity [221, 222]. Nonetheless, in the clinical studies of patients with advanced prostate cancer, inhibitor of HDAC which acetylates and activates HSP90 by acetylation, showed anti-tumor activity[223].

FOXO1 binds to AR-NTD domain and inhibits AR transcriptional activity[224]. However, PTEN deletion in prostate cancer results in AKT activation which phosphorylates the FOXO1 resulting in its nuclear exclusion [225, 226]. In a phase II clinical study, ATK inhibitor Ipatasertib combined with abiraterone showed better anti-tumor activity than abiraterone alone, especially in prostate cancer patients with PTEN deletion [227].

9.1.4 CYP17A1

CYP17 are key enzymes in the synthesis of testosterone. Ketoconazole, an antifungal drug, has been found to broadly inhibit CYP17 enzymes and has been widely used in the treatment of prostate cancer before [228]. Due to its hepatotoxicity, ketoconazole has been limited in the clinical treatment of prostate cancer[228]. Nonetheless, a CYP17A1 inhibitor, abiraterone[189]combined with low-dose glucocorticoid effectively prolonged the progression-free survival (5.6 vs 3.6 months, $P < 0.001$) of mCRPC. Abiraterone, was successfully approved by FDA to treat mCRPC in 2012.

9.1.5 5alpha reductase

5 alpha reductase can reduce testosterone to highy active dihydrotestosterone [229], so 5alpha reductase inhibitors also play a role in the treatment of prostate cancer[230].

For example, dutasteride combined with abiraterone can effectively improve the therapeutic effect of abiraterone[231].

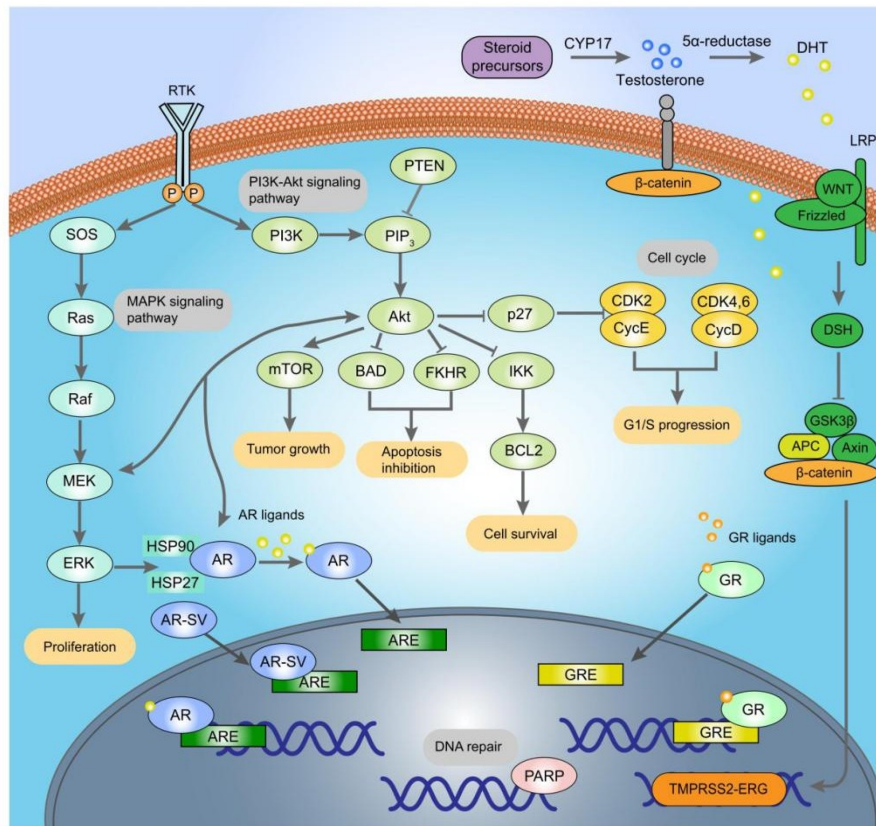

Figure 1.3 The cellular biology of prostate cancer. The complex underlying cellular biology and signaling cascades associated with prostate cancer are illustrated.

9.2 DNA repair Defects

Both germline and somatic genomic aberrations of DNA repair genes have been reported in prostate cancer. Causative germline mutations in DNA defect repair genes occur in 5% of hereditary prostate cancer carries[64], and strikingly, in 8-15% of mCRPC. Somatic mutations of DNA repair genes occur in almost 23% of mCRPC[111]. The most commonly aberrant genes are BRCA1/2 and ATM.

In 2005, studies showed that PARP inhibition could lead to death of BRCA1/2 deficient tumor cells [232]. Several studies over the past decade have demonstrated the utility of PAPR inhibitors in different types of tumors[233-235]. The PARP inhibitor olaparib has been approved by the FDA for the treatment of women with advanced ovarian cancer. While for mCRPC, recent clinical trials showed that PARP inhibitor (Lynparza/a/olaparib) significantly improved disease-free progression survival in patients with BRCA1/2, ATM mutations, compared with Enzaluramide and Arbitrone[190].

PARP (Poly ADP-ribose polymerase) detect and initiate an immediate cellular response to single-strand DNA breaks. Inhibition of PARP in the DNA repair deficient tumor results in synthetic lethality of the tumor cells [232, 236]. There are also studies suggesting that PARP inhibitors may have broader anti-tumor activity as PARP interacting with ETS and AR [237, 238].

In addition, DNA mismatch genes aberrations such as MSH6, MSH2, PSM2 occur in prostate cancer and potentially other defects in DNA repair [90, 239, 240]. These aberrations impair DNA mismatch repair, lead to an increase in mutation burden of tumors. There is also evidence showing microsatellite instability in tumors associated with DNA mismatch mutation, therefore, MMR defects can sensitize cancer to immunotherapy (Anti-PD-1, anti-CTLA4 therapeutics)[241, 242]. So far, patient selection approaches in CRPC for immune-checkpoint targeting have yet to be pursued.

9.3 PIK3-AKT pathway

PIK3-AKT signaling pathway is altered in 19% of local prostate cancer and 30% ofmCRPC. PTEN deletion is the most common aberration affecting 12% of local prostate cancers and 25% of mCRPC[112, 243]. Animal model experiments confirmed that the deletion of PTEN resulted in the formation of precursor prostate cancer lesions [244] and promoted disease progression when such features are combined with abnormalities in ERG, TPP53 [245, 246].

In tumors lacking PTEN, PI3KCA activity is suppressed while PI3KCB signaling is active. But in clinical studies with PI3KCB inhibitors, whether it was used alone or in

combination with docetaxel, the overall 1-year survival rate have not been shown to improve[247]. The reason for this may be that the inhibition of PI3KCB only results in the inhibition of AKT-mechanistic target of mTOR signaling which relives the feedback inhibition onupstream substrates and thus causes activation of PI3KCA and a rebound in downstream signaling[248]. However, Dactolisib, an inhibitor of multiple targets including PI3K and mTOR also have not shown any therapeutic advantage [249].

Inconsistent with PI3K and mTOR inhibitors, ATK inhibitor Ipatasertib combined with abiraterone showed better anti-tumor activity than abiraterone alone, especially in prostate cancer patients with PTEN deletion [227]. Studies have shown that the anti-tumor activity of ATK inhibitors benefit from PTEN-AKT-FOXO1 axis instead of AKT-mTOR signaling [225].

Recent studies also indicate that PTEN loss induces cellular senescence and myeloid-derived suppressor cell infiltration can block this senescence [245, 250, 251]. In the PTEN-null mouse model, infiltration of CD11b+, glucocorticoid receptor 1-positive myeloid cells protect a population of proliferating tumor cells from senescence. These myeloid-derived suppressor cells appear to infiltrate the prostate along a chemokine-chemokine receptor (CXCR2), and release IL-1 receptor antagonist, which inhibits senescence and drives proliferation. These findings

suggested that, targeting innate immunity may be a new therapeutic approach for PTEN-loss prostate cancer [252, 253].

9.4 ETS gene rearrangements

The transcription factors of ETS family such as ERG, ETV1, ETV4, FLI1 have important oncogenic roles in many prostate cancers. ETS rearrangements were found in about 40-60% prostate cancer patients of European ancestry. The most common rearrangement are ERG (46%), followed by ETV1 (8%), ETV4 (4%)and FL1(1%)[254].

ERG usually fuses with TMPRSS2 which is regulated by an androgen -regulated promoter element[103]. TMPRSS2-ERG fusion leads to ERG overexpression, resulting in AR expression and tumor cell proliferation[255]. Overexpression of ETS induces the formation of prostatic intraepithelial neoplasia(PIN) in a genetically engineered mouse model[153]. When combined with increased AR signaling or PTEN loss，Overexpression of ETS leads the progression of tumors[153]. Therefore, inhibiting ETS oncogene signaling is a promising therapeutic strategy to treat prostate cancer.

Transcription factors are generally considered as undruggable targets. However, new strategies of modulating the activity of transcription factors have shown promise, including disrupting the interaction between transcription factors and other proteins andthe interaction between proteins and DNA, or restricting the binding of transcription factors by epigenetic modification of chromosome[256]. Currently, such as dithiophene diamidine compounds and DB1255 which inhibit ERG-DNA interactionsare under development[257]. In addition, clinical trials using optimized liposome-encapsulated siRNA to silence ERG expression in prostate cancer is ongoing[258].Another potential therapeutic strategy is to target the downstream effectors of TMPRSS2-ERG. Evidence shows that PLA2G7 is up-regulated in ERG-positive cancer and PLA2G7 silencing by siRNA sensitized ERG-rearrangement-positive VCaP cells to oxidative stress, reducing cell viability [259, 260].

In addition to ERG rearrangement, YK-4-279, a small molecule drug targeting the FLI1 rearrangement, has reached phase I clinical trials for Ewing sarcoma treatment [261]. YK-4-279 inhibits the binding of EWS-FL1 to RNA helicase and induces apoptosis of cancer cells.YK-4-279 also shows the ability to inhibit ERG and ETV1 rearrangements in vitro [262].

9.5 TP53

As the most common aberration, TP53 mutations are carried in 41% of pan-cancers, especially in HGSOC (high-grade serous ovarian cancer) (98%), esophageal adenocarcinoma (89%) and small cell lung cancer (85%). Almost 15% of mCRPCcarry TP53 mutation[263].

TP53 is a tumor suppression gene, encodes the p53 protein which maintain a low level by MDM2 regulated post-translational ubiquitin degradation in normal cells. When DNA is damaged, cellular stress induces phosphorylation of MDM2 and acetylation of p53, leading to accumulation and activation of p53. Activated p53 proteins stopproliferating cells in G1/S phase to repair the DNA damage[264]. Oncogenic stress triggers a DNA damage response involving p53, which constitutes a major barrier against tumor development. However, recent studies have shown that this effect of p53 is dispensable in tumors, and that p53 maintaining the homeostasis of cellular metabolism and redox balance in cell is even more important[265]. Moreover, many mutant p53 proteins have acquired gain-of-function (GOF) activities [266-268], which enable them to, for example, inactivate other p53 family members, in particular the tumor proteins p63 and p73 [269].(Figure 1.4)

In contrast to aberrations in tumor suppressor genes such as RB1, adenomatous polyposis coli (APC) and PTEN, most TP53 mutations are missense mutations [113].

Therefore, a drug development strategy for p53 is to use small molecules that promote proper folding and/or reactivation of common missense-mutant p53 proteins. Several of these compounds show significant anti-tumor activity in vitro and in vivo models. Clinical trials of two of the mutant-p53-targeting compounds are ongoing. APR-246 is being tested in phase II trial s[270], while the molecule COTI-2 is being studied in a phase I trial [271]. Although the current clinical trials do not include the treatment of prostate cancer, the efficacy of these drugs in the treatment of prostate cancer with TP53 mutation should also be investigated in the future.

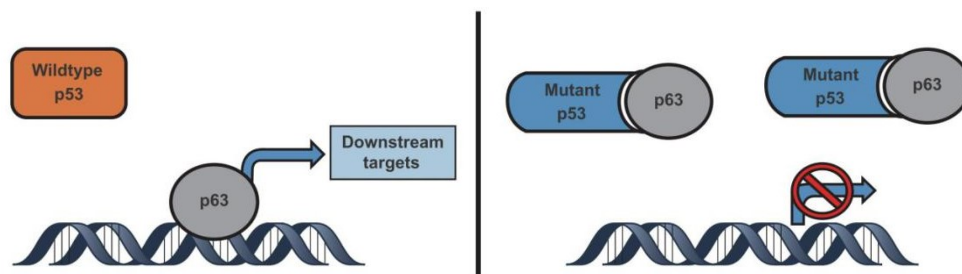

Figure1.4Mutant p53 proteins have acquired gain-of-function (GOF) activities, inactivate other p53 family members, in particular the tumor proteins p63 and inhibit the transcription activities of p63.

9.6 WNT signaling

Aberrations that result in WNT pathway activation, such as loss of function of APC (adenomatous polyposis coli protein), mutations in genes encoding beta-catenin and

mutations in RNF43 have been reported in 15% of the mCRPC[254]. RNA-seq revealed that WNT beta-catenin signaling is a functionally important pathway for androgen-independent prostate cancer progression [272, 273]. Therefore, targeting WNT signaling in the subset of the mCRPC with activation of this pathway is promising. There are multiple compounds engaged in clinical trials for solid tumors [274]. As well the efficacy of these drugs to treat mCRPC with WNT pathwayaberrations should also be investigated in the future.

9.7 The RAS-RAF-MEK signaling

Arguably less common in prostate cancer, but nevertheless still clinically relevant and potentially targetable, is oncogenic activation of RAS-RAF-MEK signaling [275, 276]. Such activation includes uncommon (1-2%) recurrent BRAF and RAF1 rearrangements as well as rare mutations of these genes [277] and other aberrations of genes activating this pathway, including HRAS, SPRED, SPROUTY, FGF, and FGFR[112, 113].

As ETS proteins are downstream effectors of RAS-RAF-MEK-extracellular signal-regulated kinase(ERK) signaling, resistance to AR blockade in ETS-rearrangedprostate cancer has been postulated to involve RAS-RAF-MEK signaling[278]. Activation of the MAPK pathway could also activate ETS signaling in

some ETS-rearrangement-negative tumors[279]. Studies of RAS-RAF-MEK inhibitors in CRPC are now needed to understand which subtypes of these cancers are driven by MEK, to further elucidate the importance of this pathway in CRPC.

10. Objectives

The incidence of prostate cancer in Europe also varies greatly in different regions, with a high incidence in Western and Northern Europe and a low incidence in Southern Europe. Sardinia is an island in the middle of the Mediterranean Sea, north to the European mainland and south to North Africa[2]. The incidence of prostate cancer in Sardinia is lower than in mainland Europe. Most recently Sardinian prostate cancer incidence is about 44 per 100,00, while Southern Europe is about 58 per 100,00 and Western and North Europe is up close to 100 per 100,00 [280, 281]. What's more, Chiang et al. demonstrated that Sardinia is a genetically isolated Mediterranean population and a purported refuge population of Neolithic ancestry. The evolutionary divergence from the European mainland population appeared to 143.3±1.3 generations (~4,300 years ago) which was much earlier than the divergence between Southern and Northern Europe [282]. Therefore, the genetic disparities may possibly explain of the disparities of prostate cancer incidence between Sardinian and other ethnic or populations.

Race and ethnicity are risk factors for prostate cancer. The effects of race and ethnicity on prostate cancer are not only reflected in different incidence but also in different frequencies of ETS family fusion in different groups. ETS family fusions is the most common alteration in prostate cancer of Caucasian men at a frequency of ~50%, however, they are lower in African Americans and Chinese at 20-30% and 10-20% in respectively. So far, most of the genomic prostate cancer studies are focused on cohorts of European ancestry, leaving minority groups underrepresented. Furthermore, in racial mixing, the ethnic contribution to risk is unclear. These problems pose a serious challenge to a compressive understanding of genetic risk, diagnosis and treatment of prostate cancer patients [133].

My aim is to determine the genomic landscape of prostate cancer of Sardinian population. This will better define prostate cancer risk management and treatment. It will also provide a better understanding of the effects of genetic factors on prostate cancer in the genetically unique Sardinian population.

My Specific Objectives ARE:

1.Characterize the somatic mutation and indels in prostate cancer of Sardinia

2.Characterize the Copy number variation in prostate cancer of Sardinia

3.Characterize the fusion events in prostate cancer of Sardinia

4.Characterize the germline risk in prostate cancer of Sardinia

Notes: To unveil the ethnic disparities of molecular basis of prostate cancer, we will also conduct a genomic study in Chinese prostate cancer cohort which is briefly summarized in the discussion chapter."A brief summary of a genomic study in Chinese prostate cancer cohort".

ChapterⅡ. Material and methods

1. Patients and Samples

This research was approved by the Ethical Committee of the Sassari AOU. Our study includes 30 patient tumor resections before ADT treatment. FFPE sections were stained with hematoxylin and eosin and reviewed by experienced pathologists to determine the Gleason score and mark a boundary betweentumor lesions and corresponding tumor adjacent tissues. Paraffin blocks were divided into two sections based on histological findings. DNA extraction was conducted separately on each divided paraffin block.

This research on Chinese prostate cohort was approved by the Ethical Committee of the Shantou University Medical College.

2. DNA extraction from FFPET tissue

QIAGEN Gene-Read DNA FFPE Kit enables purification of high-quality genomic DNA and removes artificial C>T mutations and was utilized to extract DNA from the formalin-fixed paraffin-embedded tissue (FFPET). All the procedures followed the manufacturer protocols. The concentration and quality of DNA were determined by Qubit3.0 and Agilent 2100. The qualified DNA with a fragment size more than 800 bp and the total amount of 100 ng was used for library preparation.

3. Whole Exome Sequencing Library Preparation

100~200ng DNA per sample was used for library construction using the KAPA Hyper-Plus Kit. Enzymatic fragmentation was performed according to the manufacturer instructions. Fragments with a size of 180-220bp before adapter ligation were selected for capture using the Roche Seq-Cap EZ Med-Exome system resulting in a total capture of 67Mb. Libraries were analyzed for size distribution by Agilent 2100 Bioanalyzer and quantified by real-time PCR. The qualified libraries were sequenced by the Illumina X-ten platform with 8 samples per lane.

4. Sequence data Quality control.

The original fluorescence image files obtained from X-ten platform were transformed to short reads (Raw data) by base calling and the short reads were saved in FASTQ format, containing sequence information and corresponding sequencing quality information. Reads were filtered as follows:1) Discard paired reads if either read contained adapter contamination (>10 nucleotides aligned to the adapter, allowing ≤ 10% mismatches); 2) Discard paired reads if more than 10% of bases are uncertain in either read; 3) Discard paired reads if the proportion of low quality (Phred quality <5) bases is over 50% in either read. All downstream bioinformatics analyses were based on the high-quality cleaned data.

5. Read mapping and processing.

Sequencing data was mapped to the reference human genome (UCSC hg19) using the Burrows-Wheeler Aligner (BWA) software to obtain the original mapping results in BAM format [283]. Then, SAMtools, Picard, and GATK tool kits were used to sort BAM files and duplicate marking, local realignment, and base quality recalibration to generate final BAM files for somatic and germline SNVs and indels calling[284, 285].

6. Somatic SNP and INDEL calling and annotation

Somatic SNVs were identified by GATK muTect1. Somatic INDels were identified by Pindel based on paired bam files that were generated from paired tumor and tumor adjacent normal tissues [286]. The minimal depth for high confidence SNVs was set as 10, while a depth of 20 was set for INDELs. Based on high sensitivity of the Pindel algorithm and the possible damage of DNA due to historical FFPETs (>5 years), single nucleotide deletions were removed from final results. Annotation of the somatic SNVs and INDELs was performed using Oncotator[287]. The annotated MAF files were used for downstream analysis.

7. Identification of mutation drivers

To obtain putative driver mutations in our Sardinian prostate cancer cohort,SNPs and indels were mapped to OncoKB database. [288]. OncoDriveCluster was used to predict significantly mutated genes in our Sardinia prostate cancer cohort [289] that may represent novel driver mutations. Novel mutations as well as the identified OncoKBc identified genes were mapped to cancerhotspots and the 3Dhotspots database, and then analyzed with MutationAssessor, SIFT, Polyphen2, and FATHMM [290-292]. Candidate driver mutations were defined as, listed in the Cancer hot spots, or 3D hot spots, or annotated as damage mutations by at least three of the four

methods, MutationAssessor, SIFT, Polyphen2, and FATHMM. Protein 3D structures were downloaded from the PDB database and virtualized by PYMOL [293].

8. Germline SNVs and INDELs calling and candidate germline risk identification

GATK-HaplotypeCaller was utilized to call germline SNVs and INDELs from tumor adjacent tissues [294]. Mutations with QUAL>200 were identified and annotated by ANNOVAR. Mutations with a MAF>0.01 in either 1000 Genomes or Exome Sequencing Project (ESP) databases were removed. Candidate germline risk mutations were then assessed as possible damage mutations using four methods, MutationAssessor, SIFT, Polyphen2, and FATHMM. Here we considered only mutations as damaging if identified by all four methods. Truncating mutations in tumor suppressor genes were considered as candidate germline risk mutations as well. Mutations identified in more than three samples or annotated by the FamilialCancerDatabase were considered as high confidence mutations [295].

9. Fusion calling, filtering, ORF prediction and visualization

Fusion event calling from raw read files of tumor adjacent tissues and tumor lesions was performed using FusionMap which is designed to detect and align fusion junction-spanning reads to the genome directly [296]. Fusions were filtered out if 1)

seed reads were <=3 or were listed in the in-family analysis or found in a paralog gene list, 2) fusions associated with uncharacterized genes, immunoglobin genes, mitochondrial genes or repeat regions, and 3) fusions that were reported in normal samples. High confidence fusions were input into FusionHub to search for reported fusions and predict the fusion effects. GO, KEGG and Reactome pathway enrichment analysis of fusion genes were performed by metascape[297].

10. Copy number variation calling, filtering and driver copy number variation identification

Control-FREEC was used to detect CNV with paired pileup files that were generated from unsorted bam files [298]. High confidence somatic CNVs were identified with a p-values less than 0.01 for both Wilcoxon and the Kolmogorov-Smirnov tests. They were next annotated by ANNOVAR [299]. Intergenicregions were then removed. Copy number variation reoccurring analysis was performed using GISTIC 2.0 [300].

Sardinian prostate cancer driver copy number variations already present in OncoKBannotatedalterations were classified as Putative copy number variations. Classification of Sardinian “candidate driver copy number variations” was based on curation in OncoKB onco-gene (amplification) or in OncoKB tumor suppressor gene (loss).

11. Integrative analysis of SNP, INDEL, fusion and copy number variation

The TCGA prostate cancer cohort dataset was downloaded from cBioPortal. Somatic SNPs, INDELs and copynumbervariations of the TCGA cohortandour Sardinian cohort were used to identify differential gene mutations, amplifications and deletions using the Fisher test. Adjusted P value of less than 0.05 was considered as significant.

12. Differential gene expression and genomic alterations in prostate cancer.

Differential expression of our candidate driver genes in matched TCGA normal samples and GTEx (Genotype-Tissue Expression Project) tissues was performed with GEPIA2 [301]. The frequency of the candidate driver alterations in other prostate cancer genomic studies was obtained from cbioportal[110].

13. PCR-based ERG fusion detection

Frequencies of different breakpoints for TMPRSS2-ERG fusions was based on the COSMIC database. We designed 5 pairs of primer using Primer5 for detection of the amplification and detection of TMPRSS2-ERG fusions based on the top five most frequent break points found in the COSMIC database. DNA of 19 historical paraffin

blocks of tumors from Sardinian prostate cancer patients was extracted and used for PCR screening forTMPRSS2-ERG.

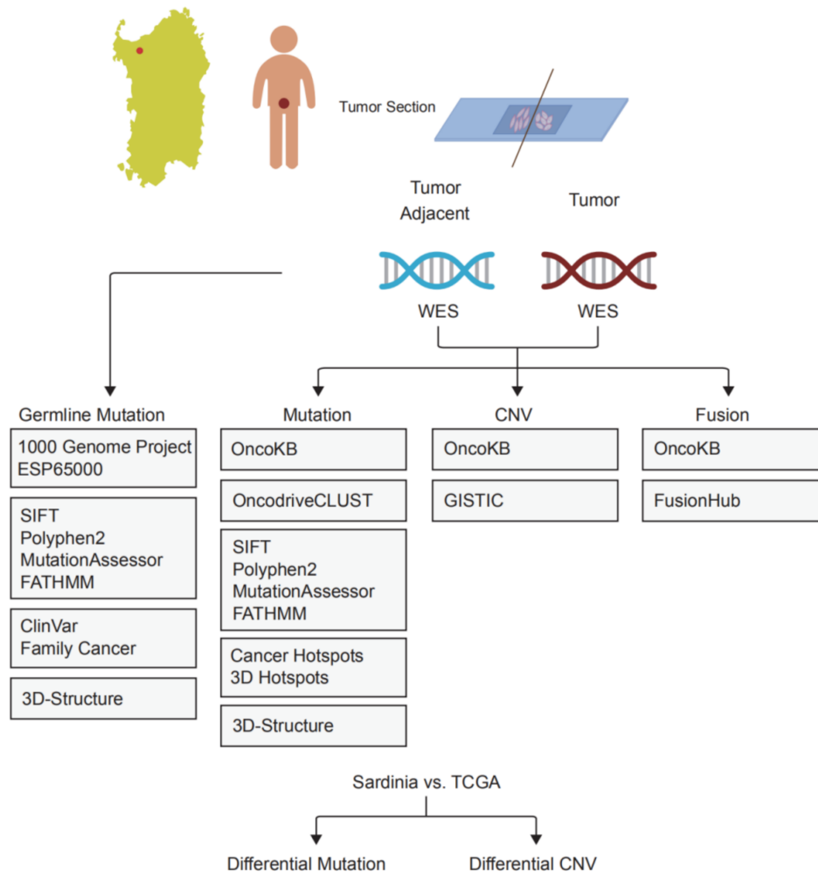

Figure 2.1 Flow chart of project of Genomic Landscape of Local Prostate Cancer in Sardinian Population

Chapter Ⅲ. Results: Genomic landscape of Localized Prostate Cancer in a Sardinian cohort

1. General summary of clinical and sequencing parameters

A total of 30 patients from the Pathology section of the Department of Experimental Medicine of the University of Sassari (Sassari, Sardinia, Italy) and diagnosed with prostate cancer during 2010 were included in the study. Most patients had levels of PSA<10 ng/ml and Gleason scores <=7（low risk, 23 out of 30 patients), six patients had PSA levels between 10 ng/mg ~20 ng/ml and Gleason scores<=7 (intermediate risk, 6 out of 30), one patient had a Gleason score of 9 (high risk, 1 out of 30). The age distribution of the patients ranges from 54 to 74. Among them, 20 patients were 60~69 years old, four patients were 54 to 59 years old, six patients were 70 to 74 years old. Survival records show that one patient died within one year after diagnosis, while another 4 patients died 5 years after diagnosis. All patients had prostate resection. Clinical information is summarized in **Table 1**.

