Integrative analysis of deregulated networks in prostate cancer for personalized treatment

Yasenya Kasikci

UNIVERSITÉ DE STRASBOURG

ÉCOLE DOCTORALE DES SCIENCE DE LA VIE ET DE LA SANTÉ

Institut de Génétique et de Biologie Moléculaire et Cellulaire (IGBMC)

CNRS UMR 7104 – Inserm U 1258

THÈSE présentée par:

Yasenya KASIKCI

Soutenue le: **19 Septembre 2023**

pour obtenir le grade de : **Docteur de l'université de Strasbourg**

Discipline/ Spécialité:Sciences de la vie et de la santé

Integrative Analysis of Deregulated Networks in Prostate Cancer for Personalized Treatment

THÈSE dirigée par:

Dr. CHAN Susan Directeur de recherche, IGBMC, Strasbourg

Dr. GRONEMEYER Hinrich Émérite, IGBMC, Strasbourg

RAPPORTEURS:

Prof. Dr. PLASS Christoph Chef de département, DKFZ, Heidelberg

Dr. WALTZER Lucas Directeur de recherche, UCA, Clermont-Ferrand

---

EXAMINATEUR:

Dr. NEBIGIL-DESAUBRY Canan Directeur de recherche, CNRS, Strasbourg

'I am not leaving a spiritual legacy of dogmas, unchangeable petrified directives. My spiritual legacy is science and reason. In this world, claiming to bring never changing judgment means denying the progression of reason and science. If those people who wish to follow me after I am gone take reason and science as their guides they will be my true spiritual heirs.'

Mustafa Kemal Atatürk

Ever since I came to know myself, I have been raised with the words of Atatürk, influenced by both my family and my teachers. I have committed myself to walk the path he illuminated, guided by the light of science. I extend my heartfelt gratitude to all those who have supported and guided me on this journey thus far.

ACKNOWLEDGMENTS

As I come to the end of writing this manuscript and reflect on the intense years that led me here, I want to express my gratitude to the people who have supported me along the way. Without them, I wouldn't have made it so far.

First and foremost, I extend my deepest appreciation to my PhD supervisors, Dr. ***Susan Chan*** and Dr. ***Hinrich Gronemeyer***, for giving me this amazing opportunity to work on an interesting project in their team. Their decision to accept me as a PhD student, their unwavering support and guidance throughout my doctoral journey, and their positive attitude have been invaluable.

I would also like to extend my heartfelt thanks to all the members of my jury, Dr. ***Christoph Plass***, Dr. ***Lucas Waltzer***, and Dr. ***Canan Nebigil-Desaubry***, for graciously agreeing and dedicating their time to read and evaluate my PhD work.

Additionally, I am profoundly grateful for the assistance provided by the IGBMC facilities, technicians, namely ***Cathie Erb*** and ***Patricia Marchal***, and the administration, particularly ***Valérie Schon***. They were kind, friendly, and always ready to support and help me.

To my friends both within and outside of IGBMC, I express my sincere appreciation for standing by my side and providing support and for the times we spent together, enjoying a few beers to lift my spirits. While it would be an extensive list to enumerate each of your names, I want to keep it concise and simply convey my love and gratitude to all.

I would also like to thank another person who has been very important to me and has had a huge impact on my life, ***Stuart Gill***. Without your support, this journey would have been considerably more challenging. Thank you for continuously placing your trust in me, supporting my accomplishments, and remaining by my side through all my ups and downs, particularly during difficult times.

Lastly, but most importantly, I want to express my deep gratitude to my mother, ***Ebru Yel***. You always believed in me, supported me every day, and gave me the energy, courage, and love I needed. *İyi ki benim annemsin ve seni çok seviyorum.*

I dedicate this work to the memory of my friends and relatives who tragically lost their lives in the earthquake on May 6, 2023, in Turkey. During this difficult time, we managed to carry on with our lives and achievements, but you left us far too soon. May you all rest in peace.

RÉSUMÉ EN FRANÇAIS

Le cancer de la prostate est le deuxième cancer le plus courant chez les hommes, et plus de 50 % des cancers de la prostate avancés présentent une fusion TMPRSS2-ERG. Malgré les vastes données génomiques et transcriptomiques sur le cancer, on en sait peu sur l'impact des mutations et de la transcription altérée sur les réseaux régulateurs chez les patients atteints de cancer de la prostate (PCa) individuels. Les études traditionnelles en biologie du cancer se concentrent souvent sur l'analyse de grands groupes de patients pour identifier les altérations moléculaires communes et les cibles potentielles pour la thérapie. Bien que cette approche ait permis des avancées significatives dans le traitement du cancer, elle néglige la diversité inhérente et l'hétérogénéité présente chez les patients individuels. La biologie de la tumeur de chaque patient est unique, et il peut y avoir une variation substantielle dans les signatures moléculaires et les voies qui favorisent la progression du cancer. Cette étude de thèse suggère qu'en adoptant une approche d'analyse en réseau, il devient possible de démêler l'interaction complexe des gènes, des protéines et des voies de signalisation au sein de la tumeur d'un individu. L'analyse en réseau implique la construction d'une carte complète des interactions moléculaires, telles que les interactions protéine-protéine, les réseaux de régulation génique et les cascades de signalisation, pour obtenir une compréhension globale de la maladie. Cette approche permet d'identifier les facteurs clés, les nœuds critiques et les cibles thérapeutiques potentielles spécifiques à chaque patient. Dans cette étude, nous avons utilisé 15 tumeurs (T) primaires de la prostate ERG-positives non traitées et des tissus normaux (N) appariés sélectionnés par un pathologiste expert. Les échantillons ont été sélectionnés sur la base de sections consécutives de la même biopsie, qui différaient seulement légèrement en termes de cellules tumorales (>80 % de cellules tumorales), tandis que la section des biopsies N de la même prostate ne présentait aucune cellule tumorale. De plus, l'immunohistochimie et l'ARN-Seq (duplicats biologiques) ont confirmé la surexpression de l'ERG par rapport aux échantillons N appariés, et tous les échantillons ont révélé une augmentation des niveaux de récepteur des androgènes (AR). Cela a permis une comparaison plus précise entre les tissus tumoraux et normaux.

Spécifiquement, des échantillons appariés de tissus normaux et tumoraux de patients ont été utilisés pour établir les variations somatiques et les profils transcriptomiques différentiels des cancers de la prostate primaires ERG-positifs. Pour l'identification des variations somatiques, nous avons réalisé un séquençage complet de l'exome et utilisé MuTect2 pour l'appel des variants. Cette analyse a révélé de 49 à 114 mutations dans chaque cancer par

rapport au tissu normal de la prostate correspondant. De plus, nous avons observé des mutations spécifiques à chaque patient, précédemment non signalées, ainsi que des mutations classiques dans tous les échantillons. Les mutations survenant dans des éléments de régulation tels que les activateurs, ainsi que des facteurs tels que les facteurs de transcription, les modulateurs épigénétiques et les enzymes, peuvent avoir un impact sur l'expression globale des gènes. Pour intégrer l'influence de ces mutations dans notre analyse en réseau, nous avons réalisé un séquençage de l'ARN total haut débit en mode brin spécifique par extrémités appariées. Nous avons effectué ce séquençage après avoir éliminé l'ARN ribosomique des sections de biopsie T et N appariées, en garantissant la duplication biologique. La présence de ces mutations génétiques significativement divergentes et les modifications conséquentes des schémas d'expression génique soulignent l'importance de comprendre les systèmes dérégulés complexes chez chaque patient individuel.