Table 1 Clinical parameters of 30 prostate cancer patients

Clinical parameters	Number of Patients
Gleason Score	
3+3	6
3+4	13
4+3	9
4+3	1
4+3+5	1
PSA	
1~4	4
4~10	20
10~20	6
Age	
54~59	4
60~69	20
70~74	6
Risk Stratification	
low risk	23
intermediate risk	6
high risk	1
Suvival	
dead within 5 years	1
dead after 5 years	5
alive	24
ADT Treatment Naive	30

Paired tumor and tumor-adjacent FFPETs sections were obtained from historical paraffin blocks of tumors and adjacent normal tissue from resected prostates. DNA was extracted from paraffin sections and libraries constructed for whole exome sequencing. Sequencing coverage of tumor tissues was between 39-176, with the median of 67. The sequencing coverage of histologically normal tissue adjacent to tumor tissue was between 30-156, the median was 58 **(Figure3.1)**.

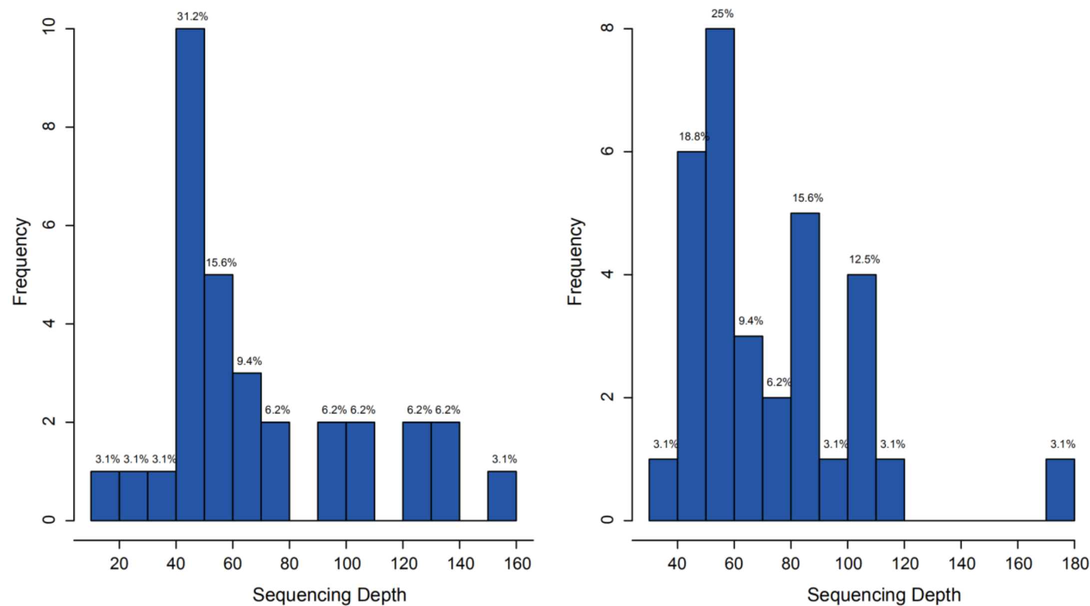

Figure 3.1Sequence depth of tumor and tumor adjacent tissues. A) Sequence depth distribution of tumor adjacent tissues. **B)** Sequence depth distribution of tumor tissues.

2. Somatic driver mutations in Sardinia prostate cancer

To identify somatic mutation in the Prostate Cancer cohort in Sardinia, we sequenced the whole exons of paired-samples from 30 patients. Somatic SNPs and Indels were called by GATK mutect1 and Pindel respectively. A total of 911 missense SNPs, 37 splice site mutations, 119 truncating mutations and 140 in-frame indels were identified in a total 30 tumor tissues. The number of somatic mutations per patient varied from 7 to 114. The median number was 28 mutations per patient. **(Figure 3.2)**

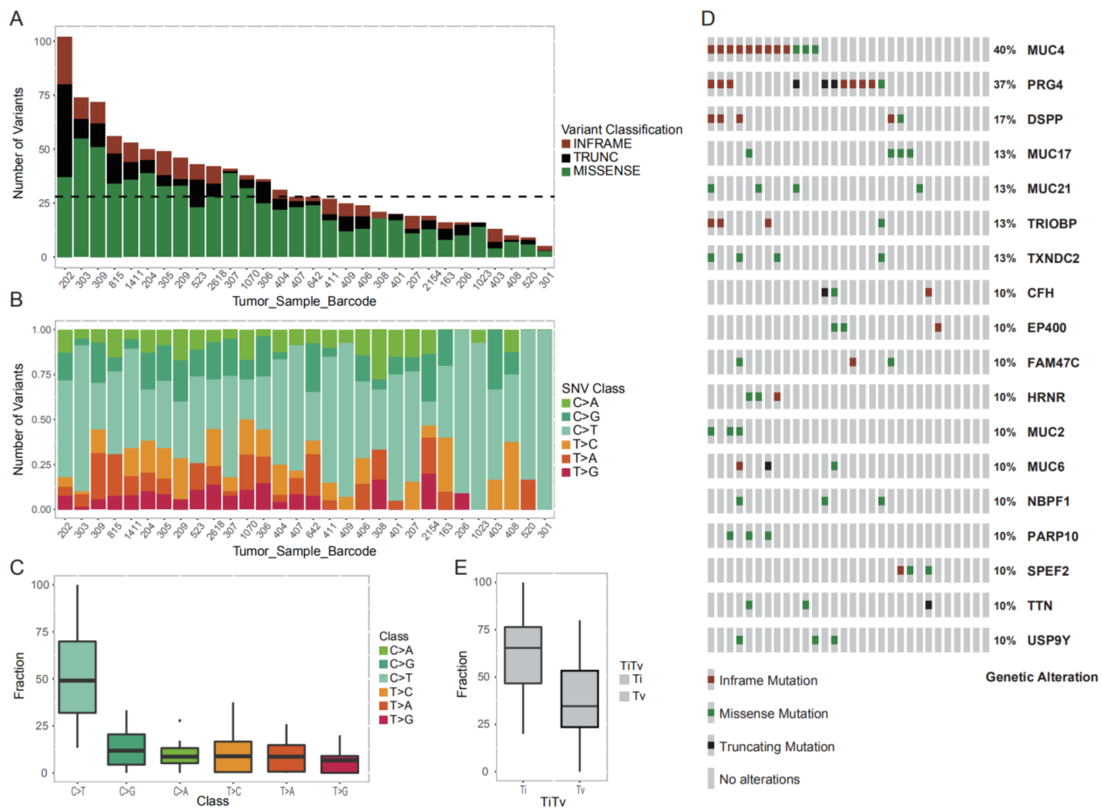

Figure 3.2 Summary of somatic mutation in 30 prostate cancer patients in Sardinia. A) Number of somatic mutations in each patient. **B)**Fraction of six different conversions in each sample.**C)** Overall distribution of six different conversions. **D)** Overall distribution of Transitions and Transversions.**E)** Oncoplot of gene with mutation in at least 3 of the 30 patients.

To identify potential driver mutations, we first mapped all mutations to the 4457 annotated alterations in the OncoKB database. We found 15 putative driver mutations present in 13 genes **(Figure 3.3)**. SPOP -Y87C and KMT2D-TRUNC mutations occurred in two samples. Single mutations were found for BRAF-G469A,

FLT3-R834Q, APC-TRUNC, BCOR-TRUNC, CDKN2C-TRUNC, FBXO11-TRUNC, KEAP1-TRUNC, FBXO11- TRUNC, NF1-TRUNC, TP53-TRUNC, RBM10-TRUNC, ZFHX3-TRUNC **(Figure 3.5)**.

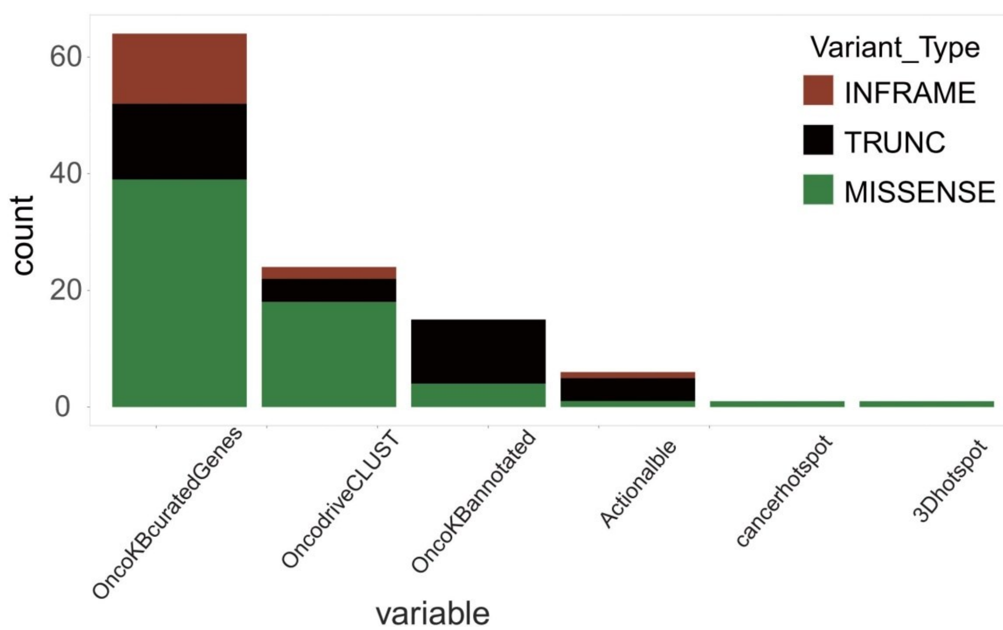

Figure 3.3 All the 1067 somatic mutations of 30 prostate cancer patients in Sardinia were annotated as clinical actionable mutations, putative drivers, cancer hot spot mutations and 3D hot spot mutations,respectively. Novel drivers were predicted by OncoDrieverCluster. Five mutations are clinical actionable, 15 mutations are cancer putative drivers (OncoKB annotated alterations). Sixty-five mutations are observed in OncokB curated genes and 22 mutations are observed in significant mutated genes that predicted by OncoDriverCluster.

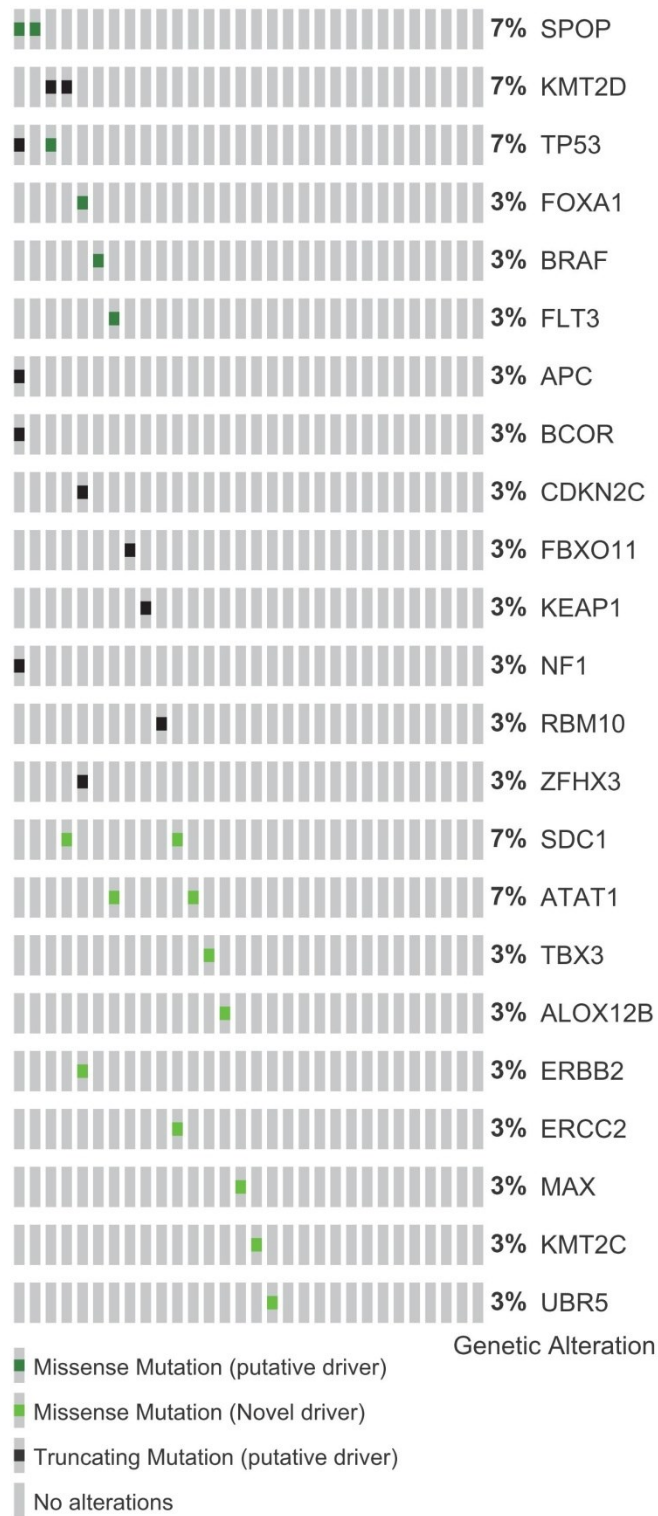
7% SPOP
7% KMT2D
7% TP53
3% FOXA1
3% BRAF
3% FLT3
3% APC
3% BCOR
3% CDKN2C
3% FBXO11
3% KEAP1
3% NF1
3% RBM10
3% ZFHX3
7% SDC1
7% ATAT1
3% TBX3
3% ALOX12B
3% ERBB2
3% ERCC2
3% MAX
3% KMT2C
3% UBR5
Genetic Alteration
Missense Mutation (putative driver)
Missense Mutation (Novel driver)
Truncating Mutation (putative driver)
No alterations

Figure 3.4 Oncoplot of putative and novel somatic drivers of prostate cancer from 30 Sardinian local prostate cancer samples wereobtainedbywholeexomesequencingof paired tumorandtumor-adjacent tissues. Each green dot represents a patient with a putative missense mutation, black dot represents a patient with putative truncating mutation, purple dot represents a patient with fusion of the specified gene.

In addition, we found that TP53-F106C was a cancer hot spot mutation and FOXA1-S250P was a 3Dhotspot mutation **(Figure 3.3)**. We next mapped these mutations to the CIVIC database of actionable mutations and found BRAF- G469A, NF1-TRUNC, APC-TRUNC and KMT2D-TRUNC clinically actionable mutations **(Table2)**.

Table 2 Clinical actionable mutations in 30 prostate cancer patients

Patient	Gene	Type of mutation	Drug-Sensitivity
401	BRAF	p.G469A	Vemurafenib, Cetuximab, Erlotinib
202	NF1	p_1172fs	Dabrafenib，Binimetinib，JQ1
202	APC	p_1212fs	G007-LK,JW55
303	KMT2D	p.GSYTDPYAQPPL2372fs	AR-42
305	KMT2D	p.QEPPP2350fs	AR-42

To identify potential novel driver candidateswhich were not identified in the OncoKB database, we next used the OncoDriverCluster algorithm to predict significantly

mutated genes in our cohort. A total of 12 genes including the putative driver SPOP were found to be significant in our cohort **(Figure 3.4)**.

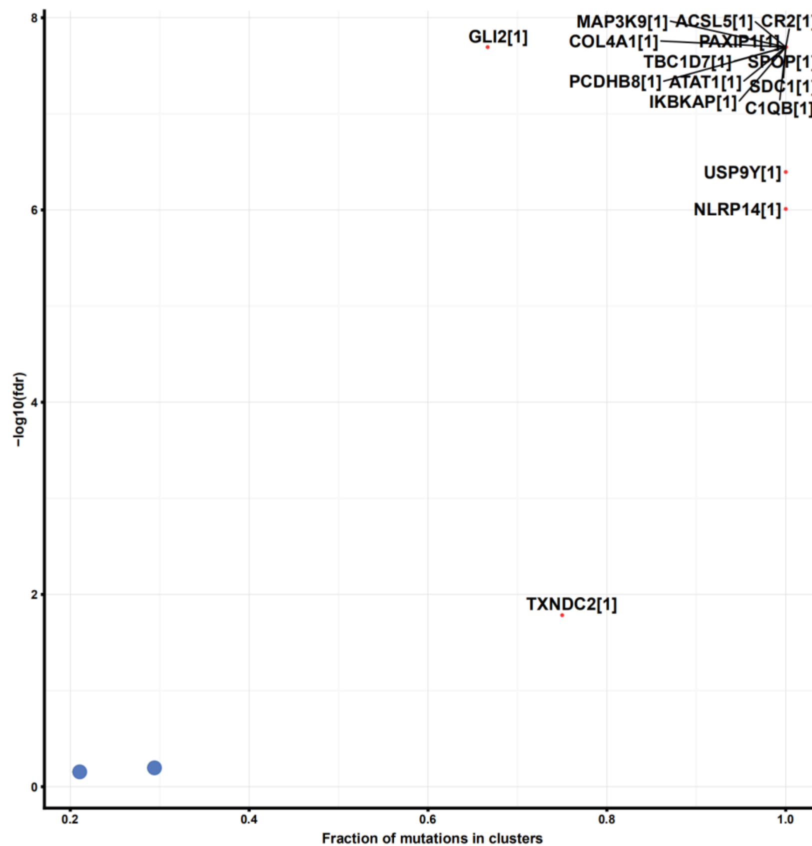

Figure 3.5 All the somatic mutations of 30 Sardinian prostate cancer were input in OncoDriverCluster algorithm to predict significantly mutated genes in our cohort. The y-axis is the log10(fdr), the x-axis represents that the fraction of mutations in a particular gene were predicted as cancer driver.

Mutations in these 12 genes and mutations in OncoKB curated genes were annotated for functional disruption using MutationAssessor, SIFT, Polyphen2, and FATHMM. Nine mutations were predicted to be deleterious mutations by at least three methods, and four of them were simultaneously predicted to be deleterious mutations by all the four methods **(Figure 3.6)**.

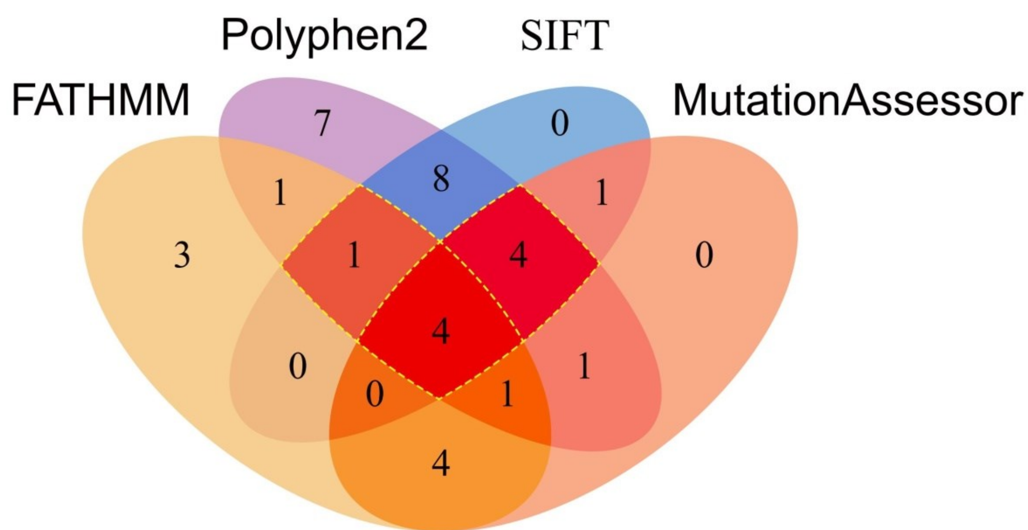

Figure 3.6 Four mutations of all the 1067 somatic mutation of 30 Sardinian prostate cancer patients have been annotated as damage mutations by all of the following, MutationAssessor, SIFT, Polyphen2, FATHMM methods, another 5 mutations have been annotated as damage by at least three of the four methods.

Six of the 9 domains containing the above mutations have known 3D structures. They include, the ALOX12B protein (R422W, **Figure 3.7A)**, the ATAT1 protein (D19V

mutation, **Figure 3.7B)**, the ERBB2 Receptor L domain harboring the C53R mutation **(Figure 3.7C)**, the ERCC2 protein (R497C mutation, **Figure 3.7D)**, the MAX protein (D87N mutation**, Figure 3.7F)**, the TBX3 T-box domain (V202D mutation, **Figure 3.7H)**.

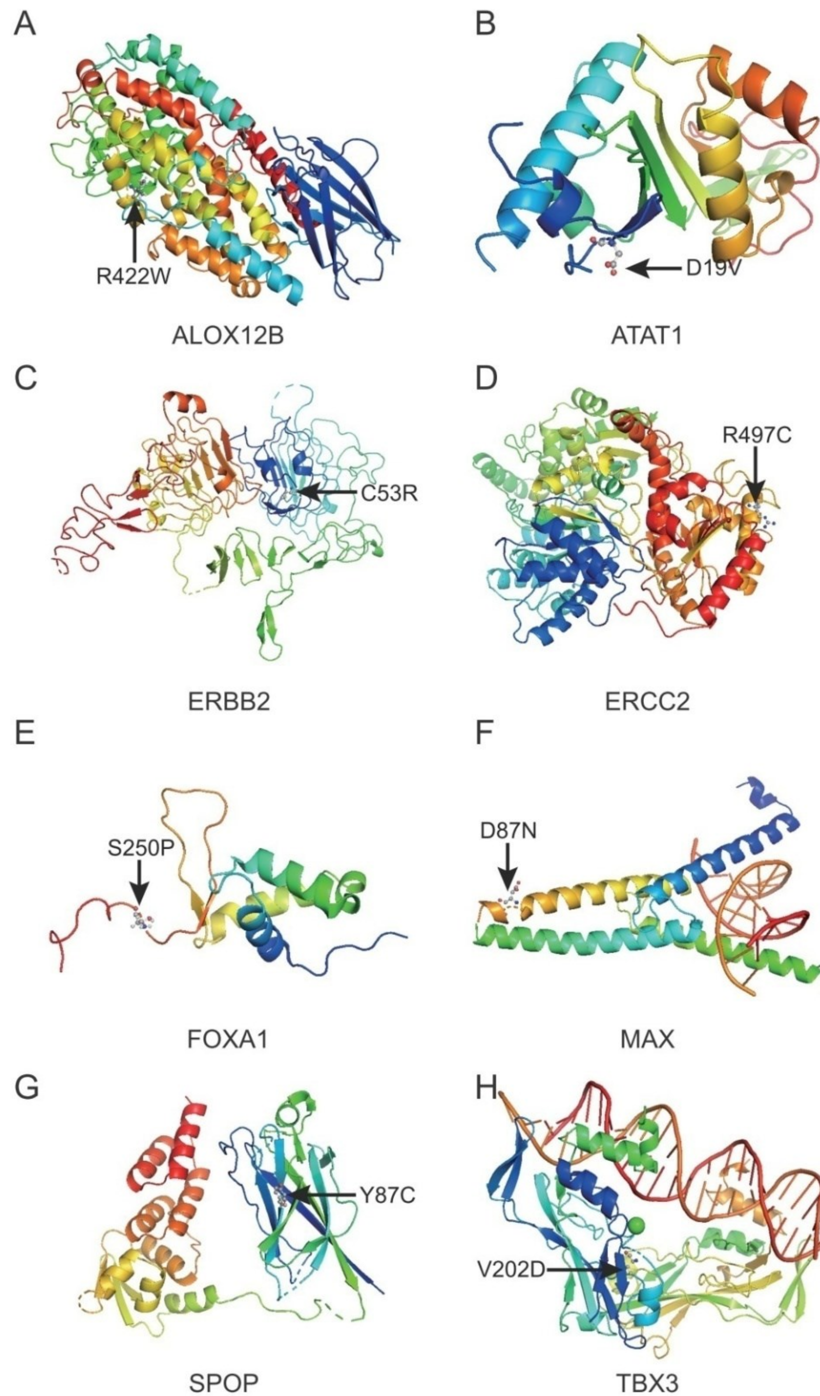

Figure 3.7 A) ALOX12B Lipoxygenase domain R422W mutation.**B)** ATAT1 D19V mutation.**C)** ERBB2 Receptor L domain C53R mutation.**D)** ERCC2 R497C

mutation.**E)** FOXA1 Forkhead domain S250P. **F)** MAX D87N mutation. **G)** SPOP MATH domain Y87C.**H)** TBX3 T-box V202D mutation.