Dans un premier temps, afin d'intégrer les différentes fonctions dysrégulées au sein de chaque tumeur, nous avons réalisé des analyses d'enrichissement de voies centrées sur le patient pour les gènes différentiellement exprimés spécifiques aux tumeurs (TS-DEGs) à l'aide de Panther dans l'environnement GeneCodis3. Bien que plusieurs voies étaient communément dysrégulées dans le cancer de la prostate de plusieurs patients, notamment les voies de signalisation des cadhérines, Wnt et intégrines, ces analyses ont également démontré que dans chaque patient, différents ensembles de voies étaient dérégulés. Cette découverte a été confirmée par l'analyse de 52 patients supplémentaires provenant du dépôt TCGA. Afin d'obtenir une meilleure compréhension de la communication altérée par les dérégulations et mutations spécifiques aux patients, nous avons reconstruit des réseaux maîtres. Nous avons utilisé les données obtenues à partir du séquençage de l'exome et du RNA-sequencing pour élucider la connectivité entre les gènes différentiellement exprimés (DEGs) et les gènes mutés, ainsi que pour identifier les cibles clés dans les voies de signalisation et les réseaux spécifiques aux patients. Pour construire les réseaux maîtres, nous avons intégré les connexions fournies par des outils validés tels que STRING (pour les interactions protéine-protéine) et CellNet (pour les gènes cibles des facteurs de transcription), puis nous avons visualisé les réseaux à l'aide de Cytoscape. Au sein de ces réseaux, nous avons analysé certaines voies qui étaient à la fois significativement affectées et communément dérégulées chez nos patients, puis nous avons observé que les cibles individuelles dérégulées au sein de ces voies, comme la voie Wnt, présentaient une variation considérable. Ainsi, il est insuffisant de se fier uniquement à un inhibiteur général de la voie Wnt, par exemple, car certains facteurs Wnt peuvent déjà être

inhibés tandis que d'autres sont exprimés. Par conséquent, il devient difficile d'identifier une thérapie unifiée, ce qui souligne l'importance d'un traitement personnalisé pour les patients. Pour valider nos résultats, nous avons réalisé une analyse de RNA-seq en duplicata et confirmé plusieurs mutations par séquençage de Sanger. De plus, nous avons effectué une validation in silico pour recueillir des preuves externes corroborantes de l'exactitude des réseaux de gènes différentiellement exprimés (DEG).

Étant donné que nous avons effectué un séquençage complet de l'ARN total en double brin spécifique (Ribo-moins), ce qui nous a permis d'évaluer non seulement les ARNm mais tous les ARN non codants, l'étude a également exploré le potentiel de développement de réseaux de miARN personnalisés pour augmenter la complexité de la réseau. Les miARN jouent un rôle central dans la régulation post-transcriptionnelle des gènes, car ils se lient à la région 3' non traduite (UTR) des ARNm cibles, entraînant une dégradation de l'ARNm ou une répression traductionnelle. Ce mécanisme de régulation ajoute une couche supplémentaire de contrôle à l'expression des gènes au-delà de la régulation transcriptionnelle. En intégrant les interactions miARN-ARNm dans le réseau, nous avons pu mieux comprendre comment ces molécules coordonnent collectivement les processus et les voies cellulaires. Cette compréhension au niveau du système aide à démêler les relations complexes et la diaphonie entre les éléments de codage et de non-codage, conduisant à une vue plus complète du paysage réglementaire. Dans cette étude, en identifiant les miARN dérégulés et leurs cibles grâce à la bioinformatique et en les intégrant dans les réseaux différentiels établis pour chaque patient, nous avons découvert de grandes différences interindividuelles entre les tissus appariés et les tumeurs normales adjacentes de la prostate, similaires aux ARNm. À l'aide de ces informations, nous avons intégré les ARN non codants dérégulés dans des réseaux différentiels globaux établis pour chaque patient à l'aide d'outils informatiques. Nous avons ensuite effectué une recherche bioinformatique dans les bases de données correspondantes pour identifier les cibles des miARN et les localiser dans les réseaux différentiels des patients. Une fois les cibles définies, nous avons identifié les gènes cibles des miARN dérégulés dans les tumeurs. Nous nous sommes concentrés à la fois sur les miARN surexprimés, dont les cibles étaient régulées négativement dans les réseaux différentiels, et sur les miARN sous-exprimés, dont les cibles étaient régulées positivement. Pour explorer les implications fonctionnelles des interactions miARN-cible identifiées au sein des réseaux différentiels des patients, nous avons effectué des analyses complètes. Nous avons effectué une analyse d'enrichissement fonctionnel sur les gènes cibles des miARN dérégulés à l'aide de la base de données de la voie Panther sur le

serveur Web ShinyGO. Pour mieux comprendre les similitudes et les différences de ces gènes et voies dérégulés chez les patients, nous avons procédé à un examen plus détaillé des voies sélectionnées. Étant donné que les voies Wnt étaient couramment dérégulées entre certains patients, nous avons intégré les DEmiR dans ces réseaux Wnt spécifiques aux patients afin de mieux comprendre les mécanismes potentiels de ces miARN. L'importance de ces gènes réside dans la possibilité qu'un seul miARN puisse être responsable de la dérégulation de plusieurs gènes. Au lieu de se concentrer sur chaque gène individuellement, nous pouvons potentiellement cibler plusieurs gènes en nous concentrant sur les miARN.

Après avoir identifié des cibles potentielles de miARN et établi des réseaux de régulation miARN-ARNm, nous avons procédé à la validation des effets régulateurs des miARN sur leurs gènes cibles par le biais d'essais expérimentaux. Nous avons effectué des dosages doubles de la luciférase et des dosages RT-PCR à cette fin. Nous nous sommes concentrés sur deux facteurs de transcription, GATA6 et TOX3, car ils jouent un rôle crucial dans la régulation de l'expression génique et ont déjà été dérégulés chez plus de deux patients. Lorsque les miARN ciblent ces facteurs de transcription, leurs niveaux d'expression et leur activité peuvent être modulés, influençant ainsi l'activation ou la répression d'autres gènes au sein du réseau de régulation. Par conséquent, les miARN peuvent affecter indirectement l'expression des gènes en aval, entraînant des effets en cascade sur les processus cellulaires et contribuant potentiellement à la progression du cancer de la prostate.

Pour lancer la validation expérimentale, nous avons cloné les régions 3' UTR de GATA6 et TOX3, qui sont les gènes cibles prédits de miR-27a et miR23a dans le vecteur pmirGlo, respectivement. Ensuite, nous avons co-transfecté ces cibles clonées avec des mimiques miR-27a et miR-23a dans des lignées cellulaires PC3. En analysant leurs interactions directes, nous avons confirmé que ces deux gènes étaient bien des cibles directes des miARN respectifs. Pour valider davantage nos résultats, nous avons effectué des analyses RT-PCR pour examiner la dérégulation d'autres gènes cibles suite à la manipulation d'imitateurs ou d'inhibiteurs de miR-27a et miR-23a. Les résultats de l'analyse ont indiqué que les gènes cibles présentaient une dérégulation significative, ce qui confirme davantage nos résultats.

En conséquence, grâce à l'intégration de fonctions dérégulées au sein de chaque tumeur, nous avons identifié des facteurs critiques au sein des réseaux de régulation qui présentent une dérégulation et peuvent être ciblés par des médicaments. Ces précieuses informations ont le potentiel de guider le développement de thérapies personnalisées spécifiquement conçues pour

répondre aux besoins uniques de chaque patient. De plus, notre étude s'est penchée sur le potentiel de création de réseaux de miARN personnalisés et nous avons validé avec succès les interactions directes entre les miARN et leurs gènes cibles. Cette validation renforce notre compréhension des mécanismes de régulation complexes qui sous-tendent l'expression génique médiée par les miARN. Dans l'ensemble, ces découvertes contribuent à une compréhension plus profonde du paysage moléculaire complexe et ouvrent des possibilités pour des interventions thérapeutiques plus précises et efficaces.

CONTENTS

TABLE OF FIGURES

LIST OF ABBREVIATIONS

ADT	Androgen Deprivation Therapy
AR	Androgen Receptor
ASAP	Atypical Small Acinar Proliferation
BPH	Benign Prostatic Hyperplasia
ceRNAs	Competing Endogenous RNAs
CRPC	Castration-Resistant Prostate Cancer
DEGs	Differentially Expressed Genes
DElncRs	Differentially Expressed microRNAs
DEmiRs	Differentially Expressed Long Non-Coding RNAs
DEncRNAs	Differentially Expressed Non-Coding RNAs
dMMR	Mismatch Repair Deficient
DRE	Digital Rectal Examination
EBRT	External Beam Radiation Therapy
EMT	Epithelial-To-Mesenchymal Transition
FSH	Follicle Stimulating Hormone
GM-CSF	Granulocyte-Macrophage Colony-Stimulating Factor
GnRH	Gonadotrophin Releasing Hormone
GRNs	Gene Regulatory Networks
GWAS	Genome-Wide Association Studies
HGPIN	High-Grade Prostatic Intraepithelial Neoplasia
HIFU	High-Intensity Focused Ultrasound
HR	Homologous Recombination
LHRH	Luteinizing Hormone-Releasing Hormone
LNAs	Locked Nucleic Acids
LOH	Loss Of Heterozygosity
mpMRI	Multiparametric Magnetic Resonance Imaging
MSI-H	Microsatellite Instability-High
NLEs	Neutral Lipid Emulsions
PAP	Prostatic Acid Phosphatase
PARP	Poly(ADP-Ribose) Polymerase
PBMCs	Peripheral Blood Mononuclear Cells
PCa	Prostate Cancer
PIN	Prostatic Intraepithelial Neoplasia
PI-RADS v2	Prostate Imaging-Reporting And Data System Version 2
PSA	Prostate-Specific Antigen
RISC	RNA-Induced Silencing Complex
SNPs	Single Nucleotide Polymorphisms
TMPRSS2-ERG	The Transmembrane Protease Serine 2 : V-Ets Erythroblastosis Virus E26 Oncogene Homolog
TNM	Tumor, Node, And Metastasis
TRUS	Transrectal Ultrasound
TURP	Transurethral Resection of The Prostate
UTRs	Untranslated Regions