3. Somatic copy number variation in Sardinia prostate cancer

Control-Freec was used to analyze somatic copy number variation in paired samples of 30 prostate cancer patient in Sardinia. A total of 784 segments of copy number variation events were identified, including 9520 gene-level events (2920 gene gains, 6600 gene losses). We found 444 gene amplifications (copy number variation is greater than or equal to 2) and 2135 deletions (copy number variation is less than or equal to -2).**(Figure 3.8)**

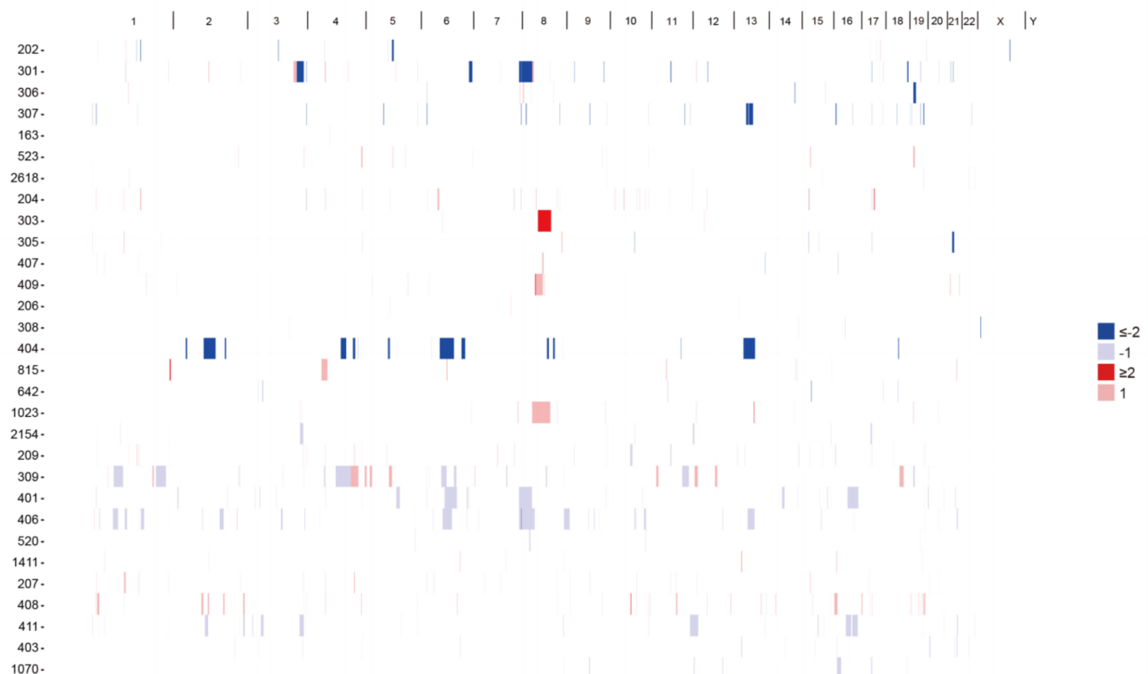

Figure 3.8 Control-Freec was used to analyze somatic copy number variation in paired samples of 30 prostate cancer patient in Sardinia. Each row represents a patient. And the column represents the chromosomes. Each dark red bar represents a patient with amplification of the region. Each dark blue bar represents a patient with deep deletion of the region. Each light red bar represents a patient with gain of the region. Each light blue bar represents a patient with loss of the region.

To identify the copy number variations that may act as "oncogenic drivers", we first mapped copy number variations to the 4457 OncoKB annotated alteration database. Six putative oncogene gain events were identified, involving six genes (BCL6, ROS1, CDK6, EGRF, ETV1, NTRK1), affecting 5 different tumor samples. Thirty-eight putative tumor suppressor gene loss events were identified, involving 30 genes, affecting 11 samples. Among them, CDKN1B-loss, DUSP4-loss, and PRDM1-loss occurred three times each, and PTEN-loss and RB1-loss occur twice. **(Table 3)**

Table 3 Clinical actionable copy number variation in 30 prostate cancer patients

Patient	Gene	Type of mutation	Drug-Sensitivity	Drug-Resistance or Non-Response
815	TYMS	Amplification	Pemetrexed	Pemetrexed
2154	BRCA1	Deletion	Olaparib,CX-5461	
309	EGFR	Amplification	Cetuximab, Panitumumab	Osimertinib, Rociletinib
309	FBXW7	Deletion	Rapamycin (Sirolimus)	
404	NBN	Deletion	GPI-15427	
305	PTEN	Deletion	Carboplatin, Buparlisib	Everolimus, BYL719
406	PTEN	Deletion	Carboplatin, Buparlisib	Everolimus, BYL719
404	RASA1	Deletion	Trametinib	

2154	TSC2	Deletion	MTOR Inhibitors	
307	RB1	Deletion	Palbociclib（PD0332991)	
404	RB1	Deletion	Palbociclib（PD0332991)	
301	STK11	Deletion	NA	Docetaxel, Selumetinib
401	ATXN1L	Deletion	NA	Vemurafenib, Dabrafenib
411	ATXN1L	Deletion	NA	Trametinib

In order to identify new potential driver copy number variation events, we performed copy number replay analysis using GISTIC 2.0. A total of 15 significant amplification intervals were identified along with the identification of 20 significant loss intervals **(Figure 3.9)**.

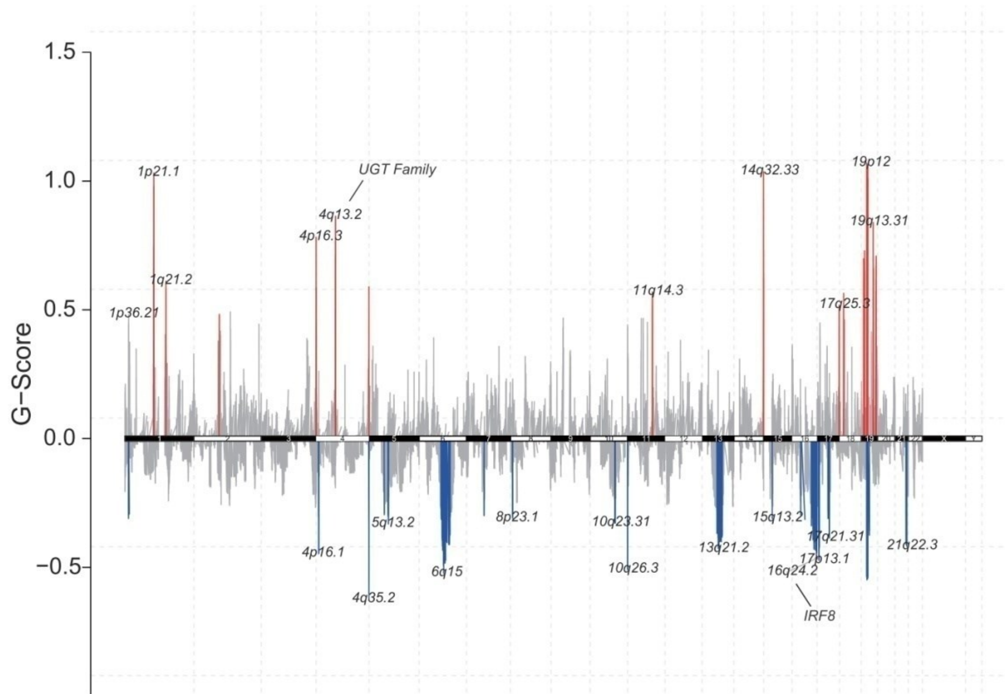

Figure 3.9G-score across the whole exome region of 30 Sardinian prostate cancers. G-scores were calculated by the GISTIC algorithm to determine the amplitude and the frequency of copy number variation.

Among the loss intervals, 6q15, 13q21 and 16q24 showed large-scale deletions. Further, we found that the 15 significant amplification regions did not contain a known onco-gene, but the 20 significant loss regions contained 21 OncoKB curated TSGs. Of the total 46 copy number deletion events, 15 deletion events contained 8 genes that were previously annotated as putative drivers.The other 31 deletion events involved 13 curated tumor suppressor genes, including BACH2-loss (4 events), NKX3-1-loss(4 events), ATP6V1B2-loss (3 events), EPHA7-loss(3 events), ESCO2-loss (3 events), SESN1-loss (3 events), TP63-loss (3 events), ERCC3-loss (2 events), IRF8-loss (2 events), FAS-loss (1 events), RECQL-loss (1 events), SDHC-loss (1 events), TBL1XR1-loss (1 events) **(Figure 3.10)**.

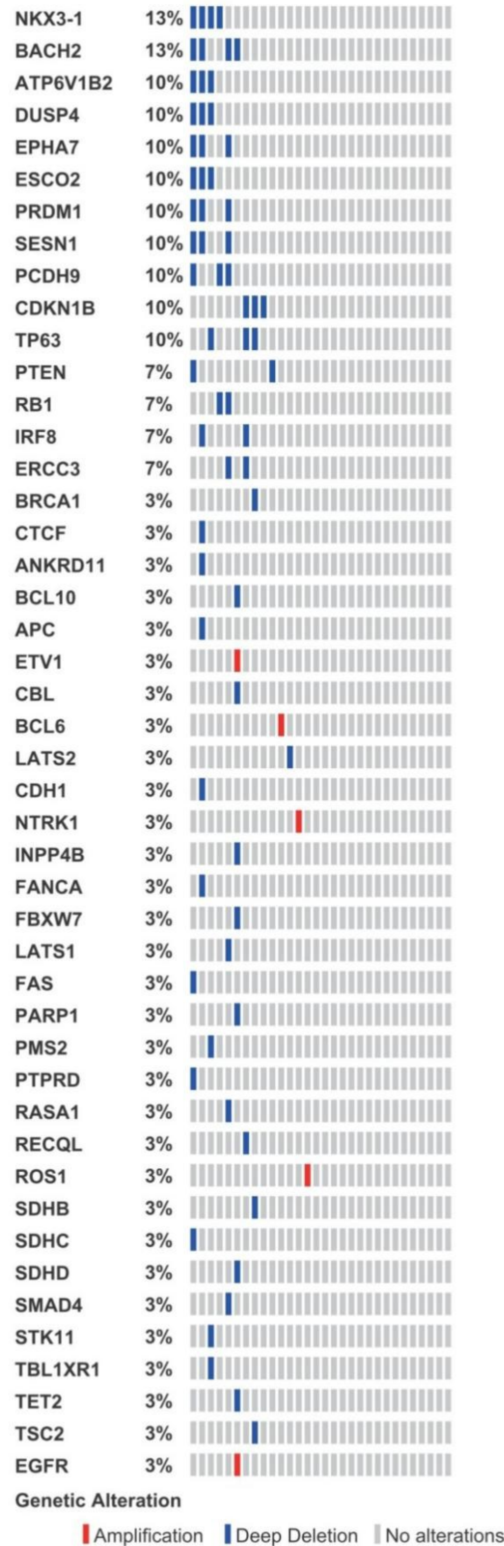
NKX3-1 13%
BACH2 13%
ATP6V1B2 10%
DUSP4 10%
EPHA7 10%
ESCO2 10%
PRDM1 10%
SESN1 10%
PCDH9 10%
CDKN1B 10%
TP63 10%
PTEN 7%
RB1 7%
IRF8 7%
ERCC3 7%
BRCA1 3%
CTCF 3%
ANKRD11 3%
BCL10 3%
APC 3%
ETV1 3%
CBL 3%
BCL6 3%
LATS2 3%
CDH1 3%
NTRK1 3%
INPP4B 3%
FANCA 3%
FBXW7 3%
LATS1 3%
FAS 3%
PARP1 3%
PMS2 3%
PTPRD 3%
RASA1 3%
RECQL 3%
ROS1 3%
SDHB 3%
SDHC 3%
SDHD 3%
SMAD4 3%
STK11 3%
TBL1XR1 3%
TET2 3%
TSC2 3%
EGFR 3%
Genetic Alteration
Amplification
Deep Deletion
No alterations

Figure 3.10Oncoplot of amplifications in oncogenes and deletion in tumor suppressor genes. Each red bar represents a patient with amplification of the specified gene on the left. Each blue bar represents a patient with deep deletion of the specified gene.

We also examined the frequency of the deletion of the genes outlined above in other prostate cancer genomic studies and found that IRF8-loss occurred in 2-8% of prostate cancer patients [104, 254, 302-305]**(Figure 3.11).** Patients with IRF8-loss have worse survival outcome compared with patients without IRF8-loss **(Figure 3.12)**. These data indicated that IRF8 deletion is a potential important driver in prostate cancer.

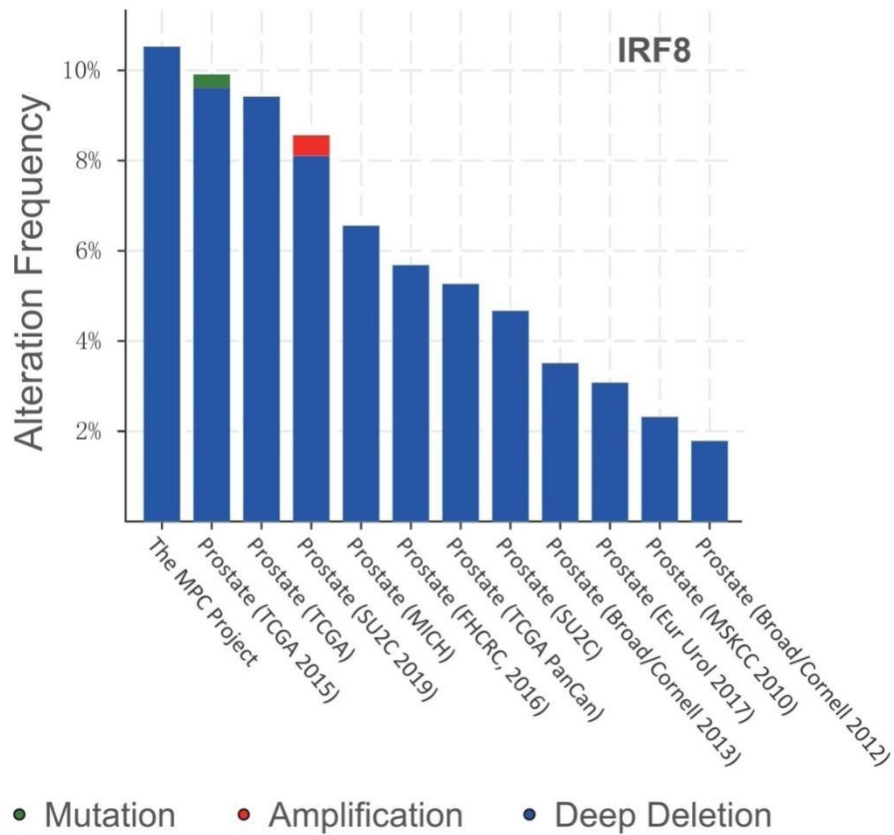

Figure 3.11 Frequencies of IRF8 deletion across the prostate cancer genomic studies that documented in cBioportal database.

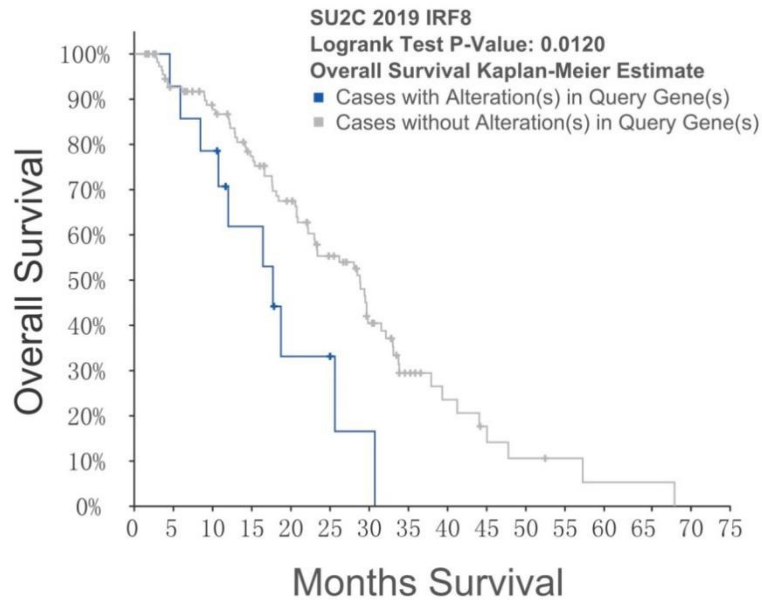

Figure 3.12 Patients with IRF8 deletion have decreased survival.

4. Comparison of gene mutations between prostate cancer in the TCGAdatabase and Sardinia cohort

To identify similarities and differences in prostate cancer mutations between Sardinia and patients of North America European ancestry, we had integrated the somatic mutation and copy number variation data from our Sardinian cohort and made a gene-level comparison to the TCGA prostate cancer dataset. The results reveal that MUC4, PRG4,DSPP SNPs and indels, 4q13.2 (UGT family genes) and 4q16.3 (ZNF595, ZNF718) amplifications, 8q23.1 (USP17L1/2/3) and 4q35.2 (DUX4) deletions are significantly enriched in the Sardinian prostate cancer cohort compared to the TCGA cohort (Fisher test, adjusted *p* value<0.05)**(Figure 3.13-3.15)**. We also

observed that both the Sardinian cohort and the TCGA data set had similar frequencies and types of tumor suppressor gene deletions **(Figure 3.15)**. MUC4 and PRG4 are large glycoprotein and mutations in them are possible passengers. The 4q16.3 amplification and 4q35.2 deletion only affect very small region of the chromosome. 8q23.1 deletion overlap with TCGA 8q23.1 deletion but have border expansion. 4q13.2 amplification including multiple UGT family members (Figure 4.1). The pick gene is UGT2B4 which is tend to have high expression in TCGA tumor samples rather than normal samples(Figure 4.3) and patients the high expression of UGT2B4 have better disease-free survival (p=0.047) (Figure 3.17). Co-expression genes of UGT2B4 and enriched in Steroid hormone biosynthesis pathways.

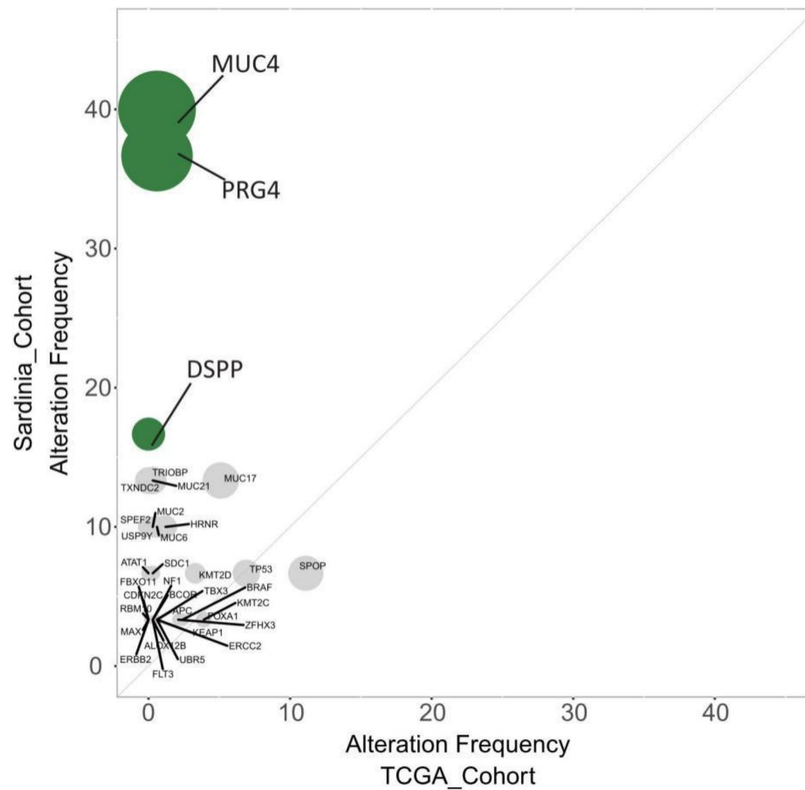

Figure 3.13 The TCGA prostate cancer cohort dataset was downloaded from cBioPortal. Somatic SNPs, INDELs of the TCGA cohortandour Sardinian cohort were used to identify differential gene mutations frequencies using the Fisher test. Adjusted P value of less than 0.05 was considered as significant.Mutations in MUC4, PRG4 and DSPP are significantly increased in Sardinia prostate cancer cohort compared to the TCGA cohort (*p* value < 0.05).

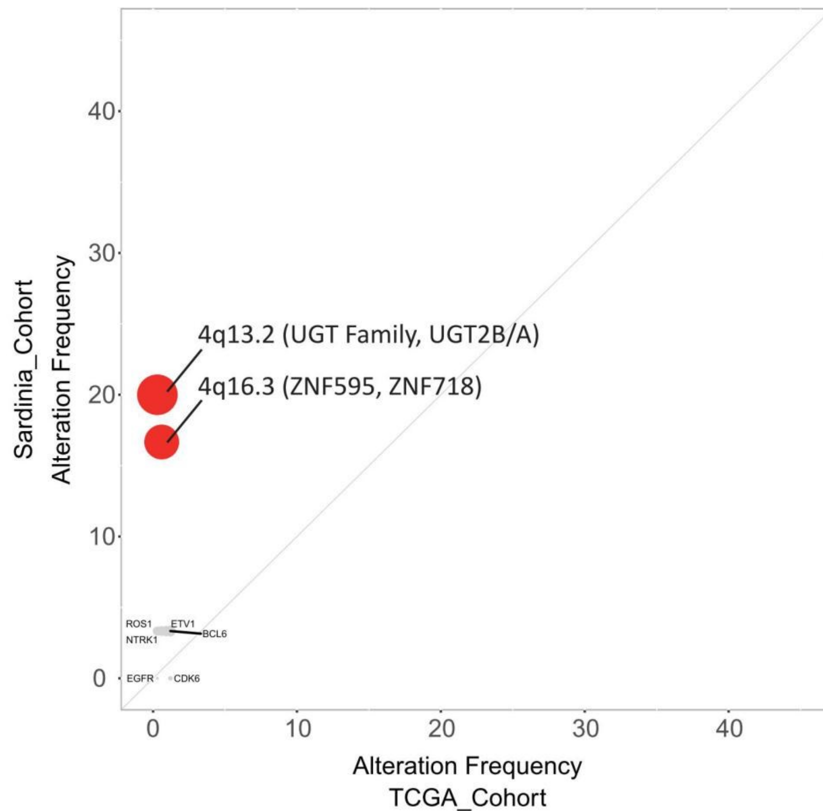

Figure 3.14 The TCGA prostate cancer cohort dataset was downloaded from cBioPortal. Amplifications of the TCGA cohortandour Sardinian cohort were used to identify differential gene mutations, amplifications frequencies using the Fisher test. Adjusted P value of less than 0.05 was considered as significant. 4q13.2 and 4q16.3 amplifications are significantly greater in the Sardinia prostate cancer cohort compared to the TCGA cohort (p value < 0.05).

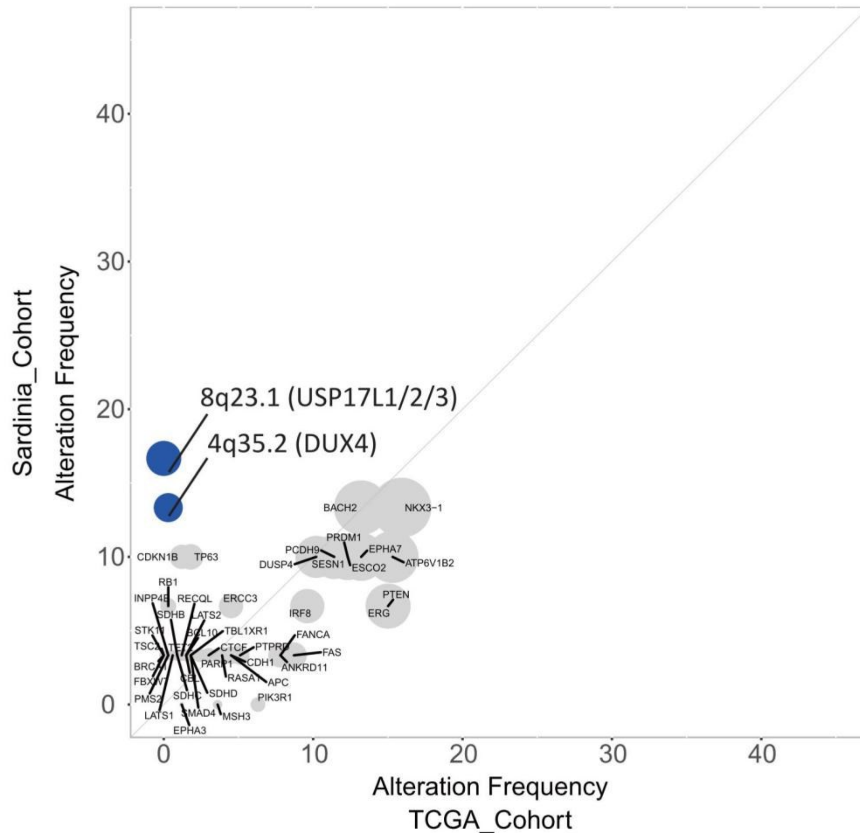

Figure 3.15 The TCGA prostate cancer cohort dataset was downloaded from cBioPortal. Amplifications of the TCGA cohortandour Sardinian cohort were used to identify differential gene mutations, deletions frequencies using the Fisher test. Adjusted P value of less than 0.05 was considered as significant. 8p23.1 and 4q35.2 deletions are significant greater in the Sardinia prostate cancer cohort compared to the TCGA cohort (*p* value < 0.05).

To further understand the potential impact of these genes on tumor development, we examined the gene expression profiles in the TCGA and the Genotype-Tissue

Expression (GTEx) databases for differential expression between prostate tumors and normal tissue. Expression profiles revealed that MUC4 is significantly down-regulated in prostate cancer tissue **(Figure 3.16).** Furthermore, patients with high expression of PRG4 have poorer disease-free survival **(Figure 3.17)** compared to patient with tumors having lower expression of PRG4. These indicate that mutations in MUC4 and PRG4 may be more than just passenger mutations. UGT gene cluster represents a family of glycosyltransferases which are able to transfer estradiol to estradiol glucuronide and are involved in the regulation of estrogen metabolism **(Figure 3.18)**[306-308].

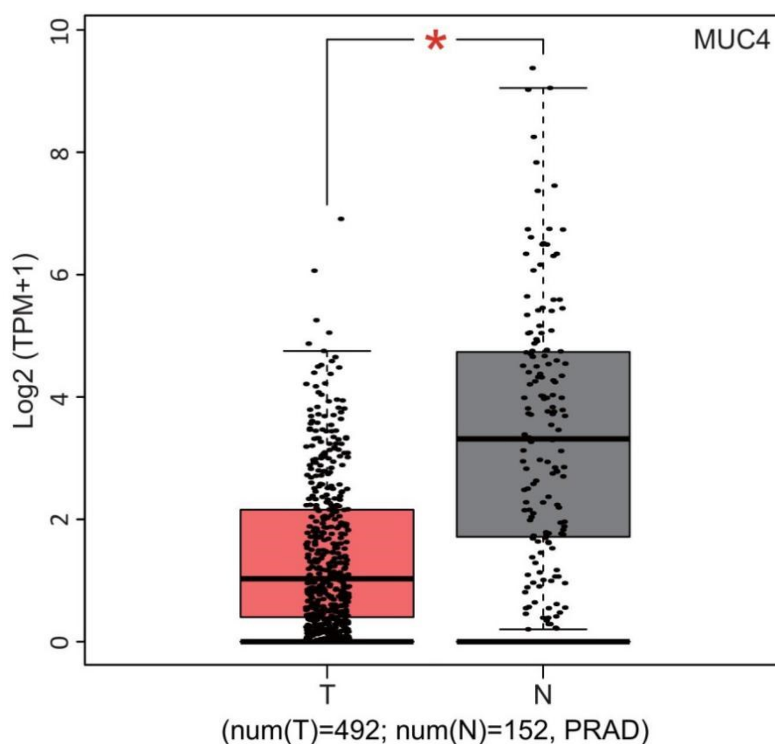

Figure 3.16MUC4 is significantly down-regulated in prostate cancer tissues (GEPIA dataportal). Expression of MUC4 was determined in 43 normal (green bar) and 497 prostate cancers (red bar) from the TCGA database. MUC4 was found to be significantly upregulated ($p<0.05$) in localized prostate tumors compared with tumor adjacent normal tissue.

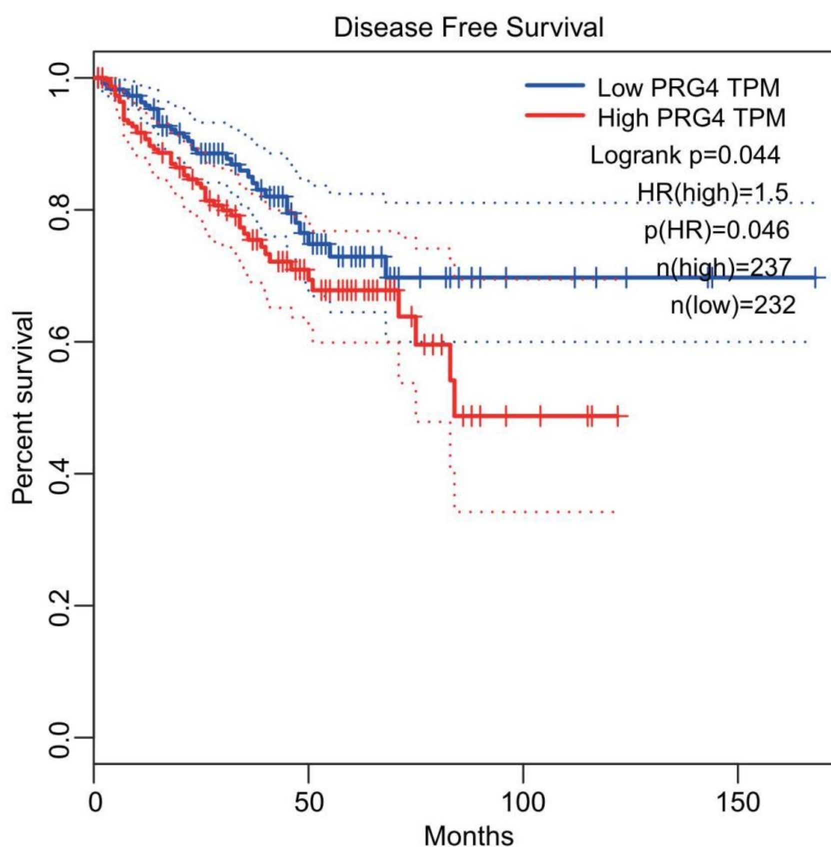

Figure 3.17 Low expression of PRG4 have better disease-free survival (GEPIA dataportal) compared to normal expression. All the 497 TCGA localized prostate cancer patients were divided into a PRG4 low expression group and a PRG4 high

expression group based on the PRG4 median expression. Between-group comparisons of DFS were performed by the Kaplan-Meier method and the log-rank test.

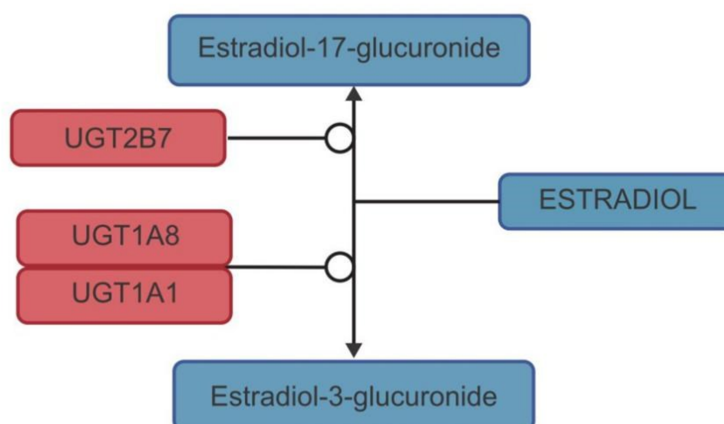

Figure 3.18 UGT family is important genes in super pathway of estrogen metabolism. (https://pathcards.genecards.org/card/estrogen_metabolism)

5. Novel BTBD7-SLC2A5 fusions and ETS family status in prostate cancer of Sardinia

We next used FusionMap to identify fusion events in whole exome data. We detected 15 in-frame fusions and 44 out of frame fusion genes in the 30 patients. Of the fusion events, only two pairings, MAGOH->TMEM220 (out of frame) and BTBD7->SLC2A5 (in frame), occurred in more than one sample, 5 and 4 times

respectively **(Figure 3.19)**.BTBD7-SLC2A5 contains exons 1-10 of BTBD7 and exons 2-8 of SLC2A5. N-terminus of the fusion protein contains the BTB/POZ domain of BTBD7, and the C-terminus contains Transmembrane region of SLC2A5 **(Figure 3.20)**.

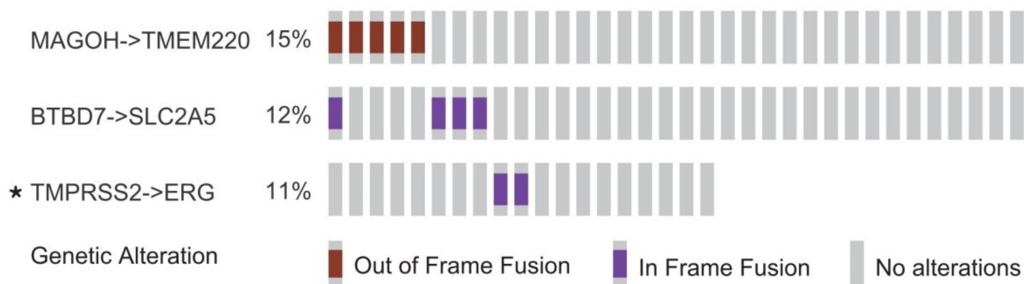

Figure 3.19 Oncoplot of fusion genes that are altered in more than 2 or more patients (ERG fusions were detected based on PCR methods and the remaining fusions were detected by FusionMap software based on WES data).

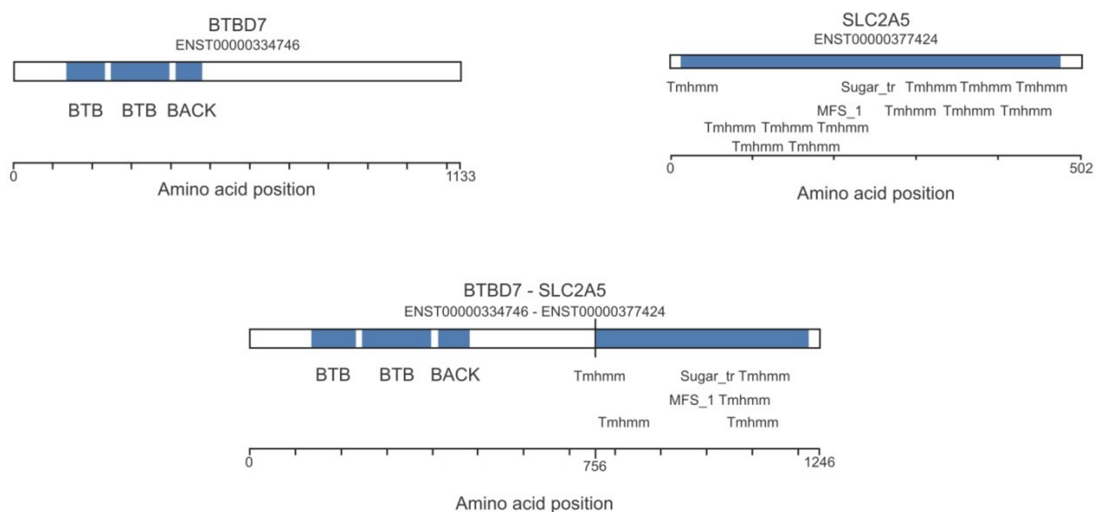

Figure 3.20 Predicted transcript and protein of the BTBD7-SLC2A5 fusion determined by Fusion hub. The predicted protein contained the entire transmembrane domain and fructose transporter domain of SLC2A5 and the BTB domain of BTBD7.

ETS family fusions are the most common fusion event associated with prostate cancer. ERG-TRMPSS2 gene fusions are found in 46% of prostate cancers, followed by ETV1 fusions in 8% and ETV4 fusions in 1% of the patients [254]. In our whole exon data of 30 Sardinian prostate cancer patients, none of the ETS fusions was detected. Possibly due to the low sensitivity of whole exon data to detect fusion events, we also used PCR to detect ERG fusions at the RNA level in 19 samples. ERG fusions were detected in only two of the 19 samples (11%) **(Figure 3.19).** The frequency of ERG fusion is still lower than the expected frequency found in patients with European ancestry. However, we did observe a significant deletion of 22q23.1 in the copy number reoccurring analysis **(Figure 3.9 & Figure 3.15).** Considering the ERG fusion event is often co-occurrence with ERG deletion, the frequency of ERG fusion events in Sardinia prostate cancer may be higher at 11%.

6. Germline risk mutations in the patients

In order to identify germline risk mutations for prostate cancer in the Sardinian prostate cancer cohort, we used the GATK best practice pipeline to perform germline SNP and INDEL calling with fastq files that generated from tumor adjacent tissue. We filtered out the mutations with a calling quality less than 200 and MAF>0.1. The remaining mutations were evaluated for the effect on protein structure/function using MutationAssessor, SIFT, Polyphen2, and FATHMM**.** A total 113 mutations were annotated as deleterious mutations by all the four methods Among these, we found a variation in ARSD-G320D **(Figure 3.21 & Figure 3.22)** affecting 53% of the patients and that in CNN2-G250V-G263S **(Figure 3.21 & Figure 3.22)** affecting 9% of the samples. In addition, four genes that were previously recorded in the Family cancer database, which contains approximately 500 cancer related hereditary disorders, have one mutation each for CHD1 (S847T), MSH2 (A382C), HPS6 (L498P), and BMPR1A (A319G).

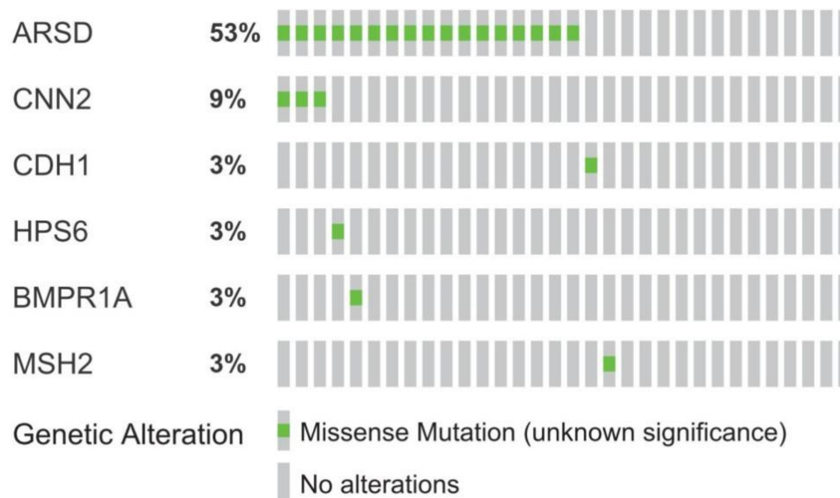

Figure 3.21 Oncoplot of germline risk mutations. Each green dot represents a patient with a germline missense mutation in a specified gene on the left.

We next examined expression of ARSD and CNN2 in prostate cancer by looking at prostate cancer gene expression profiles in the TCGA and the Genotype-Tissue Expression (GTEx) databases. We found that ARSD tends to be highly expressed in tumor tissues (data not show). Patients with high expression of ARSD have better disease-free survival than those with lower expression of ARSD **(Figure 3.23)**.

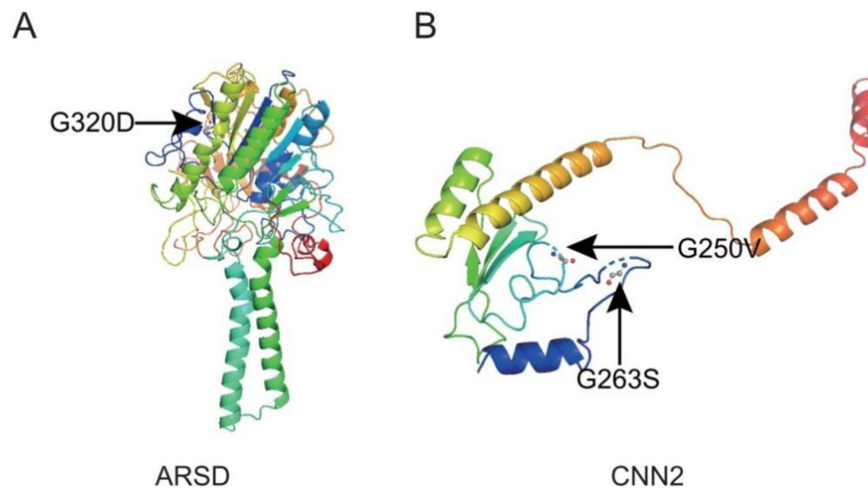

Figure 3.22 A) ARSD Sulfatase domain with the G320D mutation.**B)** CNN2 calponin domain with G250V and G263S mutations.

Considering that ARSD regulates estrogen metabolism **(Figure 3.25)**, it is possible that tumors with high ARSD expression or expressing variants of ARSD impact castration therapy (homo-therapy) hence leading to better or differential (variant) disease-free survival. CNN2 is significantly down-regulated in tumors compared to normal tissue acting as a possible tumor suppressor gene **(Figure 5.24)**. These results indicate that ARSD and CNN2 germline mutations are potentially associated with prostate cancer development or therapeutic impact in Sardinia.

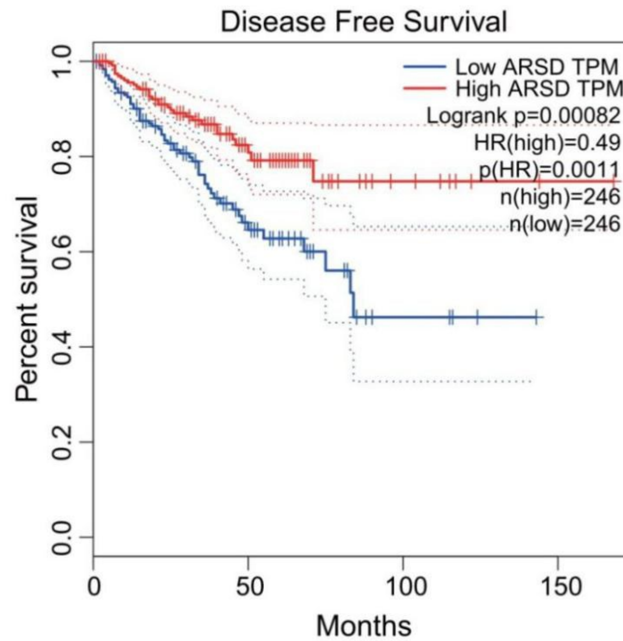

Figure 3.23 Patients with high ARSD expression have better disease-free survival (GEPIA dataportal). All the 497 TCGA localized prostate cancer patients were divided into a ARSD low expression group and a ARSD high expression group based on the ARSD median expression. Between-group comparisons of DFS were performed by the Kaplan-Meier method and the log-rank test.

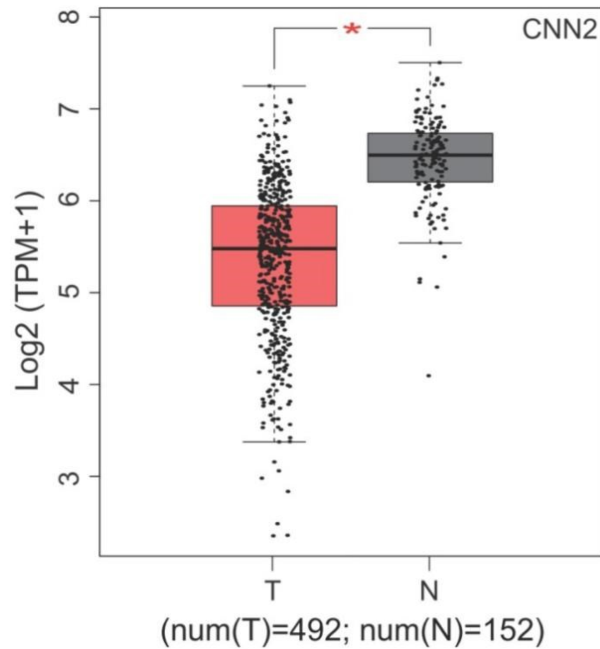

Figure 3.24 CNN2 significantly down-regulated in prostate cancer tissues (GEPIA dataportal). Expression of CNN2 was determined in 43 normal (green bar) and 497 prostate cancers (red bar) from the TCGA database. CNN2 was found to be significantly upregulated ($p<0.05$) in localized prostate tumors compared with tumor adjacent normal tissue.

Moreover, the data show that there is a tendency of mutual exclusivity (p=0.062) between the germline risk variants ARSD-G320D and somatic UGT family amplification (Figure 5G).

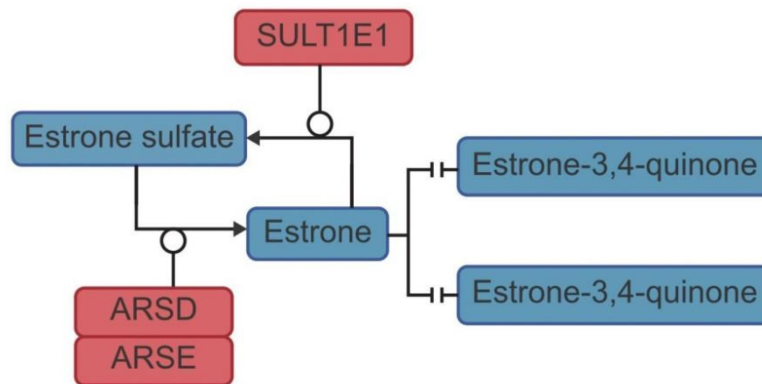

Figure 3.25 ARSD is one of the genes in super pathway of estrogen metabolism. (https://pathcards.genecards.org/card/estrogen_metabolism)

ChapterⅤ. Discussion

Emerging evidence indicates that there are remarkable disparities in prostate cancer epidemiology as well as the molecular landscape among different ethnic groups of European, North America African and Asian origin [4, 133]. The Sardinia population is an isolated Mediterranean population with evolutionary divergence from the European mainland population taking place some 143.3±1.3 generations in the past. Furthermore, prostate cancer incidence is lower in the Sardinia population compared with mainland Europe [281, 282]. However, the genomic landscape of prostate cancer in Sardinia is unknown. We performed whole exome sequencing on tumor and tumor-adjacent tissues from 30 patients diagnosed with local prostate cancer to reveal

germline risk variants, and identify somatic SNPs, INDELs, copy number variations, and fusion events in prostate cancers. Our data show both genomic disparities and similarities between prostate cancer in our Sardinian cohort and prostate cancers reported in the TCGA European ancestry cohort.

ERG gene family fusions are detected in approximately 50% of patients with European ancestry. However, recent studies show that the incidence of ERG fusion are lower in patients of North America African (20-30%) and Asian (8%-22%) ancestry [305, 309, 310]. We were unable to detect ERG fusion events in Sardinia prostate cancer cohort by whole exome sequencing, although we did detect 2 out of 19 (11%) patients having ERG fusion when using a PCR based method. ERG fusion events are accompanied with ERG loss on Chromosome 21q22.3 [311]. Nevertheless, we only observed ERG loss in only 7% of our Sardinia cohort compared to 15% reported in the TCGA cohort (most of ERG fusions occur in RNA level). Our data indicates a lower frequency of ERG-fusion events in Sardinia cancer patients.

We did find, however, a novel fusion event, BTBD7-SLC2A5, in 12% of the Sardinian patients. In addition, we didfind a novel fusion event, BTBD7-SLC2A5, in 12% of the Sardinian patients, whose breakpoints were within an exon region (Figure 6A). BTB/POZ domain-containing protein 7 (BTBD7) BTBD7 regulates the dynamics of cell adhesion and motility during epithelial branching

morphogenesis[312], and has been reported to be associated with various cancers[313-315]. BTBD7 fusions were also observed in the TCGA pan-cancer cohort. BTBD7-UBR7 and UBR7-BTBD7 were observed in breast cancer, BTBD7-RGS9 fusion in Ovarian Cancer, and TARBP1-BTBD7 fusion in Lung Adenocarcinoma (http://www.cbioprotal.org/). However, each of those fusions were detected only in a single patient. In our study, BTBD7 fusion with the fructose transporter SLC2A5 was observed in four of the prostate cancer patients in Sardinia, indicating BTBD7 fusion can be observed more frequently in tumors in a specific population.

SLC2A5 is a fructose transporter and has been reported to be associated with various cancers as well[316-318]. Lung cancers with mutations in SLC2A5 promote lung adenocarcinoma cell growth and metastasis by enhancing fructose utilization[317]. BTBD7-SLC2A5 fusions that identified in this study contained entire transmembrane domain and fructose transport domain of SLC2A5 may be involved in shifting of energy metabolism to enhance tumor growth. However, further investigated in the future is needed to determine the importance of this fusion.

BTBD7-SLC2A5 may also be involved in unregulated cell adhesion and motility or epithelial morphogenesis due to loss of exons 10 and 11. Functional studies will aid in uncovering the functional consequences of this fusion event.

Meanwhile, we found that 20% of the Sardinia samples had amplification of UGT family while only 0.3% in of the TCGA cohort had amplification in this region. UDP glucuronosyltransferase family genes catalyze the addition of the hydrophilic moiety, glucuronide, to acceptor molecules in a process called glucuronidation[319, 320]. In humans there are two major classes of UDP glucuronosyltransferase, UGT1 and UGT2, each of which contains multiple genes on chromosome 2 and 4, respectively. In Sardinian prostate cancers we observed an amplification in chromosome 4q13.2 which contains multiple UGT2 genes **(Figure4.1).**

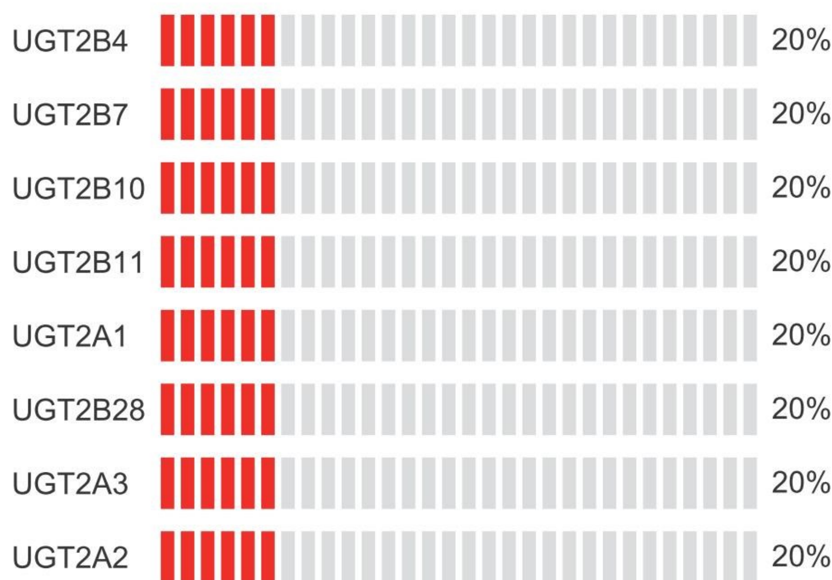

Figure 4.1 Oncoplotof amplifications of genes in Chromosome 4q13.2 across 30 Sardinian prostate cancer patients. Each red bar represents a patient with an amplification of the specified gene on the left.