INTRODUCTION

Chapter I. Prostate cancer epidemiology, pathology, and prognostic factors

Chapter I. Prostate cancer epidemiology, pathology, and prognostic factors

1.1 Prostate cancer incidence, mortality, and survival

Worldwide, prostate cancer stands as the second most frequently diagnosed cancer and ranks as the fifth leading cause of cancer-related mortality in men, with an estimated 1,414,000 newly diagnosed cancer cases and 375,304 deaths in 2020. As the most frequently diagnosed cancer in 112 countries, prostate cancer is also the leading cause of cancer death in 48 countries[1]. Prostate cancer incidence and survival in men might be impacted by factors such as behavior, genetics, family history, and treatment patterns. A better understanding of differences in incidence and survival by stage, race/ethnicity, and age can improve public health planning related to screening, treatment, and survivor care[9]. Since the 1990s, several developed countries have adopted prostate-specific antigen (PSA) screening, and there have been a number of countries that have demonstrated a downward trend in prostate cancer mortality rates, such as the United States of America (USA)[10], Canada[6,11], the United Kingdom (UK)[12], and Japan[13]. Identifying high-risk groups and developing a targeted screening program for them is an appealing approach since evidence has shown that population-wide PSA testing in general population screening is beneficial; despite this, the harms are regarded as being heavier than the benefits due to the possibility of overdiagnosis and overtreatment with PSA testing. For harms to be reduced, patients must be carefully selected for screening and diagnoses should be dissociated from treatments[14].

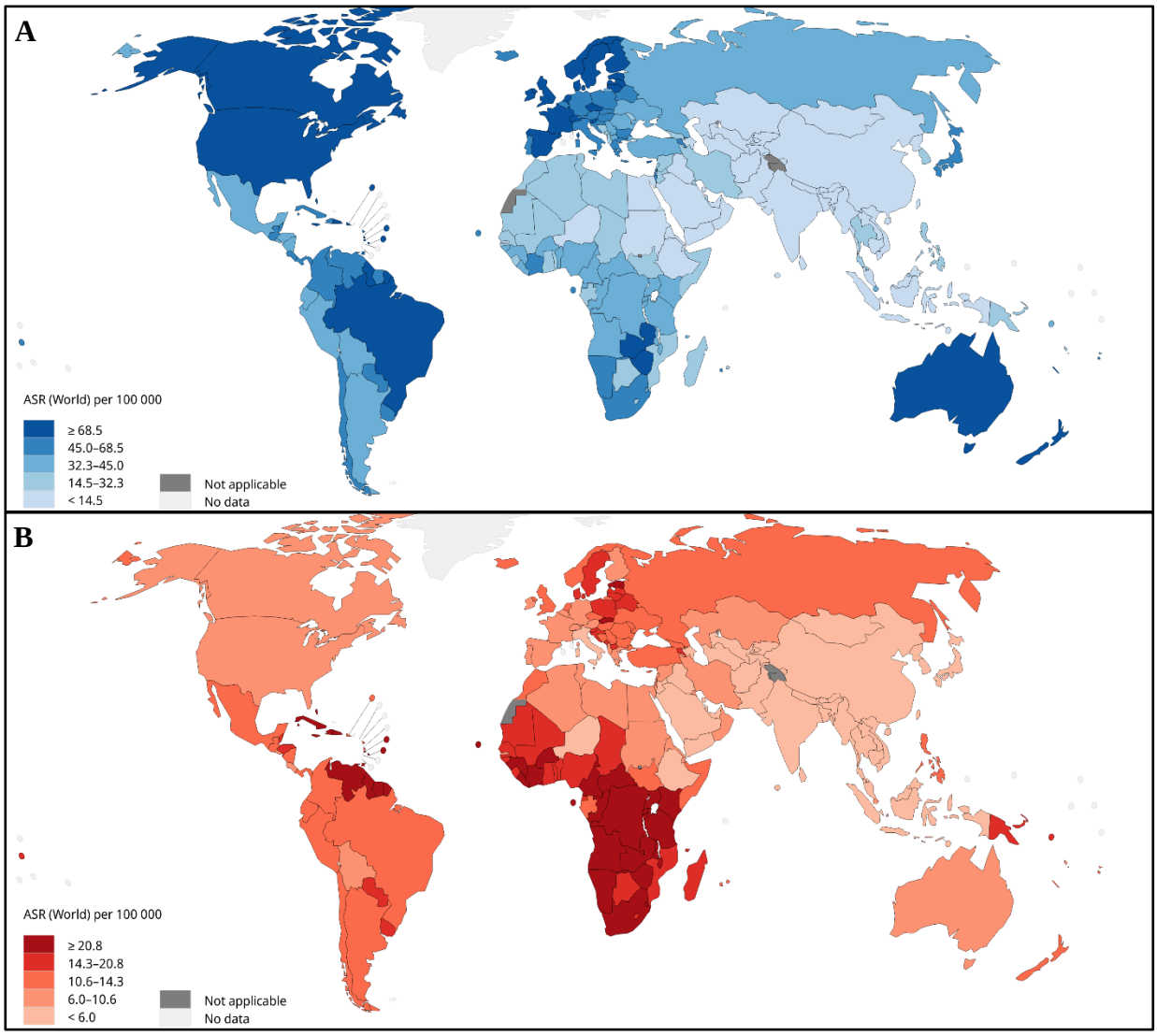


Figure 1 A) Estimated age-standardized incidence rates in World in 2020 **B)** Estimated age-standardized mortality rates in World in 2020, ASR = age-standardized rate; IARC = International Agency for Research on Cancer. Source: GLOBOCAN 2020[1]

Prostate cancer incidence and mortality vary across regions and populations. According to the GLOBOCAN 2020 data which estimates cancer incidence and mortality produced by the International Agency for Research on Cancer, the highest incidence rates are reported in North America, Oceania, and Northern and Western Europe, while lowest rates are identified in Asia and North Africa (**Figure 1**). Differences in incidence and mortality rates across countries can be attributed to various factors, including potential underdiagnosis and underreporting, disparities in healthcare access, variations in screening approaches, knowledge and awareness gaps, and attitudes towards PCa and related screening methods[15].

1.2 Normal Prostate Anatomy and Histology

1.2.1 Prostate Anatomy

The prostate is a tubulo-alveolar exocrine gland that forms part of the male reproductive system. It surrounds the neck of a man's bladder and urethra—the tube that carries urine from the bladder. It normally weighs about 15-20 grams and is shaped like a walnut (**Figure 2**).