Within this amplified region the genes UGT2B7, UGT2B15, UGT2B17 and UGT2B28 have been broadly investigated in prostate cancer because of their ability to inactive DHT and testosterone[321]. Germline inactivating mutations including deletion/insertion mutations or single nucleotide polymorphisms of these genes increases prostate cancer risk by presumably increasing levels of unconjugatedactive androgens either systemically (due to reduced hepaticmetabolism), or locally in the prostate, or both[321]. However, a number of studies haveexamined the relationship between UGT2B15 andUGT2B17 expression levels and prostate cancer progression. One study reported that UGT2B17 protein level wasincreased in prostate cancer relative to BPH, and it wasmore abundant in metastatic than benign tumors[322]. Another study reportedthat higher UGT2B17 protein levels were associated with higherGleason scores, metastasis, and progression to CRPC[323].Moreover, UGT2B17 overexpression was associated withincreased risk of biochemical recurrence during androgendeprivation therapy[324]. Recent work has identified a novel function for UGT2B17 in androgen-independent AR signaling related to the activity of c-Src kinase[323]. In contrast, UGT2B15 protein levels were reduced in prostate tumors relative to BPH. To examine the impact of UGT2B15 expression in prostate cancer, we investigated the expression data of the 497 prostate cancer patients from TCGA. Our analysis showed that the expression level of UGT2B15 in tumors had a tendency to be lower than normal tissues (but not significantly lower),

although, UGT2B15 was similar to UGT2B17, patients with higher expression of UGT2B15 had increased risk of biochemical recurrence **(Figure 4.2)**. These analyses suggest that high expression of UGT2B15 and UGT2B17 in the tumors had additional functions other than DHT and testosterone metabolism[323].

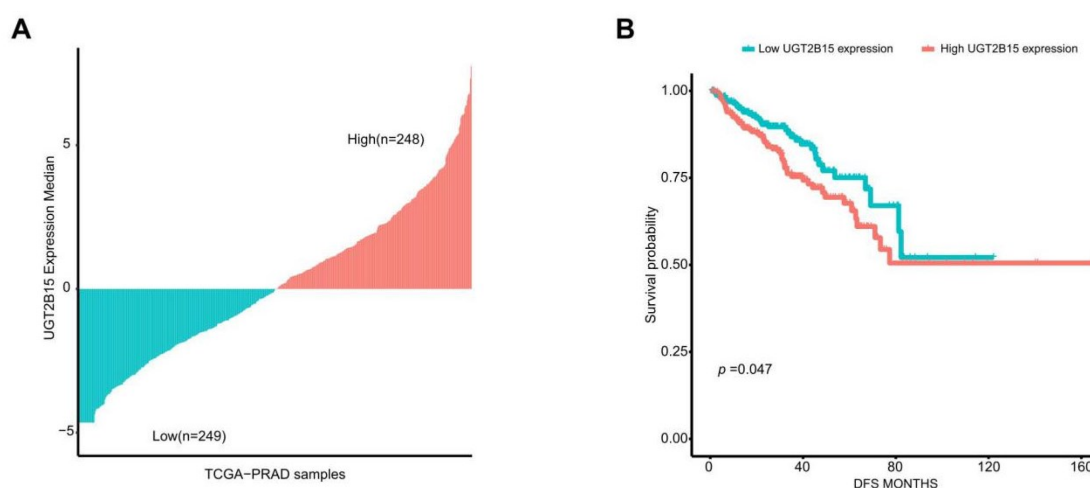

Figure 4.2 Patients with high expression of UGT2B15 had reduced disease-free survival compared with those patients with low expression of UGT2B15. **A)** All the 497 TCGA localized prostate cancer patients were divided into UGT2B15 a low expression group and a UGT2B15 high expression group based on the UGT2B15 median expression. **B)** Between-group comparisons of DFS were performed by the Kaplan-Meier method and the log-rank test.

UGT2B4 was identified as the peak of amplification in the UGT family in our present study. Interestingly, UGT2B4 is noticeable for its ability in the clearance of

estrogens[325]. It was reported that polymorphisms of UGT2B4 have been associated withincreased breast cancer risk[326]. To examine the impact of UGT2B4 expression in prostate cancer, we investigated the expression data of pan-cancer study from TCGA. Our analysis showed that the expression level of UGT2B4 in tumors of breast cancer was significantly lower than normal tissues while UGT2B4 was significantly higher in tumors than normal tissues of prostate cancer **(Figure 4.3)**.

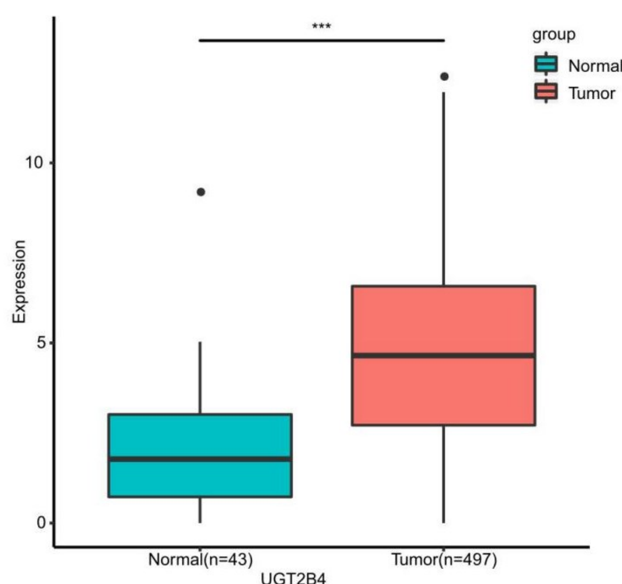

Figure 4.3 Expression of UGT2B4 was determined in 43 normal (green bar) and 497 prostate cancers (red bar) from the TCGA database. UGT2B4 was found to be significantly upregulated (p<0.001) in localized prostate tumors compared with tumor adjacent normal tissue

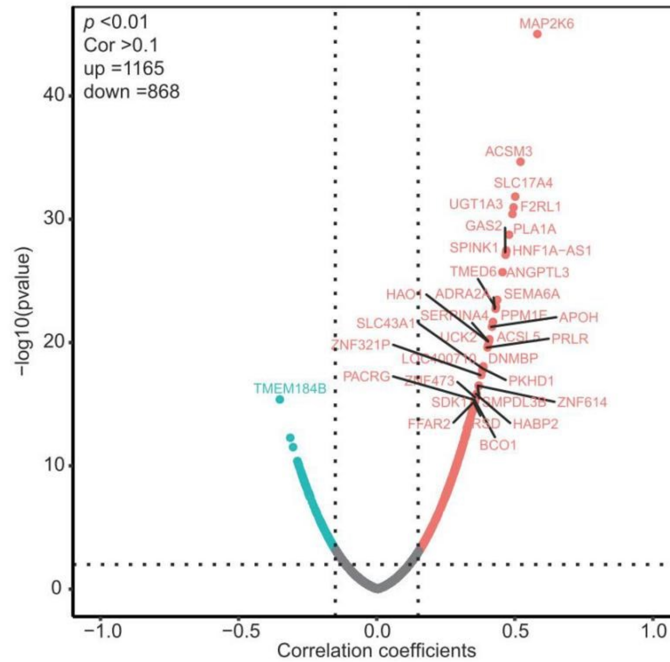

Figure 4.4 Spearman rank correlation analysis was performed on expression profiles of tumor tissues of 497 TCGA localized prostate cancer patients and UGT2B4. 1165 genes were positively co-expressed with UGT2B4 and 868 genes were negatively co-expressed with UGT2B4 i. (Spearman correlation >0.1, p<0.01)

To further examine the potential function of the UGT2B4 in prostate cancer, we performed the spearman correlation analysis of UGT2B4 in the expression data of the 497 prostate cancer patients from TCGA. We found that UGT2B4 expression was

associated with increased expression of genes that were found enriched in glutamine and monocarboxylic acid metabolic pathways. **(Figure 4.4-Figure 4.7)**.

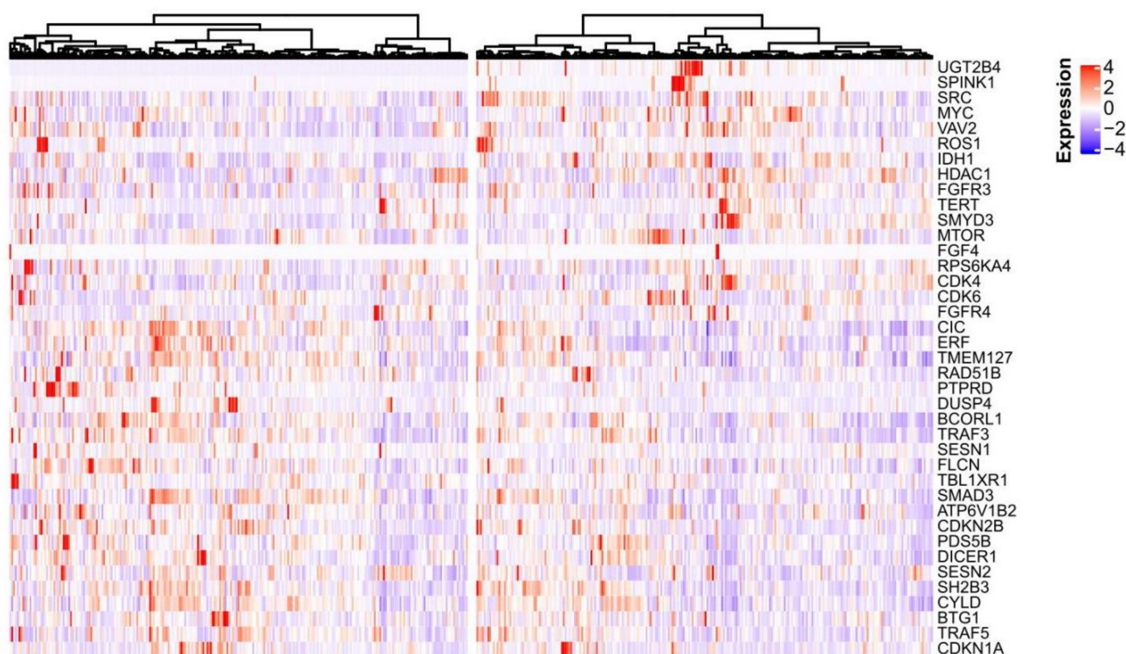

Figure 4.5 Thirty-seven curated oncogenes and tumor suppressor genes were correlated with UGT2B4 in the spearman rank correlation analysis. Heatmap of hierarchal clustering of patients based on expression of UGT2B4 and 37 UGT2B4 co-expressed oncogenes and tumor suppressor genes was described using a function Heatmap in ComplexHeatmap R package. Patients were split by median expression of UGT2B4 before hierarchal clustering. There was a cluster of prostate cancer patients with high expression of UGT2B4, SPINK1, SRC.

Thehigh frequently of somatic UDP-glucuronosyltransferase gene family amplification on Chromosome 4q13.2 in Sardinian prostate cancers may promote tumorigenesis and development by upregulating the expression of UGT2B4 and UGT2B7/15/17. Upregulation of these genes may homeostatically maintain the crosstalk of AR and ER signaling[327, 328] or other metabolic signaling pathways such as SRC and MYC signaling[323].

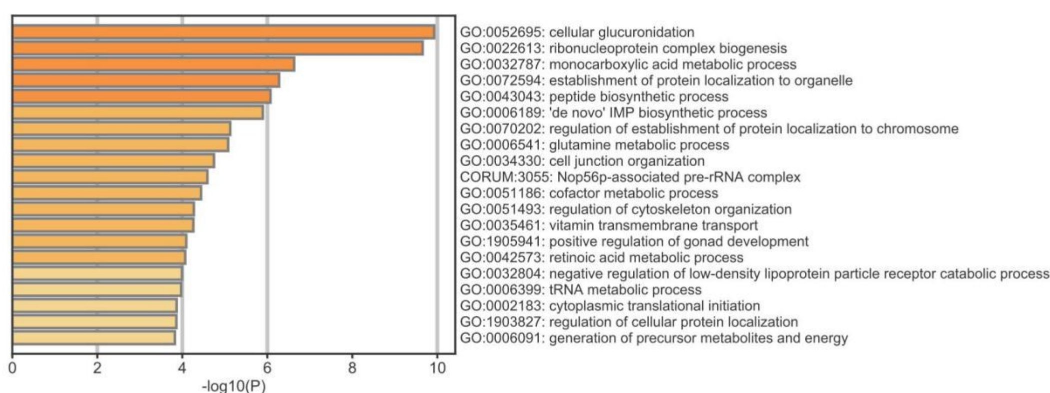

Figure 4.6 Functional enrichment including GO Biological Processes, KEGG and Reactome pathways (http://www.metascape.com/) reveals that genes co-expressed with UGT2B4 in the spearman correlation analysis were functionally enriched in ribonucleoprotein complex biogenesis, as well as glutamine, nucleotide and monocarboxylic acid metabolic pathways

Another possibility is that UDP glucuronosyltransferase family amplification is related to their important role in detoxification of xenobioticsubstrate[329]. Prostate cells with increased expression of UDP glucuronosyltransferase may have provided enhanced fitness against Sardinian environmental xenobiotic toxins.

Figure 4.7 PPI network of UGT2B4 co-expression genes in prostate cancerProtein-Protein interaction network of UGT2B4 co-expressed genes were downloaded from InWeb_IM (https://www.intomics.com/) and visualized by Cytoscape.UGT2B4 positive co-expression genes that broadly interact with SRC and MYC, functionally clustered in ribosome biogenesis, functionally clustered in nucleobase biosynthesis monocarboxylic acid metabolism, and Carbohydrate

metabolic are highlighted in red. Genes that negative co-expressed with UGT2B4 clustered in synapse organization and are blue highlighted.

In addition, we found that IRF8-loss occurred in 6% of Sardinia prostate cancer and in 2-8% of prostate cancer patients across multiple prostate cancer genomic studies [104, 254, 302-305]**(Figure 3.9).** In the most recent genomic study of mCRPC, data show that patients with IRF8-loss have poorer survival outcome compared with patients without IRF8-loss, indicating IRF8 is an important tumor suppression gene in prostate cancer.

IRF8 (also ICSBP) is a transcription factor that is a member of the interferon regulatory protein family (IRF)[330]. IRF8 is predominantly expressed in hematopoietic stem, progenitor and terminally differentiated cells including myeloid, NK and dendritic cells [331]. IRF8 functions as a transcriptional activator and repressor that is required to mediate immune cell differentiation and execution of cell-type-specific gene expression programs [332]. IRF8 also regulates the expression of genes involved in several cellular functions including adaptive immunity, cell cycle regulation and apoptosis [333, 334]. Loss of IRF8 in murine models results in a hematopoietic malignancy, in part due to STAT5 repression of IRF8 tumor suppressive activity [335]. Homozygous biallelic IRF8 mutations have been identified in NK deficiency syndromes and IRF8 missense mutations have been identified in

dendritic cell deficiency syndromes, suggesting IRF8 is required for functional NK and dendritic cell development [336]. Somatic hotspot mutations in the IRF8 DNA binding domain have been identified in pediatric-type follicular lymphoma [337] and in diffuse large B cell lymphoma[338]. In addition, IRF8 downregulation is found in hematopoietic malignancies due to epigenetic and signaling dysregulation [335, 339]. Even though the role of IRF8 in immunity has been widely investigated and the its somatic mutations have been identified in lymphoma, its role in prostate cancer is still unknown.

In keeping with several studies on prostate cancer we identified somatic mutations in TP53, SPOP, KMT2D, FOXA1 and copy number loss of NXK3-1, PTEN, RB1, PCDH9. These mutations are well recognized as important drivers of prostate cancer across multiple ethnic groups [113, 254, 305, 309].

Early genetic quantitative studies of identical and fraternal twins have demonstrated that the heritability of prostate cancer is 42% - 58%, higher than that of any other malignant tumors[52, 340]. Todate, 167 prostate cancer risk loci have been identified by genome-wide association analysis [70]. These loci with high frequency and low penetration in the population explain only about 19% of the familial risk of prostate cancer. In recent years, the effect of variants with low frequency and high penetration on prostate cancersusceptible genes has become clear, such as HOXB13,

BRCA1/BRCA2, DNA mismatch and other DNA repair pathway[72]. These variants explained only about 5% of the family risk of prostate cancer**.**Where is the missing heritability of prostate cancer?

In addition, in United States, the incidence and mortality of prostate cancer varies considerably by races and ethnicities [7, 8]. The prostate cancer incidence of the North AfricanAmerican is 208.7 per 100,000 and the mortality is 47.2 per 100,000, while prostate cancer incidence in the Asian American, Native Hawaiian and Pacific Islander (AANHIP) is only 67.8/100000 which is almost one third of North America Africans, and one half of Non-Hispanic Whites (123/100000)[8]. Even removal of effects of other factors, racial disparities in prostate cancer incidence in United State remain significant.

Therefore, racial geneticdisparity is one of the major factors contributing to the variability in the incidence of prostate cancer. Even though multiple large multi-ethnic genome-wide association studies focused on identification of ethnic specific risk locus for prostate cancer, they are insufficient to explain all of the racial disparity [119, 341]. Where is the missing racial genetic disparity on prostate cancer?

One of the reasons for the missing heritability and racial genetic disparity of prostate cancer is that most studies are based on GWAS or Whole exome sequencing which

underrepresent the full genome and epigenetic factor. The other possible reason is that the prostate cancer risk variants are too rare to be significantly identified. Even though almost 20 susceptible gene have been identified [49], we are still not able to identify the frequency of most of the specific mutations in different populations because they are very rare. Recently there are other bolder speculations that environmental factors can alter the epigenome and can be inherited in future generations [342]. If this is the case, studies based on GWAS and WGS and WES only focus on nucleotides, and do not examine markers of epigenetic heritability [343] However, the current evidence is not sufficient to indicate that the epigenome play a significant role in prostate cancer heritability.

The racial mixing and the underrepresentation from racial and ethnic minorities in current studies are possible reasons as well. In 1795, Johann Friedrich Blumenbach defined the different races of mankind. So far, the scientific community still adopts a similar fuzzy definition, which cannot accurately distinguish races, especially when racial mixing is becoming more common. In addition, most of the prostate cancer genomic studies are based on the Caucasian cohorts. The insufficient number of specimens from African Americanmen and other minority populations, together with the heterogeneity in the molecular etiology of prostate cancer across populations, challenge the generalizability of findings from these projects [133].

In the 30 Sardinian prostate cancer patients, we have identified 106 potential germline risk variants on 82 genes with five of the genes having been previously demonstrated to be associated with familial cancers. Significantly, ARSD-G320D variants were observed in 53% of the Sardinian patients. If there are germline prostate cancer risk variants in the Sardinian population, it would be striking as the highest frequency of germline risk variants were observed in HOXB13 G84E which only occurred in about 1~2% of unselected Nordic patients [344].

Even though we cannot exclude the ARSD G320D variants, as ARSD have been demonstrated to be associated with estrogen metabolism which strongly affects prostate cancer. ARSD transforms the storage form of estrogen (estrogen sulfate) to act estrogen(https://pathcards.genecards.org/card/estrogen_metabolism). Evidence that supports estrogen as a prostate cancer-causing agent includes association of elevated levels of estrogen with prostate cancer, changes in estrogen receptor status in advanced prostate cancer, and rodent models and chimeric human tissue graft models showing induction of prostate cancer using estrogen plus testosterone. Based on our data in this study, it is still difficult to link the ARSD-G320D as a risk variant for prostate cancer in Sardinia. A larger cohort and a more focused study will help resolve these questions.

Table 4.1 Important gene aberrations that enriched in Sardinian prostate cancer

Gene_Symbol	Aberration	Type	Frequency
UGT famlily	Amplification	Somatic	20%
ARSD	G320D	Germline	53%
BTBD7-SLC2A5	Fusion	Somatic	12%
CNN2	G250V&G263S	Germline	9%

Overall, our data revealed the similarities and disparities of the molecular basis of prostate cancer between Sardinians and other ethnic groups (**Table4.1).** The understanding of the genomic landscape of prostate cancer in different ethnic populations will give help in the application of precision medicine in Sardinian patients. The novel drivers we identified in prostate cancer in our Sardinian cohort are potential drug targets. Further understanding their functional mechanism of action will aid in developing drugs to benefit the patients. To achieve these goals, in the future, a retrospective research study with a larger cohort with prognosis information is necessary to investigate the clinical relevance of these gene aberrations in Sardinian prostate cancer. Furthermore experiments in vivo and vitro should be engaged to determine the biology and targetability for these genes.

Conclusion and prospect

We analyzed both germline variation and somatic mutations of 30 Sardinian prostate cancer patients and identified a novel Germline risk mutation ARSD-G320D occurring in 53 percent of the patients. We also found somatic UGT family

amplifications which occurred in 20% the patients. Interestingly, both ARSD and UGT family members regulate estrogen metabolism. A link between estrogen metabolism and prostate cancer is well established [307, 345, 346]. In addition, we identified 15 putative and 9 novel candidate driver mutations, 44 putative copy number driver events and another 31 tumor suppressor gene deletion events on 11 tumor suppressor genes. We pointed out that IRF8 deletion in 16q24.2 is a candidate driver in prostate cancer and patients with IRF8 deletion have worse prognosis. Finally, we identified one candidate out of frame fusion MAGOH-TMEM220 and one candidate in-frame fusion BTBD7-SLC2A5 occur in 15% and 12 % of the patients respectively. Our data revealed similarities and disparities of the molecular basis of prostate cancer between Sardinians and other ethnic groups and will benefit the prostate cancer risk monitoring and treatment management and development in Sardinia.

A brief summary of the Chinese prostate cancer study

To broaden our understanding of ethnicity and prostate cancer we conducted a WES study on a cohort of Chinese prostate cancer patients. In the Chinese prostate cancer cohort, we have seen greater molecular disparities from TCGA cohort than in the Sardinian prostate cancer cohort. In the Chinese study, we have sequenced tumors of 144 Chinese prostate cancer patients and identified 38 genes significantly mutated

genes. Interestingly, 20 of them have not been implicated in prostate cancer in Caucasians studies. These genes include nucleoporin 93, and cyclin D2. **(Figure 5.1)**. Comparing Chinese and Caucasian prostate cancer reveals a set of genomic markers that may be informative for ethnic disparities. **(Figure 5.2 & Figure 5.3)**. These genomic markers are summarized in **Table 5.1.**

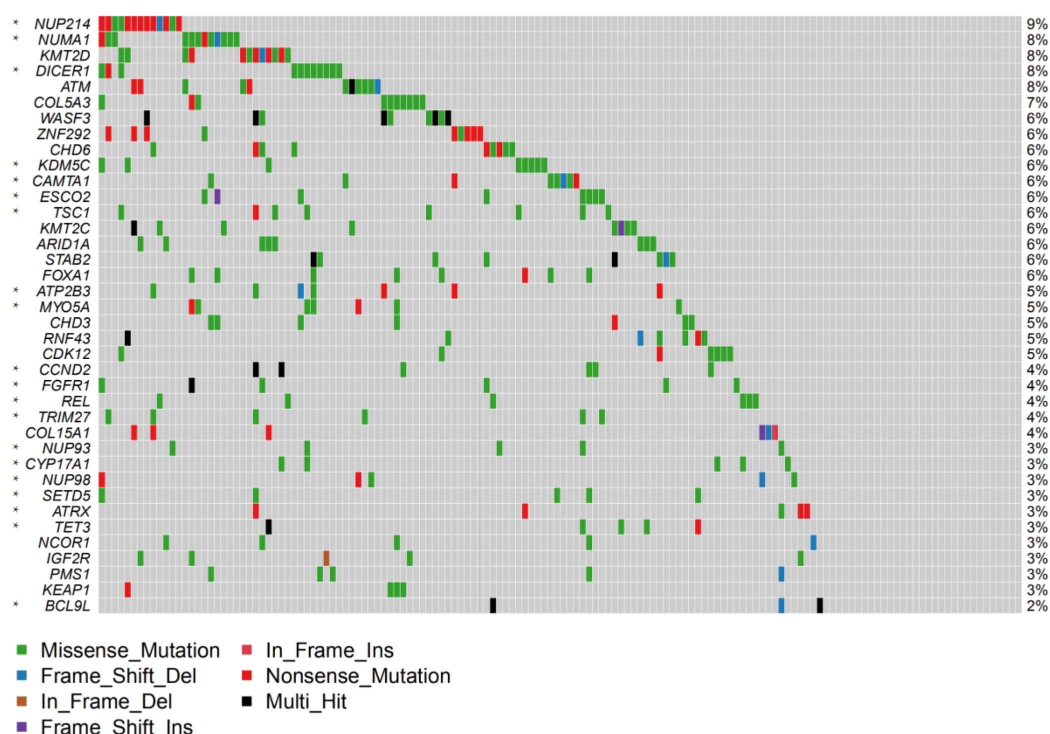

Figure 7.1 Oncoplot of 38 somatic drivers of prostate cancer from 144 Sardinian prostate cancer patients. * Novel identified drivers in prostate cancer of Chinese**.**

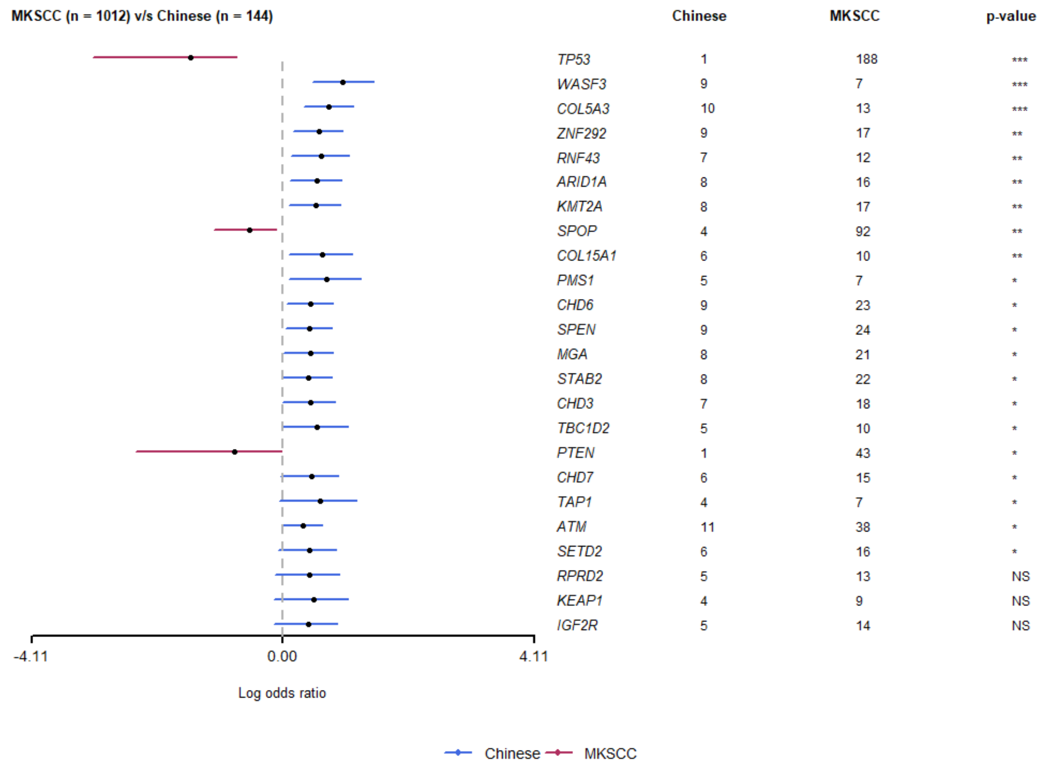

Figure 7.2 Comparison of Chinese and Caucasian prostate cancer with somatic mutations.

Overall，comparison of Chinese and Caucasian prostate cancer genome reveals a great ethnic disparities and once again emphasize the challenge toa compressive understanding of genetic risk and precision medicine of prostate cancer patients because that race and ethnic minority groups underrepresented in the prostate state genomic studies.

Data and material accession numbers

All data are available in NCBI, under study accession number: PRJNA546032.

Financial disclosures: Funding support was provided by Li Ka-Shing Foundation (DJK) and International Institute of Infection and Immunity, Shantou University Medical College (DJK) and Dalhousie Medical Research Foundation (DJK); DJK is the recipient of the Tier I Canada Research Chair in Translational Vaccinology and Inflammation.

Support was also received by the University of Sassari and the RegioneAutonoma di Sardegna to LAS.

Abbreviations

HPV	Human papillomavirus
CMV	Cytomegalovirus
HSV2	Human herpes simplex virus type 2
HHV8	Human herpesvirus type 8
EBV	Epstein-Barr virus
HPC1	Hereditary prostate cancer 1
GWAS	genome-wide association analysis
mCRPC	Metastatic castration resistance prostate cancers
MSI	Microsatellite instability
ROS	Reactive oxygen species

TNF	Tumor necrosis factor
PIA	Proliferative inflammatory atrophy
EMT	Epithelial to mesenchymal transition
PAP	Prostatic acid phosphatase
TRUS	Transrectal ultrasound
HK2	Human kallikrein 2
PHI	Prostate health index
FDA	Food and Drug Administration
AUC	Area under the curve
DRE	Non-digital rectal exam
CRPC	Castration resistant prostate cancer
OS	Overall survival
PFS	Progression-free survival
ADT	Androgen deprivation therapy
GnRH	Gonadotropin-Releasing Hormone
LHRH	Luteinising-hormone releasing hormone
LH	Luteinizing hormone
ACTH	Adreno-cortico-tropic-hormone
AR	Androgen receptor
DBD	DNA binding domain
LBD	Ligand binding domain
NTD	N-terminal domain
PROTAC	Proteolysis Targeting Chimera
PARP	Poly ADP-ribose polymerase
CXCR2	Chemokine-chemokine receptor
PIN	Prostatic intraepithelial neoplasia
HGSOC	High-grade serous ovarian cancer
GOF	Gain-of-function
APC	Adenomatous polyposis coli
ERK	Extracellular signal-regulated kinase
FFPET	Formalin-fixed paraffin-embedded tissue
BWA	Burrows-Wheeler Aligner
ESP	Exome Sequencing Project
GTEx	Genotype-Tissue Expression Project

Acknowledgment

The completion of the thesis is attributed to many peoples support andencouragement. I would like to express my gratitude to all those who have offered me valuable help in my thesis writing.

Firstly, my deepest gratitude goes first and foremost to my supervisors Leonardo Antonio Sechi and David J Kelvin, with constant encouragement and expert guidance, who have offered me valuable suggestion in academic studies and their incisive comments and constructive criticism have contributed greatly to the completion of this thesis. I thank them for the study concept and design, analysis and interpretation of data, supervision and thesis drafting. Secondly, I would like to express my heartfelt gratitude to Francesco Tanda who have instructed and helped a lot in the study concept and design, and acquisition of data. I am also greatly indebted to Maria Antonietta Fedeli for the help and support in acquisition of data. In addition, I pleased to acknowledge Dongsheng Yang, Yuyong Wang (clinical data) for their assistance in data analysis and interpretation. Also, I gratefully acknowledge the help of Giuseppe Palmieri, Xue Bei, and Lisha Jia for their administrative, technical, or material support.

Finally, in particular, my gratitude also extends to Dr Nikki Kelvin who devoted a considerable portion of her time to reading my manuscripts and making suggestion for further revisions. I am also greatly indebted to all my colleagues both in the Division

of Immunity, China and Life of Science Department, Italy who have helped me directly and indirectly in my studies, including technician Guihong Pan for libraries construction, administrative staff Jiachun Liu and Wenling Zeng for administrative support during the project processing, Dr Marco Bo and Dr Sepiedh Hosseini for help with sampling. Indeed, I have benefited a lot and academically prepared for the thesis.

Reference

1. Ferlay, J., et al., *Cancer incidence and mortality worldwide: Sources, methods and major patterns in GLOBOCAN 2012.* International Journal of Cancer, 2015. **136**(5): p. E359-E386.
2. Torre, L.A., et al., *Global Cancer Statistics, 2012.* Ca-a Cancer Journal for Clinicians, 2015. **65**(2): p. 87-108.
3. Pishgar, F., et al., *Global, Regional and National Burden of Prostate Cancer, 1990 to 2015: Results from the Global Burden of Disease Study 2015.* J Urol, 2018. **199**(5): p. 1224-1232.
4. Taitt, H.E., *Global Trends and Prostate Cancer: A Review of Incidence, Detection, and Mortality as Influenced by Race, Ethnicity, and Geographic Location.* Am J Mens Health, 2018. **12**(6): p. 1807-1823.
5. Jemal, A., et al., *Global Patterns of Cancer Incidence and Mortality Rates and Trends.* Cancer Epidemiology Biomarkers & Prevention, 2010. **19**(8): p. 1893-1907.
6. Dy, G.W., et al., *Global Burden of Urologic Cancers, 1990-2013.* Eur Urol, 2017. **71**(3): p. 437-446.
7. DeSantis, C.E., et al., *Cancer statistics for African Americans, 2016: Progress and opportunities in reducing racial disparities.* Ca-a Cancer Journal for Clinicians, 2016. **66**(4): p. 290-308.
8. Siegel, R.L., K.D. Miller, and A. Jemal, *Cancer Statistics, 2016.* Ca-a Cancer Journal for Clinicians, 2016. **66**(1): p. 7-30.
9. Chornokur, G., et al., *Disparities at Presentation, Diagnosis, Treatment, and Survival in African American Men, Affected by Prostate Cancer.* Prostate, 2011. **71**(9): p. 985-997.
10. Schwartz, K., et al., *Interplay of Race, Socioeconomic Status, and Treatment on Survival of Patients With Prostate Cancer.* Urology, 2009. **74**(6): p. 1296-1302.
11. Cheng, I., et al., *Socioeconomic status and prostate cancer incidence and mortality rates among the diverse population of California.* Cancer Causes & Control, 2009. **20**(8): p.

1431-1440.

12. Negoita, S., et al., *Annual Report to the Nation on the Status of Cancer, part II: Recent changes in prostate cancer trends and disease characteristics.* Cancer, 2018. **124**(13): p. 2801-2814.
13. Wang, M.C., Valenzuela, L. A., Murphy, G. P. & and T.M. Chu, *Purification of a human prostate specific antigen.* Invest. Urol, 1979(17): p. 159-163.
14. H. Ballentine Carter, J.D.P., E., *Longitudinal evaluation of prostate specific antigen levels in men with and without prostate disease.* JAMA, 1992.
15. Quinn, M. and P. Babb, *Patterns and trends in prostate cancer incidence, survival, prevalence and mortality. Part I: international comparisons.* Bju International, 2002. **90**(2): p. 162-173.
16. Collin, S.M., et al., *Prostate-cancer mortality in the USA and UK in 1975-2004: an ecological study.* Lancet Oncology, 2008. **9**(5): p. 445-452.
17. Moyer, V.A. and U.S.P.S.T. Force, *Screening for Prostate Cancer: US Preventive Services Task Force Recommendation Statement.* Annals of Internal Medicine, 2012. **157**(2): p. 120-+.
18. Moyer, V.A., *Screening for prostate cancer: US Preventive Services Task Force recommendation statement.* 2012.
19. Jemal, A., et al., *Prostate Cancer Incidence and PSA Testing Patterns in Relation to USPSTF Screening Recommendations.* Jama-Journal of the American Medical Association, 2015. **314**(19): p. 2054-2061.
20. Grossman, D.C., et al., *Screening for Prostate Cancer: US Preventive Services Task Force Recommendation Statement.* Jama, 2018. **319**(18): p. 1901-1913.
21. Fedewa, S.A., et al., *Prevalence of Major Risk Factors and Use of Screening Tests for Cancer in the United States.* Cancer Epidemiology Biomarkers & Prevention, 2015. **24**(4): p. 637-652.
22. Pu, Y.S., et al., *Changing trends of prostate cancer in Asia.* The aging male : the official journal of the International Society for the Study of the Aging Male, 2004. **7**(2): p. 120-32.
23. Sim, H.G. and C.W.S. Cheng, *Changing demography of prostate cancer in Asia.* European Journal of Cancer, 2005. **41**(6): p. 834-845.
24. Chen, R., et al., *Prostate cancer in Asia: A collaborative report.* Asian Journal of Urology, 2014. **1**(1): p. 15-29.
25. Kaaks, R. and P. Stattin, *Obesity, Endogenous Hormone Metabolism, and Prostate Cancer Risk: A Conundrum of "Highs" and "Lows".* Cancer Prevention Research, 2010. **3**(3): p. 259-262.
26. Parekh, N., et al., *Obesity and Prostate Cancer Detection: Insights from Three National Surveys.* American Journal of Medicine, 2010. **123**(9): p. 829-835.
27. Huncharek, M., et al., *Smoking as a Risk Factor for Prostate Cancer: A Meta-Analysis of 24 Prospective Cohort Studies.* American Journal of Public Health, 2010. **100**(4): p. 693-701.
28. Moukayed, M. and W.B. Grant, *The roles of UVB and vitamin D in reducing risk of cancer incidence and mortality: A review of the epidemiology, clinical trials, and mechanisms.* Rev Endocr Metab Disord, 2017. **18**(2): p. 167-182.
29. Patel, D.I., et al., *Nexrutine and exercise similarly prevent high grade prostate tumors in transgenic mouse model.* PLoS One, 2019. **14**(12): p. e0226187.
30. Lopez-Guarnido, O., et al., *Bioactive compounds of the Mediterranean diet and prostate cancer.* Aging Male, 2018. **21**(4): p. 251-260.

31. Howlader N, N.A., Krapcho M，Bethesda, *SEER cancer statistics review, 1975-2010.* Bethesda, MD based on November 2012

SEER data submission, posted to the SEER Web site,, 2012.

32. Vickers AJ, C.A., Bjork T, *Theconcentration of prostate specific antigen at age 60predicts lifetime risk of metastasis and death from PCa.* BMJ, 2010(341): p. 1-8.

33. Klein, E.A. and R. Silverman, *Inflammation, infection, and prostate cancer.* Current Opinion in Urology, 2008. **18**(3): p. 315-319.

34. Groom, H.C.T., et al., *No Evidence for Infection of UK Prostate Cancer Patients with XMRV, BK Virus, Trichomonas vaginalis or Human Papilloma Viruses.* Plos One, 2012. **7**(3).

35. Giorgio I. Russo, A.E.C., Rosita A. Condorelli, Guido Scalia, Giuseppe Morgia & Sandro La Vignera, *Human papillomavirus and risk of prostate cancer: a systematic review and meta-analysis.* The aging male : the official journal of the International Society for the Study of the Aging Male, 2018.

36. Dong, Q., et al., *The Microbial Communities in Male First Catch Urine Are Highly Similar to Those in Paired Urethral Swab Specimens.* Plos One, 2011. **6**(5).

37. Whiteside, S.A., et al., *The microbiome of the urinary tract-a role beyond infection.* Nature Reviews Urology, 2015. **12**(2): p. 81-90.

38. Porter, C.M., et al., *The microbiome in prostate inflammation and prostate cancer.* Prostate Cancer Prostatic Dis, 2018. **21**(3): p. 345-354.

39. Nickel, J.C. and J. Xiang, *Clinical significance of nontraditional bacterial uropathogens in the management of chronic prostatitis.* Journal of Urology, 2008. **179**(4): p. 1391-1395.

40. Cohen, R.J., et al., *Propionibacterium acnes associated with inflammation in radicial prostatectomy specimens: A possible link to cancer evolution?* Journal of Urology, 2005. **173**(6): p. 1969-1974.

41. Shinohara, D.B., et al., *A mouse model of chronic prostatic inflammation using a human prostate cancer-derived isolate of Propionibacterium acnes.* Prostate, 2013. **73**(9): p. 1007-1015.

42. Yu, H., et al., *Urinary microbiota in patients with prostate cancer and benign prostatic hyperplasia.* Archives of Medical Science, 2015. **11**(2): p. 385-394.

43. Davidsson, S., et al., *Frequency and typing of Propionibacterium acnes in prostate tissue obtained from men with and without prostate cancer.* Infectious Agents and Cancer, 2016. **11**.

44. Caini, S., et al., *Sexually transmitted infections and prostate cancer risk: A systematic review and meta-analysis.* Cancer Epidemiology, 2014. **38**(4): p. 329-338.

45. Anttila, T., et al., *Chlamydial antibodies and risk of prostate cancer.* Cancer Epidemiology Biomarkers & Prevention, 2005. **14**(2): p. 385-389.

46. Sutcliffe, S., et al., *Gonorrhea, syphilis, clinical prostatitis, and the risk of prostate cancer.* Cancer Epidemiology Biomarkers & Prevention, 2006. **15**(11): p. 2160-2166.

47. Sutcliffe, S., et al., *Plasma antibodies against Trichomonas vaginalis and subsequent risk of prostate cancer.* Cancer Epidemiology Biomarkers & Prevention, 2006. **15**(5): p. 939-945.

48. Albright, F.S., et al., *Relative Risks for Lethal Prostate Cancer Based on Complete Family History of Prostate Cancer Death.* Prostate, 2017. **77**(1): p. 41-48.

49. Giri, V.N., et al., *Role of Genetic Testing for Inherited Prostate Cancer Risk: Philadelphia*

Prostate Cancer Consensus Conference 2017. Journal of Clinical Oncology, 2018. **36**(4): p. 414-+.

50. Giri, V.N., et al., *Role of Genetic Testing for Inherited Prostate Cancer Risk: Philadelphia Prostate Cancer Consensus Conference 2017.* J Clin Oncol, 2018. **36**(4): p. 414-424.
51. Ahlbom, A., et al., *Cancer in twins: Genetic and nongenetic familial risk factors.* Journal of the National Cancer Institute, 1997. **89**(4): p. 287-293.
52. Page, W.F., et al., *Heredity and prostate cancer: A study of World War II veteran twins.* Prostate, 1997. **33**(4): p. 240-245.
53. Lichtenstein, P., et al., *Environmental and heritable factors in the causation of cancer - Analyses of cohorts of twins from Sweden, Denmark, and Finland.* New England Journal of Medicine, 2000. **343**(2): p. 78-85.
54. Smith, J.R., et al., *Major susceptibility locus for prostate cancer on chromosome 1 suggested by a genome-wide search.* Science, 1996. **274**(5291): p. 1371-1374.
55. Gronberg, H., et al., *Segregation analysis of prostate cancer in Sweden: Support for dominant inheritance.* American Journal of Epidemiology, 1997. **146**(7): p. 552-557.
56. Eeles, R.A., et al., *Linkage analysis of chromosome 1q markers in 136 prostate cancer families.* American Journal of Human Genetics, 1998. **62**(3): p. 653-658.
57. Wiklund, F., et al., *Linkage analysis of prostate cancer susceptibility: confirmation of linkage at 8p22-23.* Human Genetics, 2003. **112**(4): p. 414-418.
58. Freedman, M.L., et al., *Admixture mapping identifies 8q24 as a prostate cancer risk locus in African-American men.* Proceedings of the National Academy of Sciences of the United States of America, 2006. **103**(38): p. 14068-14073.
59. Simard, J., et al., *Prostate cancer susceptibility genes: lessons learned and challenges posed.* Endocrine-Related Cancer, 2003. **10**(2): p. 225-259.
60. Schaid, D.J., *The complex genetic epidemiology of prostate cancer.* Human Molecular Genetics, 2004. **13**: p. R103-R121.
61. Yeager, M., et al., *Genome-wide association study of prostate cancer identifies a second risk locus at 8q24.* Nature Genetics, 2007. **39**(5): p. 645-649.
62. Ewing, C.M., et al., *Germline Mutations in HOXB13 and Prostate-Cancer Risk.* New England Journal of Medicine, 2012. **366**(2): p. 141-149.
63. Auer, P.L. and G. Lettre, *Rare variant association studies: considerations, challenges and opportunities.* Genome Medicine, 2015. **7**.
64. Benafif, S. and R. Eeles, *Genetic predisposition to prostate cancer.* British Medical Bulletin, 2016. **120**(1): p. 75-89.
65. Gudmundsson, J., et al., *Genome-wide association study identifies a second prostate cancer susceptibility variant at 8q24.* Nature Genetics, 2007. **39**(5): p. 631-637.
66. Eeles, R.A., et al., *Multiple newly identified loci associated with prostate cancer susceptibility.* Nature Genetics, 2008. **40**(3): p. 316-321.
67. Thomas, G., et al., *Multiple loci identified in a genome-wide association study of prostate cancer.* Nature Genetics, 2008. **40**(3): p. 310-315.
68. Batra, J., et al., *A Replication Study Examining Novel Common Single Nucleotide Polymorphisms Identified Through a Prostate Cancer Genome-wide Association Study in a*

Japanese Population. American Journal of Epidemiology, 2011. **174**(12): p. 1391-1395.
69. Al Olama, A.A., et al., *A meta-analysis of 87,040 individuals identifies 23 new susceptibility loci for prostate cancer.* Nature Genetics, 2014. **46**(10): p. 1103-1109.
70. Schumacher, F.R., et al., *Association analyses of more than 140,000 men identify 63 new prostate cancer susceptibility loci.* Nat Genet, 2018. **50**(7): p. 928-936.
71. Sud, A., B. Kinnersley, and R.S. Houlston, *Genome-wide association studies of cancer: current insights and future perspectives.* Nature Reviews Cancer, 2017. **17**(11): p. 692-704.
72. Pilie, P.G., V.N. Giri, and K.A. Cooney, *HOXB13 and other high penetrant genes for prostate cancer.* Asian Journal of Andrology, 2016. **18**(4): p. 530-532.
73. Carter, B.S., et al., *MENDELIAN INHERITANCE OF FAMILIAL PROSTATE-CANCER.* Proceedings of the National Academy of Sciences of the United States of America, 1992. **89**(8): p. 3367-3371.
74. Nagy, R., K. Sweet, and C. Eng, *Highly penetrant hereditary cancer syndromes.* Oncogene, 2004. **23**(38): p. 6445-6470.
75. Xu, J., et al., *HOXB13 is a susceptibility gene for prostate cancer: results from the International Consortium for Prostate Cancer Genetics (ICPCG).* Human Genetics, 2013. **132**(1): p. 5-14.
76. Laitinen, V.H., et al., *HOXB13 G84E Mutation in Finland: Population-Based Analysis of Prostate, Breast, and Colorectal Cancer Risk.* Cancer Epidemiology Biomarkers & Prevention, 2013. **22**(3): p. 452-460.
77. Hoffmann, T.J., et al., *Imputation of the Rare HOXB13 G84E Mutation and Cancer Risk in a Large Population-Based Cohort.* Plos Genetics, 2015. **11**(1).
78. Lin, X., et al., *A novel Germline mutation in HOXB13 is associated with prostate cancer risk in Chinese men.* Prostate, 2013. **73**(2): p. 169-175.
79. Maia, S., et al., *Identification of Two Novel HOXB13 Germline Mutations in Portuguese Prostate Cancer Patients.* Plos One, 2015. **10**(7).
80. Dhanasekaran, S.M., et al., *Delineation of prognostic biomarkers in prostate cancer.* Nature, 2001. **412**(6849): p. 822-826.
81. Norris, J.D., et al., *The Homeodomain Protein HOXB13 Regulates the Cellular Response to Androgens.* Molecular Cell, 2009. **36**(3): p. 405-416.
82. Pomerantz, M.M., et al., *The androgen receptor cistrome is extensively reprogrammed in human prostate tumorigenesis.* Nature Genetics, 2015. **47**(11): p. 1346-+.
83. Struewing, J.P., et al., *The risk of cancer associated with specific mutations of BRCA1 and BRCA2 among Ashkenazi Jews.* New England Journal of Medicine, 1997. **336**(20): p. 1401-1408.
84. Gayther, S.A., et al., *The frequency of germ-line mutations in the breast cancer predisposition genes BRCA1 and BRCA2 in familial prostate cancer.* Cancer Research, 2000. **60**(16): p. 4513-4518.
85. Roy, R., J. Chun, and S.N. Powell, *BRCA1 and BRCA2: different roles in a common pathway of genome protection.* Nature Reviews Cancer, 2012. **12**(1): p. 68-78.
86. Deng, C.X., *BRCA1: cell cycle checkpoint, genetic instability, DNA damage response and cancer evolution.* Nucleic Acids Research, 2006. **34**(5): p. 1416-1426.
87. Moynahan, M.E., A.J. Pierce, and M. Jasin, *BRCA2 is required for homology-directed repair of chromosomal breaks.* Molecular Cell, 2001. **7**(2): p. 263-272.

88. Li, D., et al., *The role of BRCA1 and BRCA2 in prostate cancer.* Frontiers in Bioscience-Landmark, 2013. **18**: p. 1445-1459.
89. Lynch HT, L.S., Karr B, Franklin B, Lynch JF, Watson P, Tinley S, Lerman C, Carter C., *Etiology, natural history, management and molecular genetics of hereditary nonpolyposis colorectal cancer (Lynch syndromes): genetic counseling implications.* Cancer Epidemiol Biomarkers Prev, 1997. **6(12):987-91**.
90. Grindedal, E.M., et al., *Germ-Line Mutations in Mismatch Repair Genes Associated with Prostate Cancer.* Cancer Epidemiology Biomarkers & Prevention, 2009. **18**(9): p. 2460-2467.
91. Raymond, V.M., et al., *Elevated Risk of Prostate Cancer Among Men With Lynch Syndrome.* Journal of Clinical Oncology, 2013. **31**(14): p. 1713-+.
92. Dominguez-Valentin, M., et al., *Frequent mismatch-repair defects link prostate cancer to Lynch syndrome.* Bmc Urology, 2016. **16**.
93. Thibodeau, S.N., et al., *Altered expression of hMSH2 and hMLH1 in tumors with microsatellite instability and genetic alterations in mismatch repair genes.* Cancer Research, 1996. **56**(21): p. 4836-4840.
94. Pritchard, C.C., et al., *Complex MSH2 and MSH6 mutations in hypermutated microsatellite unstable advanced prostate cancer.* Nature Communications, 2014. **5**.
95. Pritchard, C.C., et al., *Inherited DNA-Repair Gene Mutations in Men with Metastatic Prostate Cancer.* New England Journal of Medicine, 2016. **375**(5): p. 443-453.
96. He, W.W., et al., *A novel human prostate-specific, androgen-regulated homeobox gene (NKX3.1) that maps to 8p21, a region frequently deleted in prostate cancer.* Genomics, 1997. **43**(1): p. 69-77.
97. Bhatia-Gaur, R., et al., *Roles for Nkx3.1 in prostate development and cancer.* Genes & Development, 1999. **13**(8): p. 966-977.
98. Dong, J.T., *Chromosomal deletions and tumor suppressor genes in prostate cancer.* Cancer and Metastasis Reviews, 2001. **20**(3-4): p. 173-193.
99. Jenkins, R.B., et al., *Detection of c-myc oncogene amplification and chromosomal anomalies in metastatic prostatic carcinoma by fluorescence in situ hybridization.* Cancer Research, 1997. **57**(3): p. 524-531.
100. Nupponen, N.N. and T. Visakorpi, *Molecular cytogenetics of prostate cancer.* Microscopy Research and Technique, 2000. **51**(5): p. 456-463.
101. Visakorpi, T., et al., *IN-VIVO AMPLIFICATION OF THE ANDROGEN RECEPTOR GENE AND PROGRESSION OF HUMAN PROSTATE-CANCER.* Nature Genetics, 1995. **9**(4): p. 401-406.
102. Petrovics, G., et al., *Frequent overexpression of ETS-related gene-1 (ERG1) in prostate cancer transcriptome.* Oncogene, 2005. **24**(23): p. 3847-3852.
103. Tomlins, S.A., et al., *Recurrent fusion of TMPRSS2 and ETS transcription factor genes in prostate cancer.* Science, 2005. **310**(5748): p. 644-648.
104. Taylor, B.S., et al., *Integrative genomic profiling of human prostate cancer.* Cancer Cell, 2010. **18**(1): p. 11-22.
105. Berger, M.F., et al., *The genomic complexity of primary human prostate cancer.* Nature, 2011. **470**(7333): p. 214-220.
106. Barbieri, C.E., et al., *Exome sequencing identifies recurrent SPOP, FOXA1 and MED12*

mutations in prostate cancer. Nature Genetics, 2012. **44**(6): p. 685-U107.