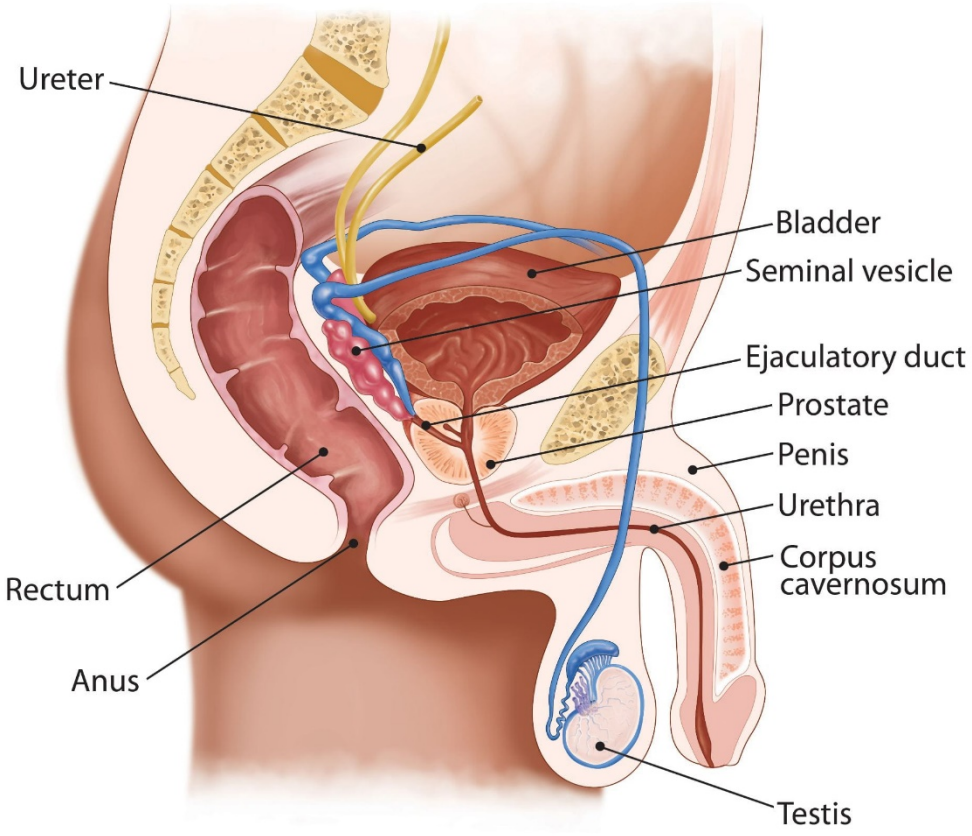


Figure 2 Anatomy of the male reproductive system [4]

McNeal JE (1981) conducted detailed studies of the normal and pathologic anatomy of the prostate and described four basic anatomic zones on the basis of biological and histological concepts[4]. There are four major zones within the normal prostate which are labeled as the peripheral zone (70% of glandular tissue), the central zone (20% of glandular tissue), the transitional zone (5-10% of glandular tissue) and the periurethral zone. The peripheral zone is the outermost region of the prostate and encircles the central zone posteroanterior and represents the most common site in the prostate for developing prostate carcinomas. The central zone surrounds the ejaculatory ducts posterior to the preprostatic urethra and makes up the majority of the prostatic base. The transition zone is the most central part of the prostate that circumscribes the distal end of the periprostatic urethra and represents the region where benign

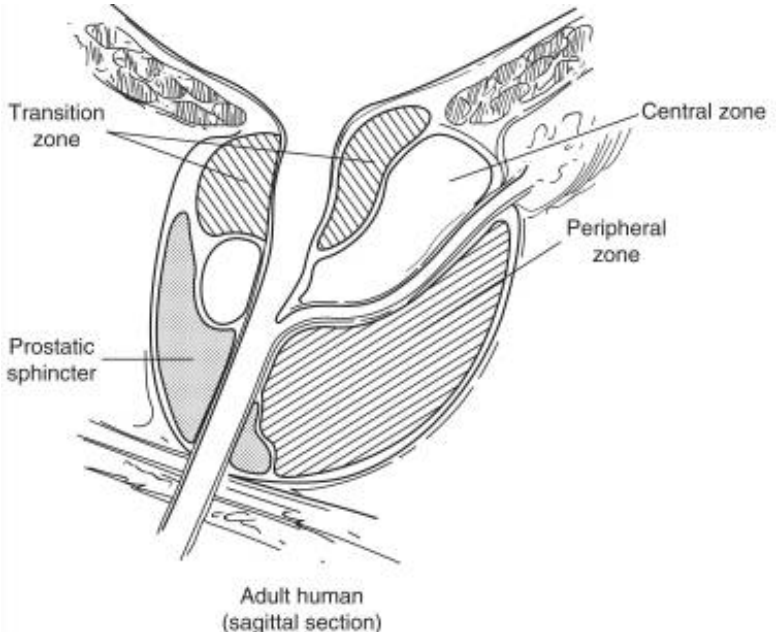


Figure 3 Schematic illustration of human prostate showing the location of the four zones[5]

prostatic hyperplasia (BPH) primarily originates. Compared to peripheral zone cancers, carcinomas that originate in the transition zone have been suggested to be of lower malignant potential. The fourth part of the prostatic urethra that is not enclosed within the peripheral zone is covered by the anterior myoelastic/fibromuscular stroma[4].

1.2.2 Prostate Histology

Prostate histology is closely tied to the structural and functional aspects of the gland, as demonstrated by the unique characteristics of each of the four zones of the prostate. The prostate is composed of branching glands, with ducts that are lined with luminal epithelial cells, basal cells and neuroendocrine cells[16]. Luminal cells are specialized cells that represent the major cell type in the gland, are androgen-dependent for growth and secrete a variety of products like prostatic acid phosphatase (PAP), PSA and human kallikrein-2 into the lumen, which contribute to the formation of the seminal fluid[17]. Basal cells are adjacent to the basement membrane and have oval nuclei. These low, cubic epithelial cells have low expression of androgen receptor (AR) and are believed to contain the stem cell population for the epithelial prostate cells. The presence of these cells can generally be identified by careful inspection of H&E (hematoxylin and eosin) sections, but easier identification is possible with immunohistochemistry for nuclear p63 (nuclear) and high-molecular cytokeratins (cytoplasmic)[5]. The third epithelial cell type found in the prostate is neuroendocrine, which cannot be reliably identified in H&E sections but can be detected by immunohistochemistry by neuroendocrine markers such as chromogranin and synaptophysin immunostains[5]. They do

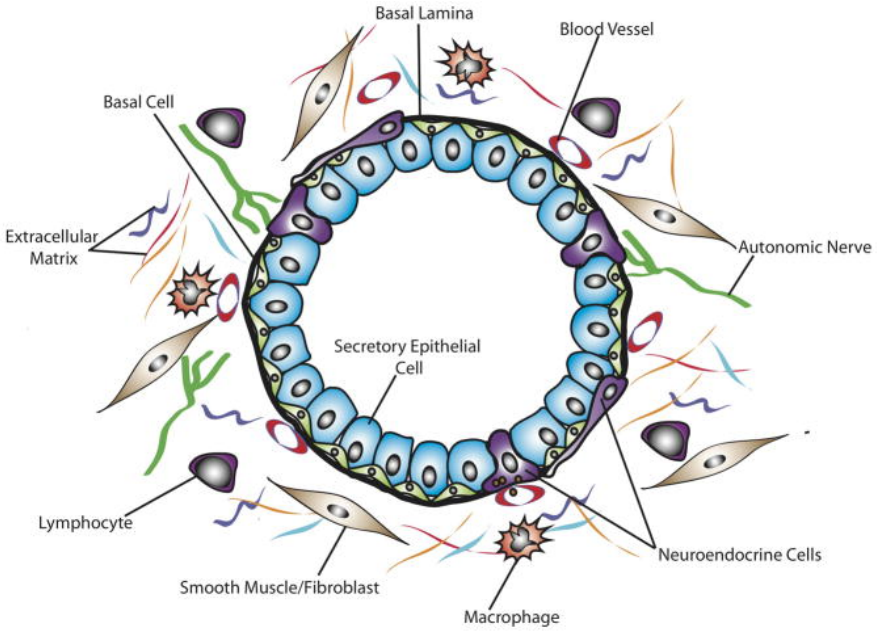


Figure 4 Normal histology of the prostate[7]

not express AR and their role is not fully understood. Surrounding the gland is a stroma that includes abundant smooth muscle cells mixed with fibroblasts, endothelial cells, dendritic cells, nerve cells, and inflammatory infiltrates. No adipose tissue is present in the prostate.

1.3 Prostate Pathology

1.3.1 Benign Prostatic Hyperplasia

BPH leads to an increase in the number of epithelial cells in the prostate glands. This proliferation often results in the formation of nodules or glandular hyperplasia within the gland, which is accompanied by changes in the surrounding stroma. The stromal cells undergo hypertrophy and hyperplasia, contributing to the overall enlargement of the prostate gland, and that causes urinary symptoms such as difficulty starting or stopping the flow of urine, a weak urine stream, or the need to urinate more frequently, especially at night[18]. The exact cause of BPH is not well understood, but it is thought to be related to changes in hormone levels that occur as men age. The prevalence of BPH increases with age, and it is estimated that approximately 50% of men in their 60s and up to 90% of men in their 70s and 80s have some degree of BPH[19]. BPH is not a cancerous condition, and it does not increase the risk of prostate cancer[20]. Treatment for BPH includes conservative management with lifestyle advice initially where appropriate, before progressing to medical treatment and/or surgery if required. The medications include alpha blockers, which relax the muscles in the prostate and bladder neck to improve urine flow, and 5-alpha reductase inhibitors, which shrink the prostate by blocking the production of a hormone that can cause the prostate to enlarge[21]. Surgery is another option

for treating BPH. One common surgical procedure is transurethral resection of the prostate (TURP), which involves removing excess prostate tissue through the urethra. Other surgical options include open prostatectomy, in which an incision is made in the lower abdomen to remove the excess prostate tissue, and prostate artery embolization, in which the blood supply to the prostate is blocked to shrink the gland[22]. The choice of treatment for BPH will depend on the severity of the symptoms and the size of the prostate.

1.3.2 High-grade Prostatic Intraepithelial Neoplasia

High-grade prostatic intraepithelial neoplasia (HGPIN) is a type of precancerous lesion that is found in the prostate gland. It is characterized by abnormal growth and changes in the cells of the prostate gland, specifically in the prostate epithelium[23,24]. HGPIN is characterized by increased proliferation of luminal epithelial cells, with the basal cell layer being partially or completely preserved, but without invasion into the surrounding stroma[25]. Similar to prostate cancer, HGPIN is most commonly identified in the peripheral zone and can be multifocal. HGPIN is considered a precancerous condition because it can progress to prostate cancer in some cases[26,27]. The prevalence of HGPIN increases with age and it does not cause an elevation of serum PSA levels[28]. The diagnosis of HGPIN in prostate biopsies is clinically significant as it increases the likelihood of detecting prostate carcinoma in subsequent rebiopsies. Patients with HGPIN in three or more biopsy sites, despite an initially negative prostate biopsy, are recommended to have a repeat biopsy[29]. The extent of HGPIN is a strong predictor of cancer in subsequent biopsies[30]. Therefore, it is crucial to recognize HGPIN as it is closely associated with prostate carcinoma.

1.3.3 Atypical Small Acinar Proliferation

Atypical small acinar proliferation (ASAP) is a condition that affects the prostate gland and is characterized by the presence of small, atypical glands in the prostate tissue[31]. The significance of ASAP in pathological reports is that it indicates the presence of abnormal or atypical glands in the prostate tissue, which can be a precursor to prostate cancer. Approximately 5% of prostate needle biopsies reveal the presence of ASAP, and previous studies have shown that 30-50% of patients initially diagnosed with ASAP progress to prostate cancer on subsequent biopsies[32,33]. Therefore, the detection of ASAP on a prostate biopsy is typically concerning and requires further monitoring and follow-up.

1.3.4 Prostate Cancer

Prostate acinar adenocarcinoma is the most common form of prostate cancer, accounting for over 90% of all cases. It typically originates in the peripheral zone of the prostate, although it can also occur in the central and transitional zones, albeit less frequently. Histologically, acinar adenocarcinoma is characterized by the presence of enlarged nuclei with prominent nucleoli, luminal mucin, crystalloids, and amorphous eosinophilic secretions. The absence of basal cells is also a key feature of acinar adenocarcinoma. Additional characteristics may include glomerulation, mucinous fibroplasia, and perineural invasion[34].

The remaining 5-10% of prostate cancers are non-acinar types, which include various subtypes such as ductal adenocarcinoma, sarcomatoid carcinoma, squamous cell carcinoma, adenosquamous carcinoma, transitional cell carcinoma, small cell carcinoma, basal cell carcinoma, and clear cell carcinoma. Each subtype exhibits distinct histological and clinical features, and varies in incidence and prognosis[34].

1.4 Prostate Cancer Risk Factors

Like all types of cancer, the exact cause of prostate cancer is not easy to determine. In many cases, multiple factors may be involved, including genetics and exposure to environmental toxins, like certain chemicals or radiation. One of the primary risk factors may be obesity, as obese men have been found to have a 34% higher death rate from prostate cancer than normal-weight men[35].

There are, however, several non-modifiable risk factors have been identified that are thought to increase the probability of developing prostate cancer.

1.4.1 Age

Numerous studies have explored the correlation between age and the risk of developing prostate cancer. Prostate cancer is uncommon in men younger than 45, but becomes more common with advancing age. Cancer statistics indicate that the risk of prostate cancer increases exponentially with age, with the most substantial increase in risk occurring after the age of 50. The risk of prostate cancer is 5 times higher in men over the age of 60 compared to men under the age of 60[36]. According to the American Cancer Society, the incidence rate of prostate cancer among men under the age of 49 in the United States is 1 in 451. However, this rate increases to 1 in 55 among men between the ages of 50-59 and further elevates to 1 in 20 among

men aged between 60-69. It is noteworthy that a significant proportion of prostate cancer cases are diagnosed in men above 65 years of age[36].

1.4.2 Ethnicity

Cancer occurrence and outcomes vary considerably between racial and ethnic groups, largely because of inequalities in wealth that lead to differences in risk factor exposures and barriers to high-quality cancer prevention, early detection, and treatment[37]. According to the American Cancer Society, African American men are more than twice as likely to develop prostate cancer and 2.5 times more likely to die from it compared to white men. In contrast, Asian and Hispanic men have a lower incidence rate of prostate cancer[36]. Similar trends are also observed in UK[38].

1.4.3 Genetic Factors

Genetics may affect risk, as suggested by associations with race, family, and specific gene variants. Having a father or brother with prostate cancer more than doubles a man's risk of developing the disease. The risk is even higher for men who have multiple family members with prostate cancer, especially if their relatives were diagnosed at a young age. Additionally, men of African descent are more likely to have a family history of prostate cancer, which may partially explain the higher incidence rate of prostate cancer among African American men[39]. Twin studies in Scandinavia suggest that 58% of prostate cancer risk can be explained by inherited factors[40].

As of 2018, genome-wide association studies (GWAS) has identified 63 susceptibility loci for prostate cancer, which are regions of the genome that are associated with an increased risk of developing prostate cancer[41]. Although the majority of these variations have low prevalence, when combined, they account for approximately 30% of hereditary prostate cancer risk[42]. One of the most well-established susceptibility loci for prostate cancer is located on chromosome 8q24[43,44]. This region contains several genes, including POU5F1P1, originally thought to be a pseudogene, that can encode a functional protein that contributes to carcinogenesis through its role as a weak transcriptional activator[45]. Another gene in proximity to this region is MYC, a proto-oncogene frequently disrupted in various cancers[46,47]. In another study, interactions between the 8q24 locus and genes involved in the Wnt signaling pathway have also been observed, suggesting that certain risk single nucleotide polymorphisms (SNPs) may affect the expression of multiple genes and influence several cell signaling pathways[48]. Other susceptibility loci identified by GWAS include those located on chromosomes 10q11.23,

17q12, and 19q13.33, among others. These variants have been associated with an increased risk of prostate cancer, and the genes located in these regions, such as MSMB, HNF1B, and KLK3, are known biomarkers for prostate cancer progression[49-52]. In addition, the BRCA1 and BRCA2 genes, which are well recognized for their role in increased breast cancer risk, are also associated with an increased incidence of prostate cancer. It is estimated that 2% of men diagnosed with early-onset prostate cancer have a BRCA2 mutation, compared to 1 in 300 in the general population[53,54]. Germline BRCA2 mutations are associated with an 8.6-fold increased risk of developing prostate cancer at age 65 or earlier, while germline BRCA1 mutations are associated with a more modest 3.75-fold increased risk[55,56]. Men with a BRCA2 mutation and an elevated PSA were found to have 48% prostate cancer at subsequent biopsy and were more likely to have intermediate- or high-risk disease, suggesting that targeted PSA screening is likely to be beneficial in this population[57].

Biomarkers are defined characteristics that are measured to indicate normal biological processes, pathologic processes, or responses to exposure or treatment[58]. In the past decades, PSA was the only widely-used serum biomarker for prostate cancer. However, this has led to an increase in over-diagnosis and over-treatment of prostate cancer[59]. Consequently, there has been a push toward identifying new biomarkers for prostate cancer that can be used in combination with PSA testing. However, in 2005, the TMPRSS2-ERG (The transmembrane protease serine 2 : v-ets erythroblastosis virus E26 oncogene homolog) gene fusion has been identified as a promising biomarker for prostate cancer, as it has been found to be present in 55% of prostate cancer cases, as determined by cancer outlier profile analysis[60]. This gene fusion is the result of a chromosomal rearrangement that brings together the androgen-regulated TMPRSS2 gene promoter and ERG genes, leading to the formation of a new oncogenic protein that promotes prostate cancer development and progression[61]. The TMPRSS2 gene encodes a protein belonging to the serine protease family, which functions in prostate carcinogenesis and relies on gene fusion with ETS transcription factors, such as ERG[62]. ERG, on the other hand, is an oncogene that encodes a member of the erythroblast transformation-specific family of transcription factors, which is a key regulator of cell proliferation, differentiation, angiogenesis, inflammation and apoptosis[63]. The TMPRSS2:ERG gene fusion is the most frequent genomic alteration found in prostate cancer cases and results in the overexpression of the ERG transcription factor, which has been observed in both early- and late-stage prostate cancer[60,64]. This genetic alteration leads to the

activation of ERG gene, which in turn drives the growth and proliferation of prostate cancer cells, resulting in more aggressive and invasive tumors.