107. Liu, W., et al., *Identification of novel CHD1-associated collaborative alterations of genomic structure and functional assessment of CHD1 in prostate cancer.* Oncogene, 2012. **31**(35): p. 3939-3948.
108. Ren, G., et al., *Identification of frequent BRAF copy number gain and alterations of RAF genes in chinese prostate cancer.* Genes Chromosomes & Cancer, 2012. **51**(11): p. 1014-1023.
109. Weischenfeldt, J., et al., *Integrative Genomic Analyses Reveal an Androgen-Driven Somatic Alteration Landscape in Early-Onset Prostate Cancer.* Cancer Cell, 2013. **23**(2): p. 159-170.
110. Gao, J., et al., *Integrative analysis of complex cancer genomics and clinical profiles using the cBioPortal.* Sci Signal, 2013. **6**(269): p. pl1.
111. Robinson, D., et al., *Integrative Clinical Genomics of Advanced Prostate Cancer (vol 161, pg 1215, 2015).* Cell, 2015. **162**(2): p. 454-454.
112. Cancer Genome Atlas Research, N., *The Molecular Taxonomy of Primary Prostate Cancer.* Cell, 2015. **163**(4): p. 1011-25.
113. Armenia, J., et al., *The long tail of oncogenic drivers in prostate cancer.* Nat Genet, 2018. **50**(5): p. 645-651.
114. Fraser, M., et al., *Genomic hallmarks of localized, non-indolent prostate cancer.* Nature, 2017. **541**(7637): p. 359-+.
115. Robinson, D., et al., *Integrative Clinical Genomics of Advanced Prostate Cancer.* Cell, 2015. **161**(5): p. 1215-1228.
116. Yeager, M., et al., *Identification of a new prostate cancer susceptibility locus on chromosome 8q24.* Nature Genetics, 2009. **41**(10): p. 1055-1057.
117. Haiman, C.A., et al., *Genome-wide association study of prostate cancer in men of African ancestry identifies a susceptibility locus at 17q21.* Nature Genetics, 2011. **43**(6): p. 570-U103.
118. Taioli, E., et al., *Polymorphisms in CYP17 and CYP3A4 and prostate cancer in men of African descent.* Prostate, 2013. **73**(6): p. 668-676.
119. Cheng, I., et al., *Evaluating Genetic Risk for Prostate Cancer among Japanese and Latinos.* Cancer Epidemiology Biomarkers & Prevention, 2012. **21**(11): p. 2048-2058.
120. Tan, Y.-C., et al., *Common 8q24 sequence variations are associated with Asian Indian advanced prostate cancer risk.* Cancer Epidemiology Biomarkers & Prevention, 2008. **17**(9): p. 2431-2435.
121. Takata, R., et al., *Genome-wide association study identifies five new susceptibility loci for prostate cancer in the Japanese population.* Nature Genetics, 2010. **42**(9): p. 751-U34.
122. Xu, J., et al., *Genome-wide association study in Chinese men identifies two new prostate cancer risk loci at 9q31.2 and 19q13.4.* Nature Genetics, 2012. **44**(11): p. 1231-1235.
123. Marzec, J., et al., *A genetic study and meta-analysis of the genetic predisposition of prostate cancer in a Chinese population.* Oncotarget, 2016. **7**(16): p. 21393-21403.
124. Breyer, J.P., et al., *Confirmation of the HOXB13 G84E Germline Mutation in Familial Prostate Cancer.* Cancer Epidemiology Biomarkers & Prevention, 2012. **21**(8): p. 1348-1353.
125. Koboldt, D.C., et al., *Rare Variation in TET2 Is Associated with Clinically Relevant Prostate Carcinoma in African Americans.* Cancer Epidemiology Biomarkers & Prevention, 2016. **25**(11): p. 1456-1463.

126. Magi-Galluzzi, C., et al., *TMPRSS2-ERG Gene Fusion Prevalence and Class Are Significantly Different in Prostate Cancer of Caucasian, African-American and Japanese Patients.* Prostate, 2011. **71**(5): p. 489-497.
127. Zhou, C.K., et al., *TMPRSS2:ERG Gene Fusions in Prostate Cancer of West African Men and a Meta-Analysis of Racial Differences.* Am J Epidemiol, 2017. **186**(12): p. 1352-1361.
128. Khani, F., et al., *Evidence for molecular differences in prostate cancer between African American and Caucasian men.* Clin Cancer Res, 2014. **20**(18): p. 4925-34.
129. Petrovics, G., et al., *A novel genomic alteration of LSAMP associates with aggressive prostate cancer in African American men.* Ebiomedicine, 2015. **2**(12): p. 1957-1964.
130. Bose, R., et al., *ERF mutations reveal a balance of ETS factors controlling prostate oncogenesis.* Nature, 2017. **546**(7660): p. 671-+.
131. Huang, F.W., et al., *Exome Sequencing of African-American Prostate Cancer Reveals Loss-of-Function ERF Mutations.* Cancer Discovery, 2017. **7**(9): p. 973-983.
132. Lindquist, K.J., et al., *Mutational Landscape of Aggressive Prostate Tumors in African American Men.* Cancer Res, 2016. **76**(7): p. 1860-8.
133. Tan, S.H., G. Petrovics, and S. Srivastava, *Prostate Cancer Genomics: Recent Advances and the Prevailing Underrepresentation from Racial and Ethnic Minorities.* Int J Mol Sci, 2018. **19**(4).
134. Isaacs, J.T., *PROSTATIC STRUCTURE AND FUNCTION IN RELATION TO THE ETIOLOGY OF PROSTATIC-CANCER.* Prostate, 1983. **4**(4): p. 351-366.
135. Cunha, G.R., et al., *THE ENDOCRINOLOGY AND DEVELOPMENTAL BIOLOGY OF THE PROSTATE.* Endocrine Reviews, 1987. **8**(3): p. 338-362.
136. McNeal, J.E., *NORMAL HISTOLOGY OF THE PROSTATE.* American Journal of Surgical Pathology, 1988. **12**(8): p. 619-633.
137. McNeal, J.E., *THE ZONAL ANATOMY OF THE PROSTATE.* Prostate, 1981. **2**(1): p. 35-49.
138. Karthaus, W.R., et al., *Identification of Multipotent Luminal Progenitor Cells in Human Prostate Organoid Cultures.* Cell, 2014. **159**(1): p. 163-175.
139. McNeal, J.E., et al., *ZONAL DISTRIBUTION OF PROSTATIC ADENOCARCINOMA - CORRELATION WITH HISTOLOGIC PATTERN AND DIRECTION OF SPREAD.* American Journal of Surgical Pathology, 1988. **12**(12): p. 897-906.
140. Drachenberg, C.B. and J.C. Papadimitriou, *Prostatic corpora amylacea and crystalloids: Similarities and difference on ultrastructural and histochemical studies.* Journal of Submicroscopic Cytology and Pathology, 1996. **28**(2): p. 141-150.
141. De Marzo, A.M., et al., *Proliferative inflammatory atrophy of the prostate - Implications for prostatic carcinogenesis.* American Journal of Pathology, 1999. **155**(6): p. 1985-1992.
142. Montironi, R., R. Mazzucchelli, and M. Scarpelli, *Precancerous lesions and conditions of the prostate - From morphological and biological characterization to chemoprevention*, in *Hormone-Related Tumors: Novel Approaches to Prevention and Treatment*, L. Castanetta, et al., Editors. 2002. p. 169-184.
143. Strand, D.W. and A.S. Goldstein, *The many ways to make a luminal cell and a prostate cancer cell.* Endocrine-Related Cancer, 2015. **22**(6): p. T187-T197.
144. Wang, G., et al., *Genetics and biology of prostate cancer.* Genes Dev, 2018. **32**(17-18): p. 1105-1140.

145. Franks, L.M., *ATROPHY AND HYPERPLASIA IN THE PROSTATE PROPER.* Journal of Pathology and Bacteriology, 1954. **68**(2): p. 617-&.
146. Ruska, K.M., J. Sauvageot, and J.I. Epstein, *Histology and cellular kinetics of prostatic atrophy.* American Journal of Surgical Pathology, 1998. **22**(9): p. 1073-1077.
147. van Leenders, G., et al., *Intermediate cells in human prostate epithelium are enriched in proliferative inflammatory atrophy.* American Journal of Pathology, 2003. **162**(5): p. 1529-1537.
148. De Marzo, A.M., et al., *Inflammation in prostate carcinogenesis.* Nature Reviews Cancer, 2007. **7**(4): p. 256-269.
149. Ouyang, X., et al., *Loss-of-function of Nkx3.1 promotes increased oxidative damage in prostate carcinogenesis.* Cancer Research, 2005. **65**(15): p. 6773-6779.
150. Elkahwaji, J.E., et al., *Chronic bacterial infection and inflammation incite reactive hyperplasia in a mouse model of chronic prostatitis.* Prostate, 2007. **67**(1): p. 14-21.
151. Sfanos, K.S., et al., *Acute inflammatory proteins constitute the organic matrix of prostatic corpora amylacea and calculi in men with prostate cancer.* Proceedings of the National Academy of Sciences of the United States of America, 2009. **106**(9): p. 3443-3448.
152. Mani, R.S., et al., *Inflammation-Induced Oxidative Stress Mediates Gene Fusion Formation in Prostate Cancer.* Cell Reports, 2016. **17**(10): p. 2620-2631.
153. Kasper, S., *Survey of genetically engineered mouse models for prostate cancer: Analyzing the molecular basis of prostate cancer development, progression, and metastasis.* Journal of Cellular Biochemistry, 2005. **94**(2): p. 279-297.
154. Norrish, A.E., et al., *Heterocyclic amine content of cooked meat and risk of prostate cancer.* Jnci-Journal of the National Cancer Institute, 1999. **91**(23): p. 2038-2044.
155. Sugimura, T., et al., *Heterocyclic amines: Mutagens/carcinogens produced during cooking of meat and fish.* Cancer Science, 2004. **95**(4): p. 290-299.
156. Wu, J.B., et al., *Monoamine oxidase A mediates prostate tumorigenesis and cancer metastasis.* Journal of Clinical Investigation, 2014. **124**(7): p. 2891-2908.
157. Brabletz, T., et al., *EMT in cancer.* Nature Reviews Cancer, 2018. **18**(2): p. 128-+.
158. Mittal, V., *Epithelial Mesenchymal Transition in Tumor Metastasis*, in *Annual Review of Pathology: Mechanisms of Disease, Vol 13*, A.K. Abbas and J.C. Aster, Editors. 2018. p. 395-412.
159. Yap, T.A., et al., *Circulating Tumor Cells: A Multifunctional Biomarker.* Clinical Cancer Research, 2014. **20**(10): p. 2553-2568.
160. Aceto, N., et al., *En Route to Metastasis: Circulating Tumor Cell Clusters and Epithelial-to-Mesenchymal Transition.* Trends in Cancer, 2015. **1**(1): p. 44-52.
161. Hoshino, A., et al., *Tumour exosome integrins determine organotropic metastasis.* Nature, 2015. **527**(7578): p. 329-+.
162. Catalona, W.J., et al., *COMPARISON OF DIGITAL RECTAL EXAMINATION AND SERUM PROSTATE-SPECIFIC ANTIGEN IN THE EARLY DETECTION OF PROSTATE-CANCER - RESULTS OF A MULTICENTER CLINICAL-TRIAL OF 6,630 MEN.* Journal of Urology, 1994. **151**(5): p. 1283-1290.
163. Cooner, W.H., et al., *PROSTATE-CANCER DETECTION IN A CLINICAL UROLOGICAL PRACTICE BY ULTRASONOGRAPHY, DIGITAL RECTAL EXAMINATION AND PROSTATE SPECIFIC ANTIGEN.*

Journal of Urology, 1990. **143**(6): p. 1146-1154.

164. Catalona, W.J., et al., *Measurement of prostate-specific antigen in serum as a screening test for prostate cancer.* N Engl J Med, 1991. **324**(17): p. 1156-61.

165. Carter, H.B., et al., *LONGITUDINAL EVALUATION OF PROSTATE-SPECIFIC ANTIGEN LEVELS IN MEN WITH AND WITHOUT PROSTATE DISEASE.* Jama-Journal of the American Medical Association, 1992. **267**(16): p. 2215-2220.

166. Catalona, W.J., et al., *MEASUREMENT OF PROSTATE-SPECIFIC ANTIGEN IN SERUM AS A SCREENING-TEST FOR PROSTATE-CANCER.* New England Journal of Medicine, 1991. **324**(17): p. 1156-1161.

167. Laxman, B., et al., *A first-generation multiplex biomarker analysis of urine for the early detection of prostate cancer.* Cancer Research, 2008. **68**(3): p. 645-649.

168. Hodge, K.K., J.E. McNeal, and T.A. Stamey, *ULTRASOUND GUIDED TRANS-RECTAL CORE BIOPSIES OF THE PALPABLY ABNORMAL PROSTATE.* Journal of Urology, 1989. **142**(1): p. 66-70.

169. Levine, M.A., et al., *Two consecutive sets of transrectal ultrasound guided sextant biopsies of the prostate for the detection of prostate cancer.* Journal of Urology, 1998. **159**(2): p. 471-475.

170. Presti, J.C., et al., *Extended peripheral zone biopsy schemes increase cancer detection rates and minimize variance in prostate specific antigen and age related cancer rates: Results of a community multi-practice study.* Journal of Urology, 2003. **169**(1): p. 125-129.

171. Boorjian, S.A., et al., *Mayo Clinic validation of the D'Amico risk group classification for predicting survival following radical prostatectomy.* Journal of Urology, 2008. **179**(4): p. 1354-1360.

172. Lee, S.E., et al., *Application of the Epstein criteria for prediction of clinically insignificant prostate cancer in Korean men.* Bju International, 2010. **105**(11): p. 1526-1530.

173. Partin, M.A.K.A.W., *Evolving Role of Pro-PSA as a New Serum Marker for the Early Detection of Prostate Cancer.* Rev Urol, 2002(4(4)): p. 198–200

174. Guazzoni, G., et al., *Prostate-specific antigen (PSA) isoform p2PSA significantly improves the prediction of prostate cancer at initial extended prostate biopsies in patients with total PSA between 2.0 and 10 ng/ml: results of a prospective study in a clinical setting.* Eur Urol, 2011. **60**.

175. Fossati, N., et al., *Preoperative prostate-specific antigen isoform p2PSA and its derivatives, %p2PSA and prostate health index, predict pathologic outcomes in patients undergoing radical prostatectomy for prostate cancer: Results from a Multicentric European Prospective Study.* Eur Urol, 2015. **68**.

176. Konety, B., et al., *The 4Kscore(R) Test Reduces Prostate Biopsy Rates in Community and Academic Urology Practices.* Rev Urol, 2015. **17**(4): p. 231-40.

177. Parekh, D.J., et al., *A multi-institutional prospective trial in the USA confirms that the 4Kscore accurately identifies men with high-grade prostate cancer.* Eur Urol, 2015. **68**(3): p. 464-70.

178. Borque-Fernando, A., et al., *A Preliminary Study of the Ability of the 4Kscore test, the Prostate Cancer Prevention Trial-Risk Calculator and the European Research Screening Prostate-Risk Calculator for Predicting High-Grade Prostate Cancer.* Actas Urol Esp, 2016. **40**(3): p. 155-63.

179. Lin, D.W., et al., *Evaluating the Four Kallikrein Panel of the 4Kscore for Prediction of*

High-grade Prostate Cancer in Men in the Canary Prostate Active Surveillance Study. Eur Urol, 2017. **72**(3): p. 448-454.

180. Vickers, A.J., E.A. Vertosick, and D.D. Sjoberg, *Value of a Statistical Model Based on Four Kallikrein Markers in Blood, Commercially Available as 4Kscore, in All Reasonable Prostate Biopsy Subgroups.* Eur Urol, 2018. **74**(4): p. 535-536.
181. L, V.N., *Detection of High-grade Prostate Cancer Using a Urinary Molecular Biomarker.* Eur Urol, 2016. **70**(5): p. 740-748.
182. TM, G., *Cost-Effectiveness of Urinary Biomarker Panel in Prostate Cancer Risk Assessment.* J Urol, 2018(10): p. 10-16.
183. McKiernan, J., et al., *A novel urine exosome gene expression assay to predict high-grade prostate cancer at initial biopsy.* JAMA Oncol, 2016. **2**.
184. McKiernan, J., et al., *A Prospective Adaptive Utility Trial to Validate Performance of a Novel Urine Exosome Gene Expression Assay to Predict High-grade Prostate Cancer in Patients with Prostate-specific Antigen 2-10ng/ml at Initial Biopsy.* Eur Urol, 2018. **74**(6): p. 731-738.
185. Fujita, K. and N. Nonomura, *Urinary biomarkers of prostate cancer.* Int J Urol, 2018. **25**(9): p. 770-779.
186. Sanda, M.G., et al., *Clinically Localized Prostate Cancer: AUA/ASTRO/SUO Guideline. Part I: Risk Stratification, Shared Decision Making, and Care Options.* Journal of Urology, 2018. **199**(3): p. 683-690.
187. Lowrance, W.T., et al., *Castration-Resistant Prostate Cancer: AUA Guideline Amendment 2015.* J Urol, 2016. **195**(5): p. 1444-1452.
188. Scher, H.I., et al., *Increased Survival with Enzalutamide in Prostate Cancer after Chemotherapy.* New England Journal of Medicine, 2012. **367**(13): p. 1187-1197.
189. Fizazi, K., et al., *Abiraterone acetate for treatment of metastatic castration-resistant prostate cancer: final overall survival analysis of the COU-AA-301 randomised, double-blind, placebo-controlled phase 3 study.* Lancet Oncology, 2012. **13**(10): p. 983-992.
190. Martin, G.A., A.H. Chen, and K. Parikh, *A Novel Use of Olaparib for the Treatment of Metastatic Castration-Recurrent Prostate Cancer.* Pharmacotherapy, 2017. **37**(11): p. 1406-1414.
191. Kantoff, P.W., et al., *Sipuleucel-T Immunotherapy for Castration-Resistant Prostate Cancer.* New England Journal of Medicine, 2010. **363**(5): p. 411-422.
192. Sheikh, N.A., et al., *Sipuleucel-T immune parameters correlate with survival: an analysis of the randomized phase 3 clinical trials in men with castration-resistant prostate cancer.* Cancer Immunology Immunotherapy, 2013. **62**(1): p. 137-147.
193. Tannock, I.F., et al., *Docetaxel plus prednisone or mitoxantrone plus prednisone for advanced prostate cancer.* New England Journal of Medicine, 2004. **351**(15): p. 1502-1512.
194. James, N.D., et al., *Docetaxel and/or zoledronic acid for hormone-naive prostate cancer: First overall survival results from STAMPEDE (NCT00268476).* Journal of Clinical Oncology, 2015. **33**(15).
195. Parker, C., et al., *Alpha Emitter Radium-223 and Survival in Metastatic Prostate Cancer.* New England Journal of Medicine, 2013. **369**(3): p. 213-223.
196. Wedge, D.C., et al., *Sequencing of prostate cancers identifies new cancer genes, routes of*

progression and drug targets. Nat Genet, 2018. **50**(5): p. 682-692.
197. Kittai, A.S., J. Blank, and J.N. Graff, *Gonadotropin-Releasing Hormone Antagonists in Prostate Cancer.* Oncology (Williston Park), 2018. **32**(12): p. 599-602, 604-6.
198. Ahmann, F.R., et al., *ZOLADEX - A SUSTAINED-RELEASE, MONTHLY LUTEINIZING-HORMONE-RELEASING HORMONE ANALOG FOR THE TREATMENT OF ADVANCED PROSTATE-CANCER.* Journal of Clinical Oncology, 1987. **5**(6): p. 912-917.
199. Conn, P.M. and W.F. Crowley, Jr., *Gonadotropin-releasing hormone and its analogues.* N Engl J Med, 1991. **324**(2): p. 93-103.
200. Baba, Y., H. Matsuo, and A.V. Schally, *Structure of the porcine LH- and FSH-releasing hormone. II. Confirmation of the proposed structure by conventional sequential analyses.* Biochem Biophys Res Commun, 1971. **44**(2): p. 459-63.
201. Parmar, H., et al., *Randomised controlled study of orchidectomy vs long-acting D-Trp-6-LHRH microcapsules in advanced prostatic carcinoma.* Lancet, 1985. **2**(8466): p. 1201-5.
202. Cook, T. and W.P. Sheridan, *Development of GnRH antagonists for prostate cancer: new approaches to treatment.* Oncologist, 2000. **5**(2): p. 162-8.
203. Boccon-Gibod, L., E. van der Meulen, and B.E. Persson, *An update on the use of gonadotropin-releasing hormone antagonists in prostate cancer.* Ther Adv Urol, 2011. **3**(3): p. 127-40.
204. Goldspiel, B.R. and D.R. Kohler, *Flutamide: an antiandrogen for advanced prostate cancer.* Dicp, 1990. **24**(6): p. 616-23.
205. Blackledge, G.R., I.D. Cockshott, and B.J. Furr, *Casodex (bicalutamide): overview of a new antiandrogen developed for the treatment of prostate cancer.* Eur Urol, 1997. **31 Suppl 2**: p. 30-9.
206. *Maximum androgen blockade in advanced prostate cancer: an overview of 22 randomised trials with 3283 deaths in 5710 patients. Prostate Cancer Trialists' Collaborative Group.* Lancet, 1995. **346**(8970): p. 265-9.
207. Sack, J.S., et al., *Crystallographic structures of the ligand-binding domains of the androgen receptor and its T877A mutant complexed with the natural agonist dihydrotestosterone.* Proceedings of the National Academy of Sciences of the United States of America, 2001. **98**(9): p. 4904-4909.
208. Hara, T., et al., *Possible role of adaptive mutation in resistance to antiandrogen in prostate cancer cells.* Prostate, 2005. **65**(3): p. 268-275.
209. Joseph, J.D., et al., *A Clinically Relevant Androgen Receptor Mutation Confers Resistance to Second-Generation Antiandrogens Enzalutamide and ARN-509.* Cancer Discovery, 2013. **3**(9).
210. Smith, M.R., et al., *Apalutamide Treatment and Metastasis-free Survival in Prostate Cancer.* N Engl J Med, 2018. **378**(15): p. 1408-1418.
211. Smith, M.R., et al., *Apalutamide Treatment and Metastasis-free Survival in Prostate Cancer.* New England Journal of Medicine, 2018. **378**(15): p. 1408-1418.
212. Wade, C.A. and N. Kyprianou, *Profiling Prostate Cancer Therapeutic Resistance.* Int J Mol Sci, 2018. **19**(3).
213. Shore, N.D., *Darolutamide (ODM-201) for the treatment of prostate cancer.* Expert Opin Pharmacother, 2017. **18**(9): p. 945-952.