Numerous studies have evaluated the significance of TMPRSS2-ERG in prostate cancer patients with varying results and they demonstrated that TMPRSS2-ERG fusion was associated with an increased risk of prostate cancer mortality[65-67]. In this study, we analyzed the expression pattern of the TMPRSS2-ERG fusion gene in human pan-cancers, including prostate cancer, by using the publically available data from cBioPortal. Among prostate cancer

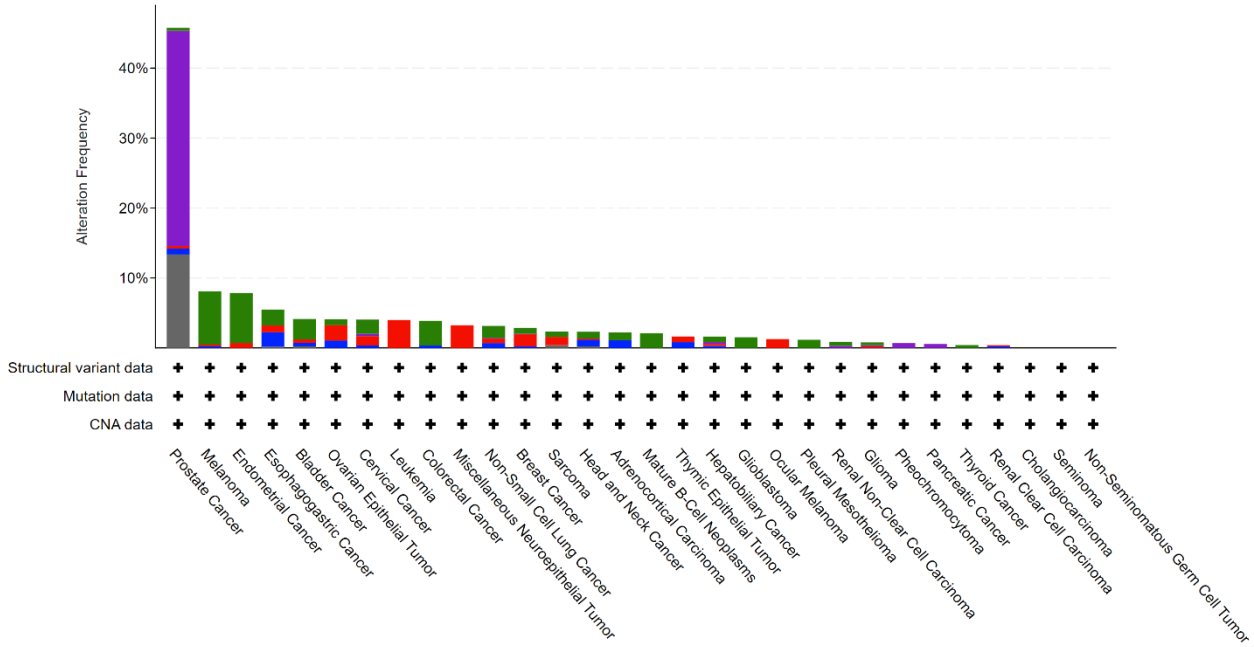


Figure 5 Alteration frequency analysis of TMPRSS2 and ERG gene in human cancers using cBioPortal. The alteration frequencies of TMPRSS2 and ERG across 32 cancer studies (TCGA PanCancer Atlas Studies). Structural variant (gene fusion) is the highest for both gene in prostate cancer patients. The red bars indicate gene amplification, blue bars are deep deletions, green bars are mutations, gray bars are multiple alterations and purple bars indicate structural variant.

studies, the most frequent alteration of TMPRSS2 and ERG was gene fusion **(Figure 5)**. Amplification, missense mutation and deep deletion were less frequently observed. Studies have shown that TMPRSS2 and ERG were altered in over 45% in prostate cancers, with approximately 30% of these alterations being gene fusions. It should also be emphasized that the activation of the AR serves as a trigger for TMPRSS2:ERG rearrangement. Based on this evidence, TMPRSS2: ERG fusions are observed more frequently in younger patients who exhibit elevated AR levels[68]. Conversely, the expression of TMPRSS2:ERG contributes to the activation of critical oncogenic pathways, including EZH2 and MYC[62,69]. The intricate interplay between TMPRSS2:ERG fusion, AR activation, and downstream pathway activation underscores the complexity associated with the development and progression of prostate cancer.

1.5 Management principles of prostate cancer

In order to determine the presence of prostate cancer, a histological examination of prostate tissue is typically conducted, most commonly using transrectal ultrasound (TRUS) guidance. The decision to perform a biopsy should be based on the combination of the results of the digital rectal examination (DRE), PSA test, and other relevant factors such as the patient's age, overall health, and family history of prostate cancer. However, it is generally not recommended to perform a biopsy in patients with a high clinical suspicion of prostate cancer and evidence of bone metastases, as biopsy is typically used to confirm the diagnosis and the presence of bone metastases likely indicates the cancer has already spread beyond the prostate. The Royal College of Pathologists has established guidelines for the reporting of TRUS biopsies of the prostate in order to standardize practice and provide consistent information to clinicians[70]. These guidelines include recommendations for the reporting of the number and location of cores obtained, the histological type and grading of the tumor using both the Gleason grading system and grade group, an estimate of tumor extent, the presence of perineural invasion, and evidence of extra-prostatic extension[70].

While TRUS biopsy is considered the gold standard for diagnosing prostate cancer, it carries risks of complications such as bleeding, infection and pain, and it is also prone to sampling error and has a high rate of detecting clinically insignificant prostate cancer, which may not require treatment[71]. Conversely, magnetic resonance imaging (MRI), a non-invasive imaging technique utilizing magnetic fields and radio waves, has gained popularity in detecting prostate cancer due to its ability to provide detailed images of suspicious areas, particularly high-risk cancers, and its ability to detect areas not visible on TRUS biopsy. Furthermore, multiparametric MRI (mpMRI) is more specific and less sensitive than TRUS biopsy in detecting clinically significant prostate cancer[72]. The Prostate Imaging-Reporting and Data System version 2 (PI-RADS v2) is used to score findings on mpMRI and defines standards of clinical service for mpMRI, including image creation and reporting[73]. It is estimated that performing mpMRI before prostate biopsy could reduce the number of TRUS biopsies by up to 25%[74,75].

It is important to note that while mpMRI has many advantages, inter-observer variability remains a major limitation, which PI-RADS v2 aims to resolve. A meta-analysis has shown that this scoring system is highly sensitive and specific for prostate cancer detection despite moderate inter-observer variability[76,77]. PI-RADS v2 scoring has shown a specificity

and sensitivity of 73% and 95%, respectively, for detection of prostate cancer[78]. Therefore, the National Institute for Clinical Excellence (NICE) is currently considering the use of mpMRI before TRUS biopsy as a standard of care[79]. However, in order to enhance the early detection

Biomarker/Test	Components	Sample Type	Clinical Use	FDA approved
PSA	Kallikrein-3 enzyme	Serum	Initial or Repeat Biopsy	Yes
PHI (Prostate Health Index)	PSA, fPSA, 2ProPSA	Serum	Initial or Repeat Biopsy	Yes
4K Score	PSA, fPSA, iPSA, kallikrein-like peptidase 2 (kK2)	Serum	Initial or Repeat Biopsy	No
PCA3	Prostate Cancer Antigen-3 ncRNA	Urinary	Repeat Biopsy	Yes
MPS (Mi-Prostate Score)	PSA + PCA3a + TMPRSS2-ERG	Urinary + Serum	Initial or Repeat Biopsy	No
SelectMDx Score	mRNA expression levels of DLX1 and HOXC6	Urinary	Initial Biopsy	No
ExoDx Prostate Intelliscore	Exosomal mRNA of PCA3 and ERG + SPDEF	Urinary	Initial or Repeat Biopsy	No
Confirm MDx	Methylation of GSTP1, APC, and RASSF1 genes	Tissue based	Repeat Biopsy	No
Decipher	Microarray-based expression of 22 Genes	Tissue based	Repeat Biopsy	No
Oncotype Dx	Gene expression analysis of 17 genes (12 cancer specific and 5 housekeeping genes)	Tissue based	Repeat Biopsy	No
Polaris	Gene expression analysis of 46 genes (31 cancer specific and 15 housekeeping genes)	Tissue based	Repeat Biopsy	No

Figure 6 FDA approved and experimental PCa biomarkers and tests

and risk assessment of clinically significant prostate cancer (csPCa), other molecular biomarkers can also be used alongside PSA and mpMRI to address cases with indeterminate findings. Serum biomarkers such as the 4K score[80] and prostate health index (phi)[81], as well as urine biomarkers like SelectMDx[82], ExoDx Prostate Intelliscore[83], Mi-ProstateScore (MPS)[84], and PCa antigen 3 (PCA3)[85], can assist in making decisions about initial or repeat biopsies. Additionally, tissue-based biomarkers such as Confirm MDx[86], Decipher[87], Oncotype Dx[88], and Polaris[89] can be utilized to assess aggressiveness and predict outcomes, particularly in men with low- or intermediate-risk profiles. These biomarkers are based on gene panels associated with cancer. The availability of multiple biomarkers has led to the development of personalized treatment approaches for men with prostate cancer, allowing urologists to make informed decisions regarding appropriate treatment and predict the risk of recurrence and progression after localized treatment.

Cambridge Prognostic Group	Criteria
1	Gleason score 6 (grade group 1) **and** PSA less than 10 µg/l **and** Stages T1–T2
2	Gleason score 3 + 4 = 7 (grade group 2) **or** PSA 10 µg/l to 20 µg/l **and** Stages T1–T2
3	Gleason score 3 + 4 = 7 (grade group 2) **and** PSA 10 µg/l to 20 µg/l **and** Stages T1–T2 **or** Gleason 4 + 3 = 7 (grade group 3) and Stages T1–T2
4	One of : Gleason score 8 (grade group 4) **and** PSA more than 20 µg/l **and** Stage T3
5	Two or more of : Gleason score 8 (grade group 4) **and** PSA more than 20 µg/l **and** Stage T3 **or** Gleason score 9 to 10 (grade group 5) **or** Stage T4

Figure 7 Risk stratification for people with localised or locally advanced prostate cancer. NICE recommends using the serum PSA, Gleason score, and clinical stage at diagnosis to assign a risk category to men with localized or locally advanced prostate cancer. This can help clinicians decide on the most appropriate treatment for the patient.

In addition to these diagnostic steps, the cancer tissue that is obtained during the biopsy is also evaluated by a pathologist using the Gleason Grading system, which gives a score of 2-10, indicating the aggressiveness of the cancer[90]. The score is based on how closely the cancer cells resemble normal cells and how abnormal they appear under a microscope. This information is used to classify the cancer into a grade group (1-5) which helps to determine the prognosis and the treatment plan. Once the diagnosis of prostate cancer is confirmed, the stage of the cancer is determined. Determining the stage of prostate cancer is crucial in understanding the prognosis and determining the most appropriate treatment options. The TNM system, which stands for Tumor, Node, and Metastasis, is widely used and categorizes the cancer based on the size of the tumor, the involvement of lymph nodes and the presence of metastases[79]. The TNM system is divided into 4 stages, with T1 and T2 indicating that the cancer is confined to the prostate, while T3 and T4 indicating that the cancer has spread beyond the prostate. Understanding the stage of the cancer is essential in making informed decisions about the patient's treatment plan.

1.6 Management of prostate cancer

As I mentioned above, management for prostate cancer depend largely on the stage of the disease, the Gleason score, and the PSA level, expression of relevant biomarkers, as well as the patient's overall health and preferences. Patients with localized or early-stage cancer are typically offered curative treatments such as surgery or radiation therapy, while patients with advanced or metastatic cancer are typically offered palliative treatments such as hormone therapy or chemotherapy.

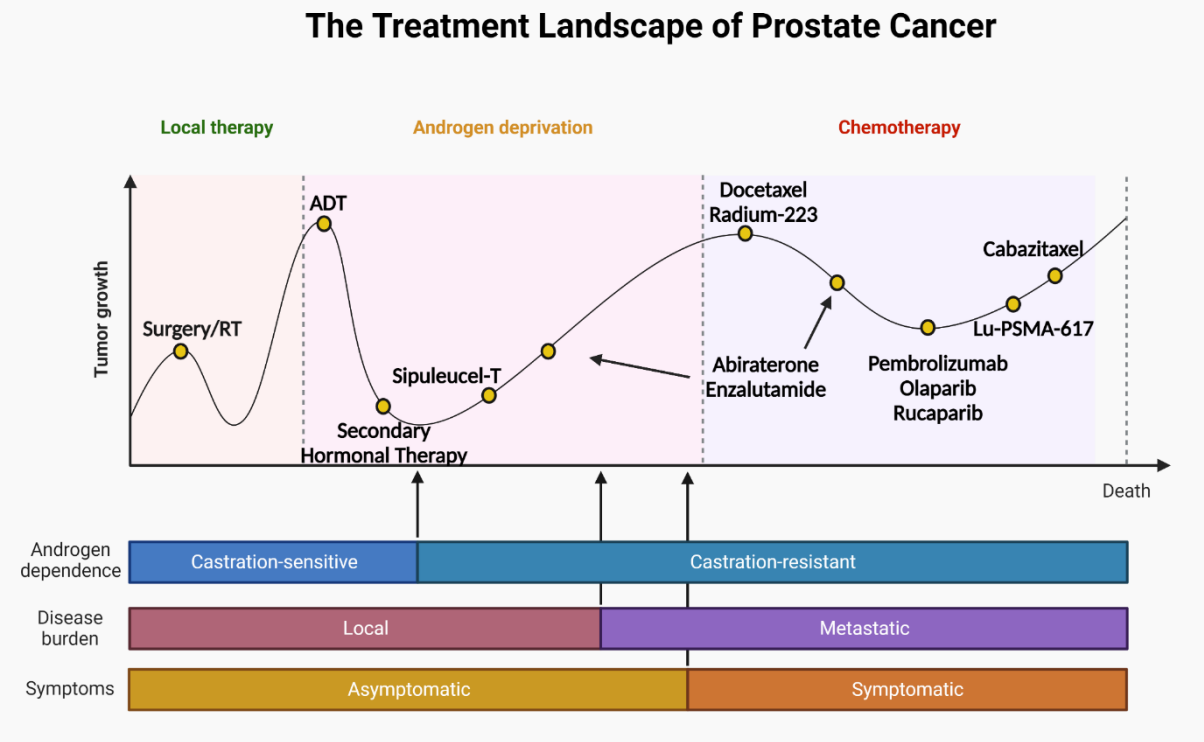


Figure 8 The prostate cancer progression and the corresponding treatments. Prostate cancer can be treated with surgery or radiotherapy for localized tumors. For advanced prostate cancer, androgen deprivation therapy (ADT) is the primary treatment. Castration-resistant prostate cancer can be initially treated with antiandrogens such as enzalutamide, abiraterone, or immunotherapy agents such as Sipuleucel-T. If the disease progresses further, docetaxel and cabazitaxel can be added to the treatment, along with abiraterone and enzalutamide if they haven't been used previously. Other targeted therapies such as olaparib and rucaparib may also be used. *This figure was created using BioRender.com*

1.6.1 Active Surveillance

Active surveillance also referred to as "watchful waiting" is a treatment strategy for men diagnosed with low-risk prostate cancer. This approach involves regular monitoring through PSA blood tests and prostate biopsies. The goal of active surveillance is to closely observe the progression of cancer and delay or avoid the use of more aggressive treatments that may have significant side effects. It is typically recommended for patients with early-stage prostate cancer that is asymptomatic and unlikely to progress rapidly. During active surveillance, patients undergo regular follow-up examinations to assess any changes in their

cancer status. Approximately one-third of men who choose active surveillance may eventually show signs of tumor progression and may need to initiate treatment within three years[91]. A study conducted in 2011 suggested that active surveillance is the best option for older 'low-risk' patients[92].