214. Fizazi, K., et al., *Darolutamide in Nonmetastatic, Castration-Resistant Prostate Cancer.* N Engl J Med, 2019. **380**(13): p. 1235-1246.
215. Antonarakis, E.S., et al., *AR-V7 and Resistance to Enzalutamide and Abiraterone in Prostate Cancer.* New England Journal of Medicine, 2014. **371**(11): p. 1028-1038.
216. Liu, C., et al., *Niclosamide inhibits androgen receptor variants expression and overcomes enzalutamide resistance in castration-resistant prostate cancer.* Clin Cancer Res, 2014. **20**(12): p. 3198-3210.
217. Fang, Y., et al., *Hsp90 regulates androgen receptor hormone binding affinity in vivo.* J Biol Chem, 1996. **271**(45): p. 28697-702.
218. Solit, D.B., H.I. Scher, and N. Rosen, *Hsp90 as a therapeutic target in prostate cancer.* Semin Oncol, 2003. **30**(5): p. 709-16.
219. Trepel, J., et al., *Targeting the dynamic HSP90 complex in cancer.* Nat Rev Cancer, 2010. **10**(8): p. 537-49.
220. Ferraldeschi, R., et al., *Second-Generation HSP90 Inhibitor Onalespib Blocks mRNA Splicing of Androgen Receptor Variant 7 in Prostate Cancer Cells.* Cancer Research, 2016. **76**(9): p. 2731-2742.
221. Pacey, S., et al., *A Phase I Study of the Heat Shock Protein 90 Inhibitor Alvespimycin (17-DMAG) Given Intravenously to Patients with Advanced Solid Tumors.* Clinical Cancer Research, 2011. **17**(6): p. 1561-1570.
222. Thakur, M.K., et al., *A phase II trial of ganetespib, a heat shock protein 90 Hsp90) inhibitor, in patients with docetaxel-pretreated metastatic castrate-resistant prostate cancer (CRPC)-a prostate cancer clinical trials consortium (PCCTC) study.* Invest New Drugs, 2016. **34**(1): p. 112-8.
223. Ganai, S.A., *Histone deacetylase inhibitor sulforaphane: The phytochemical with vibrant activity against prostate cancer.* Biomed Pharmacother, 2016. **81**: p. 250-257.
224. Dong, X.Y., et al., *FOXO1A is a candidate for the 13q14 tumor suppressor gene inhibiting androgen receptor signaling in prostate cancer.* Cancer Res, 2006. **66**(14): p. 6998-7006.
225. Bohrer, L.R., et al., *FOXO1 binds to the TAU5 motif and inhibits constitutively active androgen receptor splice variants.* Prostate, 2013. **73**(10): p. 1017-1027.
226. Dehm, S.M., et al., *Selective role of an NH2-Terminal WxxLF motif for aberrant androgen receptor activation in androgen depletion-independent prostate cancer cells.* Cancer Research, 2007. **67**(20): p. 10067-10077.
227. de Bono, J.S., et al., *Randomized Phase II Study Evaluating Akt Blockade with Ipatasertib, in Combination with Abiraterone, in Patients with Metastatic Prostate Cancer with and without PTEN Loss.* Clin Cancer Res, 2019. **25**(3): p. 928-936.
228. Berthold, D.R., C.N. Sternberg, and I.F. Tannock, *Management of advanced prostate cancer after first-line chemotherapy.* Journal of Clinical Oncology, 2005. **23**(32): p. 8247-8252.
229. Stanbrough, M., et al., *Increased expression of genes converting adrenal androgens to testosterone in androgen-independent prostate cancer.* Cancer Res, 2006. **66**(5): p. 2815-25.
230. Kramer, B.S., et al., *Use of 5-alpha-Reductase Inhibitors for Prostate Cancer Chemoprevention: American Society of Clinical Oncology/American Urological Association 2008 Clinical Practice Guideline.* Journal of Clinical Oncology, 2009. **27**(9): p. 1502-1516.

231. Li, Z., et al., *Redirecting abiraterone metabolism to fine-tune prostate cancer anti-androgen therapy.* Nature, 2016. **533**(7604): p. 547-51.

232. Bryant, H.E., et al., *Specific killing of BRCA2-deficient tumours with inhibitors of poly(ADP-ribose) polymerase.* Nature, 2005. **434**(7035): p. 913-7.

233. Fong, P.C., et al., *Poly(ADP)-Ribose Polymerase Inhibition: Frequent Durable Responses in BRCA Carrier Ovarian Cancer Correlating With Platinum-Free Interval.* Journal of Clinical Oncology, 2010. **28**(15): p. 2512-2519.

234. Tutt, A., et al., *Oral poly(ADP-ribose) polymerase inhibitor olaparib in patients with BRCA1 or BRCA2 mutations and advanced breast cancer: a proof-of-concept trial.* Lancet, 2010. **376**(9737): p. 235-244.

235. Morales, J., et al., *Review of poly (ADP-ribose) polymerase (PARP) mechanisms of action and rationale for targeting in cancer and other diseases.* Crit Rev Eukaryot Gene Expr, 2014. **24**(1): p. 15-28.

236. Haffner, M.C., et al., *Transcription-Induced DNA Double Strand Breaks: Both Oncogenic Force and Potential Therapeutic Target?* Clinical Cancer Research, 2011. **17**(12): p. 3858-3864.

237. Han, S., et al., *Targeted radiosensitization of ETS fusion-positive prostate cancer through PARP1 inhibition.* Neoplasia, 2013. **15**(10): p. 1207-17.

238. Asim, M., et al., *Synthetic lethality between androgen receptor signalling and the PARP pathway in prostate cancer.* Nature Communications, 2017. **8**.

239. Bauer, C.M., et al., *Hereditary prostate cancer as a feature of Lynch Syndrome.* Familial Cancer, 2011. **10**(1): p. 37-42.

240. Rosty, C., et al., *High prevalence of mismatch repair deficiency in prostate cancers diagnosed in mismatch repair gene mutation carriers from the colon cancer family registry.* Familial Cancer, 2014. **13**(4): p. 573-582.

241. Le, D.T., et al., *Mismatch repair deficiency predicts response of solid tumors to PD-1 blockade.* Science, 2017. **357**(6349): p. 409-413.

242. Lu, X., et al., *Effective combinatorial immunotherapy for castration-resistant prostate cancer.* Nature, 2017. **543**(7647): p. 728-+.

243. Robinson, D., et al., *Integrative clinical genomics of advanced prostate cancer.* Cell, 2015. **161**(5): p. 1215-1228.

244. Wang, S.Y., et al., *Prostate-specific deletion of the murine Pten tumor suppressor gene leads to metastatic prostate cancer.* Cancer Cell, 2003. **4**(3): p. 209-221.

245. Chen, Z.B., et al., *Crucial role of p53-dependent cellular senescence in suppression of Pten-deficient tumorigenesis.* Nature, 2005. **436**(7051): p. 725-730.

246. Zong, Y., et al., *ETS family transcription factors collaborate with alternative signaling pathways to induce carcinoma from adult murine prostate cells.* Proceedings of the National Academy of Sciences of the United States of America, 2009. **106**(30): p. 12465-12470.

247. Mateo, J., et al., *A First-Time-in-Human Study of GSK2636771, a Phosphoinositide 3 Kinase Beta-Selective Inhibitor, in Patients with Advanced Solid Tumors.* Clinical Cancer Research, 2017. **23**(19): p. 5981-5992.

248. Schwartz, S., et al., *Feedback Suppression of PI3K alpha Signaling in PTEN-Mutated Tumors Is Relieved by Selective Inhibition of PI3K beta.* Cancer Cell, 2015. **27**(1): p. 109-122.

249. Wei, X.X., et al., *A Phase I Study of Abiraterone Acetate Combined with BEZ235, a Dual PI3K/mTOR Inhibitor, in Metastatic Castration Resistant Prostate Cancer.* Oncologist, 2017. **22**(5): p. 503-e43.
250. Ahmad, I., et al., *HER2 overcomes PTEN (loss)-induced senescence to cause aggressive prostate cancer.* Proc Natl Acad Sci U S A, 2011. **108**(39): p. 16392-7.
251. Di Mitri, D., et al., *Tumour-infiltrating Gr-1(+) myeloid cells antagonize senescence in cancer.* Nature, 2014. **515**(7525): p. 134-+.
252. Garcia, A.J., et al., *Pten Null Prostate Epithelium Promotes Localized Myeloid-Derived Suppressor Cell Expansion and Immune Suppression during Tumor Initiation and Progression.* Molecular and Cellular Biology, 2014. **34**(11): p. 2017-2028.
253. Wang, G., et al., *Targeting YAP-Dependent MDSC Infiltration Impairs Tumor Progression.* Cancer Discovery, 2016. **6**(1): p. 80-95.
254. Abeshouse, A., et al., *The Molecular Taxonomy of Primary Prostate Cancer.* Cell, 2015. **163**(4): p. 1011-1025.
255. Chinnaiyan, A.M., *Role of the TMPRSS2-ERG gene fusion in prostate cancer.* Neoplasia, 2008. **10**(2): p. 177-U23.
256. Yeh, J.E., P.A. Toniolo, and D.A. Frank, *Targeting transcription factors: promising new strategies for cancer therapy.* Current Opinion in Oncology, 2013. **25**(6): p. 652-658.
257. Nhili, R., et al., *Targeting the DNA-binding activity of the human ERG transcription factor using new heterocyclic dithiophene diamidines.* Nucleic Acids Research, 2013. **41**(1): p. 125-138.
258. Shao, L., et al., *Highly Specific Targeting of the TMPRSS2/ERG Fusion Gene Using Liposomal Nanovectors.* Clinical Cancer Research, 2012. **18**(24): p. 6648-6657.
259. Vainio, P., et al., *Arachidonic Acid Pathway Members PLA2G7, HPGD, EPHX2, and CYP4F8 Identified as Putative Novel Therapeutic Targets in Prostate Cancer.* American Journal of Pathology, 2011. **178**(2): p. 525-536.
260. Vainio, P., et al., *Phospholipase PLA2G7, associated with aggressive prostate cancer, promotes prostate cancer cell migration and invasion and is inhibited by statins.* Oncotarget, 2011. **2**(12): p. 1176-1190.
261. Erkizan, H.V., et al., *A small molecule blocking oncogenic protein EWS-FLI1 interaction with RNA helicase A inhibits growth of Ewing's sarcoma.* Nature Medicine, 2009. **15**(7): p. 750-U8.
262. Rahim, S., et al., *YK-4-279 Inhibits ERG and ETV1 Mediated Prostate Cancer Cell Invasion.* Plos One, 2011. **6**(4).
263. Leroy, B., M. Anderson, and T. Soussi, *TP53 mutations in human cancer: database reassessment and prospects for the next decade.* Hum. Mutat., 2014. **35**: p. 672-688.
264. Vousden, K.H. and C. Prives, *Blinded by the light: the growing complexity of p53.* Cell, 2009. **137**: p. 413-431.
265. Yingjian Liang, J.L.Z.F., *The regulation of cellular metabolism by tumor suppressor p53.* Cell & Bioscience volume 2013. **3**(9).
266. Walerych, D., *Proteasome machinery is instrumental in a common gain-of-function program of the p53 missense mutants in cancer.* Nat. Cell Biol., 2016. **18**: p. 897-909.
267. Xu, J., *Gain of function of mutant p53 by coaggregation with multiple tumor suppressors.* Nat.

Chem. Biol., 2011. **7**: p. 285-295.

268. Brosh, R. and V. Rotter, *When mutants gain new powers: news from the mutant p53 field.* Nat. Rev. Cancer, 2009. **9**: p. 701-713.
269. Kravchenko, J.E., *Small-molecule RETRA suppresses mutant p53-bearing cancer cells through a p73-dependent salvage pathway.* Proc. Natl Acad. Sci. Usa, 2008. **105**: p. 6302-6307.
270. Gourley, C., *PISARRO: a EUTROC phase Ib study of APR-246 in combination with carboplatin (C) and pegylated liposomal doxorubicin (PLD) in platinum sensitive relapsed high grade serous ovarian cancer (HGSOC) [abstract].* J. Clin. Oncol., 2016. **34**: p. 5571.
271. Salim, K.Y., et al., *COTI-2, a new anticancer drug currently under clinical investigation, targets mutant p53 and negatively modulates the PI3K/AKT/mTOR pathway.* Eur. J. Cancer, 2016. **69**: p. S19.
272. Rajan, P., et al., *Next-generation Sequencing of Advanced Prostate Cancer Treated with Androgen-deprivation Therapy.* European Urology, 2014. **66**(1): p. 32-39.
273. Murillo-Garzon, V. and R. Kypta, *WNT signalling in prostate cancer.* Nat Rev Urol, 2017. **14**(11): p. 683-696.
274. Liu, J., et al., *Targeting Wnt-driven cancer through the inhibition of Porcupine by LGK974.* Proceedings of the National Academy of Sciences of the United States of America, 2013. **110**(50): p. 20224-20229.
275. Mulholland, D.J., et al., *Pten loss and RAS/MAPK activation cooperate to promote EMT and metastasis initiated from prostate cancer stem/progenitor cells.* Cancer Res, 2012. **72**(7): p. 1878-89.
276. Wanjala, J., et al., *Identifying actionable targets through integrative analyses of GEM model and human prostate cancer genomic profiling.* Mol Cancer Ther, 2015. **14**(1): p. 278-88.
277. Palanisamy, N., et al., *Rearrangements of the RAF kinase pathway in prostate cancer, gastric cancer and melanoma.* Nature Medicine, 2010. **16**(7): p. 793-U94.
278. Gravina, G.L., et al., *Long-term presence of androgens and anti-androgens modulate EGF-receptor expression and MAP-kinase phosphorylation in androgen receptor-prostate positive cancer cells.* Int J Oncol, 2004. **25**(1): p. 97-104.
279. Aytes, A., et al., *ETV4 promotes metastasis in response to activation of PI3-kinase and Ras signaling in a mouse model of advanced prostate cancer.* Proceedings of the National Academy of Sciences of the United States of America, 2013. **110**(37): p. E3506-E3515.
280. Torre, L.A., et al., - *Global cancer statistics, 2012.*2015. **- 65**.
281. Budroni, M., et al., *Estimates of cancer burden in Sardinia.* Tumori, 2013. **99**(3): p. 408-15.
282. Chiang, C.W.K., et al., *Genomic history of the Sardinian population.* Nat Genet, 2018. **50**(10): p. 1426-1434.
283. Li, H. and R. Durbin, *Fast and accurate short read alignment with Burrows-Wheeler transform.* Bioinformatics, 2009. **25**(14): p. 1754-60.
284. Li, H., et al., *The Sequence Alignment/Map format and SAMtools.* Bioinformatics, 2009. **25**(16): p. 2078-9.
285. McKenna, A., et al., *The Genome Analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data.* Genome Res, 2010. **20**(9): p. 1297-303.
286. Ye, K., et al., *Pindel: a pattern growth approach to detect break points of large deletions and*

medium sized insertions from paired-end short reads. Bioinformatics, 2009. **25**(21): p. 2865-71.

287. Ramos, A.H., et al., *Oncotator: cancer variant annotation tool.* Hum Mutat, 2015. **36**(4): p. E2423-9.
288. Chakravarty, D., et al., *OncoKB: A Precision Oncology Knowledge Base.* JCO Precis Oncol, 2017. **2017**.
289. Tamborero, D., A. Gonzalez-Perez, and N. Lopez-Bigas, *OncodriveCLUST: exploiting the positional clustering of somatic mutations to identify cancer genes.* Bioinformatics, 2013. **29**(18): p. 2238-44.
290. Chang, M.T., et al., *Accelerating Discovery of Functional Mutant Alleles in Cancer.* Cancer Discov, 2018. **8**(2): p. 174-183.
291. Gao, J., et al., *3D clusters of somatic mutations in cancer reveal numerous rare mutations as functional targets.* Genome Med, 2017. **9**(1): p. 4.
292. Liu, X., et al., *dbNSFP v3.0: A One-Stop Database of Functional Predictions and Annotations for Human Nonsynonymous and Splice-Site SNVs.* Hum Mutat, 2016. **37**(3): p. 235-41.
293. Yuan, S., et al., *PyMOL and Inkscape Bridge the Data and the Data Visualization.* Structure, 2016. **24**(12): p. 2041-2042.
294. Van der Auwera, G.A., et al., *From FastQ data to high confidence variant calls: the Genome Analysis Toolkit best practices pipeline.* Curr Protoc Bioinformatics, 2013. **43**: p. 11 10 1-33.
295. Rainville, I. and J.E. Garber, *Familial cancer database online.* Lancet Oncol, 2008. **9**(10): p. 925-6.
296. Ge, H., et al., *FusionMap: detecting fusion genes from next-generation sequencing data at base-pair resolution.* Bioinformatics, 2011. **27**(14): p. 1922-8.
297. Zhou, Y., et al., *Metascape provides a biologist-oriented resource for the analysis of systems-level datasets.* Nat Commun, 2019. **10**(1): p. 1523.
298. Boeva, V., et al., *Control-FREEC: a tool for assessing copy number and allelic content using next-generation sequencing data.* Bioinformatics, 2012. **28**(3): p. 423-5.
299. Wang, K., M. Li, and H. Hakonarson, *ANNOVAR: functional annotation of genetic variants from high-throughput sequencing data.* Nucleic Acids Res, 2010. **38**(16): p. e164.
300. Mermel, C.H., et al., *GISTIC2.0 facilitates sensitive and confident localization of the targets of focal somatic copy-number alteration in human cancers.* Genome Biol, 2011. **12**(4): p. R41.
301. Tang, Z., et al., *GEPIA2: an enhanced web server for large-scale expression profiling and interactive analysis.* Nucleic Acids Res, 2019.
302. Barbieri, C.E., et al., *Exome sequencing identifies recurrent SPOP, FOXA1 and MED12 mutations in prostate cancer.* Nat Genet, 2012. **44**(6): p. 685-9.
303. Bielski, C.M., et al., *Genome doubling shapes the evolution and prognosis of advanced cancers.* Nat Genet, 2018. **50**(8): p. 1189-1195.
304. Hieronymus, H., et al., *Copy number alteration burden predicts prostate cancer relapse.* Proc Natl Acad Sci U S A, 2014. **111**(30): p. 11139-44.
305. Ren, S., et al., *Whole-genome and Transcriptome Sequencing of Prostate Cancer Identify New Genetic Alterations Driving Disease Progression.* Eur Urol, 2017.
306. Ali Shah, S.I., *Emerging potential of parenteral estrogen as androgen deprivation therapy for*

prostate cancer. South Asian J Cancer, 2015. **4**(2): p. 95-7.

307. Di Zazzo, E., et al., *Prostate cancer stem cells: the role of androgen and estrogen receptors.* Oncotarget, 2016. **7**(1): p. 193-208.

308. Nelles, J.L., W.Y. Hu, and G.S. Prins, *Estrogen action and prostate cancer.* Expert Rev Endocrinol Metab, 2011. **6**(3): p. 437-451.

309. Petrovics, G., et al., *A novel genomic alteration of LSAMP associates with aggressive prostate cancer in African American men.* EBioMedicine, 2015. **2**(12): p. 1957-64.

310. Rosen, P., et al., *Differences in frequency of ERG oncoprotein expression between index tumors of Caucasian and African American patients with prostate cancer.* Urology, 2012. **80**(4): p. 749-53.

311. Sven Perner, F.D., Rameen Beroukhim, Folke H. Schmidt, Juan-Miguel Mosquera, Sunita Setlur, Joelle Tchinda, Scott A. Tomlins, Matthias D. Hofer, Kenneth G. Pienta, Rainer Kuefer, Robert Vessella, Xiao-Wei Sun, Matthew Meyerson, *TMPRSS2:ERG Fusion-Associated Deletions Provide Insight into the Heterogeneity of Prostate Cancer.* Cancer Research, 2006. **66**(17).

312. Onodera, T., et al., *Btbd7 regulates epithelial cell dynamics and branching morphogenesis.* Science, 2010. **329**(5991): p. 562-5.

313. Fang, L.Z., et al., *Silencing of Btbd7 Inhibited Epithelial-Mesenchymal Transition and Chemoresistance in CD133(+) Lung Carcinoma A549 Cells.* Oncol Res, 2017. **25**(5): p. 819-829.

314. Grisard, E., et al., *Sleeping beauty genetic screen identifies miR-23b::BTBD7 gene interaction as crucial for colorectal cancer metastasis.* EBioMedicine, 2019.

315. Tao, Y.M., et al., *BTB/POZ domain-containing protein 7: epithelial-mesenchymal transition promoter and prognostic biomarker of hepatocellular carcinoma.* Hepatology, 2013. **57**(6): p. 2326-37.

316. Jin, X., et al., *An essential role for GLUT5-mediated fructose utilization in exacerbating the malignancy of clear cell renal cell carcinoma.* Cell Biol Toxicol, 2019.

317. Weng, Y., et al., *SLC2A5 promotes lung adenocarcinoma cell growth and metastasis by enhancing fructose utilization.* Cell Death Discov, 2018. **4**: p. 38.

318. Zhao, P., et al., *SLC2A5 overexpression in childhood philadelphia chromosome-positive acute lymphoblastic leukaemia.* Br J Haematol, 2018. **183**(2): p. 242-250.

319. Rowland, A., J.O. Miners, and P.I. Mackenzie, *The UDP-glucuronosyltransferases: their role in drug metabolism and detoxification.* Int J Biochem Cell Biol, 2013. **45**(6): p. 1121-32.

320. Dutton, G.J., *Developmental aspects of drug conjugation, with special reference to glucuronidation.* Annu Rev Pharmacol Toxicol, 1978. **18**: p. 17-35.

321. Grant, D.J., et al., *UDP-glucuronosyltransferases and biochemical recurrence in prostate cancer progression.* BMC Cancer, 2017. **17**(1): p. 463.

322. Paquet, S., et al., *Differential Expression of the Androgen-Conjugating UGT2B15 and UGT2B17 Enzymes in Prostate Tumor Cells during Cancer Progression.* Journal of Clinical Endocrinology & Metabolism, 2012. **97**(3): p. E428-E432.

323. Li, H., et al., *UGT2B17 Expedites Progression of Castration-Resistant Prostate Cancers by Promoting Ligand-Independent AR Signaling.* Cancer Research, 2016. **76**(22): p. 6701-6711.

324. Grant, D.J., et al., *UDP-glucuronosyltransferases and biochemical recurrence in prostate cancer progression.* Bmc Cancer, 2017. **17**.

325. Barre, L., et al., *Substrate specificity of the human UDP-glucuronosyltransferase UGT2B4 and UGT2B7 - Identification of a critical aromatic amino acid residue at position 33.* Febs Journal, 2007. **274**(5): p. 1256-1264.
326. Sun, C., et al., *A signature of balancing selection in the region upstream to the human UGT2B4 gene and implications for breast cancer risk.* Human Genetics, 2011. **130**(6): p. 767-775.
327. Donati, S., T. Sander, and H. Link, *Crosstalk between transcription and metabolism: how much enzyme is enough for a cell?* Wiley Interdisciplinary Reviews-Systems Biology and Medicine, 2018. **10**(1).
328. Peng, X., et al., *Molecular Characterization and Clinical Relevance of Metabolic Expression Subtypes in Human Cancers.* Cell Reports, 2018. **23**(1): p. 255-+.
329. Meech, R., et al., *The glycosidation of xenobiotics and endogenous compounds: Versatility and redundancy in the UDP glycosyltransferase superfamily.* Pharmacology & Therapeutics, 2012. **134**(2): p. 200-218.
330. Alsamman, K. and O.S. El-Masry, *Interferon regulatory factor 1 inactivation in human cancer.* Biosci Rep, 2018. **38**(3).
331. Lee, J., et al., *Lineage specification of human dendritic cells is marked by IRF8 expression in hematopoietic stem cells and multipotent progenitors.* Nat Immunol, 2017. **18**(8): p. 877-888.
332. Wang, H., et al., *A reporter mouse reveals lineage-specific and heterogeneous expression of IRF8 during lymphoid and myeloid cell differentiation.* J Immunol, 2014. **193**(4): p. 1766-77.
333. Ozawa, M.G., et al., *A study of the mutational landscape of pediatric-type follicular lymphoma and pediatric nodal marginal zone lymphoma.* Mod Pathol, 2016. **29**(10): p. 1212-20.
334. Adams, N.M., et al., *Transcription Factor IRF8 Orchestrates the Adaptive Natural Killer Cell Response.* Immunity, 2018. **48**(6): p. 1172-1182 e6.
335. Waight, J.D., et al., *Regulation of the interferon regulatory factor-8 (IRF-8) tumor suppressor gene by the signal transducer and activator of transcription 5 (STAT5) transcription factor in chronic myeloid leukemia.* J Biol Chem, 2014. **289**(22): p. 15642-52.
336. Mace, E.M., et al., *Biallelic mutations in IRF8 impair human NK cell maturation and function.* J Clin Invest, 2017. **127**(1): p. 306-320.
337. Schmidt, J., et al., *Mutations of MAP2K1 are frequent in pediatric-type follicular lymphoma and result in ERK pathway activation.* Blood, 2017. **130**(3): p. 323-327.
338. Zhang, J., et al., *Genetic heterogeneity of diffuse large B-cell lymphoma.* Proc Natl Acad Sci U S A, 2013. **110**(4): p. 1398-403.
339. Otto, N., et al., *ICSBP promoter methylation in myelodysplastic syndromes and acute myeloid leukaemia.* Leukemia, 2011. **25**(7): p. 1202-7.
340. Hjelmborg, J.B., et al., *The Heritability of Prostate Cancer in the Nordic Twin Study of Cancer.* Cancer Epidemiology Biomarkers & Prevention, 2014. **23**(11): p. 2303-2310.
341. Hoffmann, T.J., et al., *A Large Multiethnic Genome-Wide Association Study of Prostate Cancer Identifies Novel Risk Variants and Substantial Ethnic Differences.* Cancer Discovery, 2015. **5**(8): p. 878-891.
342. Spagopoulou, M.I.L.F., *Evolutionary consequences of epigenetic inheritance.* Heredity, 2018. **121**: p. 205-209.

343. Koboldt, D.C., et al., *The Next-Generation Sequencing Revolution and Its Impact on Genomics.* Cell, 2013. **155**(1): p. 27-38.
344. Stott-Miller, M., et al., *HOXB13 mutations in a population-based, case-control study of prostate cancer.* Prostate, 2013. **73**(6): p. 634-641.
345. Allam, R.M., et al., *Fingolimod interrupts the cross talk between estrogen metabolism and sphingolipid metabolism within prostate cancer cells.* Toxicol Lett, 2018. **291**: p. 77-85.
346. Tang, L., et al., *Associations between polymorphisms in genes related to estrogen metabolism and function and prostate cancer risk: results from the Prostate Cancer Prevention Trial.* Carcinogenesis, 2018. **39**(2): p. 125-133.