1.6.2 Surgery

Surgical removal of the prostate, known as prostatectomy, is a common treatment option for early-stage prostate cancer or for cancer that has not responded to radiation therapy. There are different techniques used for prostatectomy, including radical prostatectomy, robotic prostatectomy, and laparoscopic prostatectomy. Radical prostatectomy is the removal of the entire prostate gland, seminal vesicles and sometimes the lymph nodes. Robotic prostatectomy and laparoscopic prostatectomy are minimally invasive procedures. Robotic prostatectomy utilizes a robotic system to assist in the surgical removal of the prostate gland, while laparoscopic prostatectomy involves small incisions and the use of a camera to guide the removal of the prostate gland. The decision to undergo a prostatectomy is based on several factors, including the stage and grade of the cancer, as well as the patient's overall health and preferences. In general, prostatectomy is recommended for men with early-stage prostate cancer that has not spread beyond the prostate gland[93]. It's important to note that prostatectomy may carry certain risks and potential side effects. Nerve damage during the procedure can lead to changes in quality of life for prostate cancer survivors. Additionally, common side effects of prostatectomy include urinary incontinence and erectile dysfunction, which can be temporary or permanent[94].

1.6.3 Radiation therapy

Radiation therapy, also known as radiotherapy, is a commonly used treatment for prostate cancer across all stages. It can be employed as an alternative to surgery or as a follow-up treatment if surgery was not successful in curing the cancer. The aim of radiation therapy is to destroy cancer cells using high-energy beams such as X-rays or protons. There are two main types of radiation therapy for prostate cancer : brachytherapy and external beam radiation therapy (EBRT). Brachytherapy, also known as internal radiation therapy, involves the insertion of small radioactive seeds directly into the prostate gland. These seeds emit radiation to destroy cancer cells within the prostate. Brachytherapy is recommended for patients with low to intermediate-risk localized disease [95]. There are two main types of brachytherapy : low dose rate (LDR) and high dose rate (HDR). LDR brachytherapy involves the permanent

implantation of small radioactive seeds, typically made of iodine-125 or palladium-103, into the prostate, while HDR is a temporary form of brachytherapy, where a small radioactive source, typically made of iridium-192, is injected into the prostate for a short time, typically a few minutes. In LDR, the seeds remain in the prostate permanently and emit low levels of radiation over time. One study suggested that LDR brachytherapy may even be effective in controlling high-risk localized prostate cancer and may benefit patients who cannot tolerate radical prostatectomy or EBRT[96].

EBRT can be offered to all patients with localized and locally advanced prostate cancer. During EBRT, high-energy X-rays are delivered to the prostate gland from a machine called a linear accelerator. The X-rays kill the cancer cells while minimizing damage to surrounding healthy tissue. EBRT is typically administered on an outpatient basis, five days a week for several weeks, with each treatment sessions usually last around 15-30 minutes. Although radical radiotherapy is an effective treatment option for prostate cancer, it is associated with a number of possible side effects and complications. These may include urinary and bowel problems, sexual dysfunction, skin irritation and radiation-induced bowel cancer[94].

1.6.4 High-intensity focused ultrasound and cryotherapy

High-intensity focused ultrasound (HIFU) is a type of non-invasive, image-guided therapy used to treat prostate cancer. It utilizes high-energy ultrasound waves to selectively target and destroy cancerous tissue in the prostate gland. HIFU is performed by inserting a probe into the rectum and directing the ultrasound waves towards the prostate. High-energy ultrasound waves are then directed at the prostate gland, raising the temperature to over 100 degrees Celsius and destroying the cancerous tissue[97]. The procedure is guided by real-time imaging, such as ultrasound or MRI, to precisely target the cancerous areas while sparing surrounding healthy tissue. Whilst HIFU have shown promising results in the management of localised prostate cancer[98,99], it should not be seen as a replacement for traditional treatments such as surgery or radiation therapy but rather as a primary option for well-differentiated, monofocal prostate cancer. By using HIFU first, patients may be able to postpone the need for more invasive treatments that carry a higher risk of genitourinary side effects, allowing them to live without such side effects for a longer period of time[100]. One of the advantages of HIFU is its non-invasive nature, which means it does not involve surgical incisions or radiation exposure. This results in minimal blood loss, reduced pain, and faster recovery times compared to traditional treatments. Additionally, HIFU may offer a lower risk of certain side effects such

as urinary incontinence and erectile dysfunction, although further research is needed to fully understand its long-term outcomes and side effect profile.

1.6.5 Hormonal Therapy

Hormonal therapy, also known as ADT, is a treatment option for prostate cancer that aims to reduce the levels of androgens in the body. Surgical castration (orchiectomy) or medical castration is used in hormonal therapy to block prostate cancer cells from getting dihydrotestosterone (DHT), a hormone produced in the prostate and required for the growth and spread of most prostate cancer cells[101]. However, hormonal therapy rarely cures prostate cancer because cancers that initially respond to hormonal therapy typically become resistant after one to two years. Hormonal therapy is, therefore, usually used in combination with radiation therapy or surgery, or as a standalone treatment for advanced or metastatic prostate cancer [102].

Orchiectomy involves the surgical removal of the testicles. By removing the testicles, the production of androgens is significantly reduced, which can help slow the progression of the disease. Orchiectomy is considered the gold standard of treatment and have benefits such as cost savings compared to medical castration even if it is an irreversible procedure [103].

Luteinizing hormone-releasing hormone (LHRH) agonists and antagonists were also among the first therapies developed to reduce androgen signaling in PCa[104]. LHRH agonists downregulate the gonadotrophin releasing hormone (GnRH) receptors in the pituitary, producing a hypogonadal effect by reducing the secretion of luteinising hormone (LH) and follicle stimulating hormone (FSH), and subsequently reducing androgen levels[105]. LHRH agonists, such as leuprolide and goserelin, work by mimicking GnRH and binding to GnRH receptors on the pituitary gland, thereby stimulating the release of LH and FSH. This results in an initial surge of testosterone and estrogen, followed by a decrease in their production, as the negative feedback loop is activated. On the other hand, LHRH antagonists, such as degarelix, work by binding to the LHRH receptor and blocking the action of LHRH. This leads to an immediate decrease in LH and testosterone levels, without the initial surge seen with LHRH agonists. Both LHRH agonists and antagonists can be effective in the treatment of prostate cancer, but they may cause side effects, such as hot flashes, decreased libido, erectile dysfunction, and bone loss[106].

Anti-androgens are another type of hormonal therapy that act peripherally to reduce the effect of endogenous androgens by competitively binding to the AR. There are two types of

anti-androgen therapy : non-steroidal anti-androgens, such as bicalutamide and flutamide, and steroidal anti-androgens, such as cyproterone acetate and megestrol acetate. Non-steroidal anti-androgens work by blocking the androgen receptor in prostate cancer cells, while steroidal anti-androgens work by blocking the production of androgens in the testes and adrenal glands. Anti-androgen therapy is often used in combination with LHRH agonists or antagonists to achieve maximum androgen deprivation[107]. This combination therapy is known as total androgen blockade (TAB), aims to maximize androgen deprivation and improve treatment outcomes in prostate cancer patients.

However, despite hormonal therapy, disease progression may occur, leading to the development of castration-resistant prostate cancer (CRPC). CRPC has a poor prognosis, and treatment options for this stage of prostate cancer may include other targeted therapies, immunotherapy, chemotherapy, or radiopharmaceuticals. CRPC typically emerges after 2-3 years of ADT and is characterized with a poor prognosis and a median survival time ranging from 9-30 months[108].

1.6.6 Chemotherapy

Chemotherapy is a systemic treatment option for advanced prostate cancer that has spread beyond the prostate gland and has become resistant to hormone therapy[109]. It involves the use of drugs that are designed to kill or slow down the growth of cancer cells. Docetaxel and cabazitaxel are the two main chemotherapy drugs used for prostate cancer, with both showing improved overall survival rates and quality of life in clinical trials. Docetaxel is usually administered in combination with a corticosteroid, such as prednisone. This combination has become the standard first-line chemotherapy regimen for mCRPC. Studies have also shown that initiating docetaxel treatment along with ADT at an early stage of hormone-sensitive metastatic prostate cancer can lead to improved overall survival and decreased time to disease progression[110]. As per the current recommendation, men presenting with metastatic disease should be offered a combined treatment of ADT and docetaxel as the first-line therapy, provided their physical fitness is appropriate for receiving this treatment[111]. Cabazitaxel, on the other hand, is a semisynthetic tubulin-binding taxane drug that was developed after the emergence of resistance with other taxanes[112]. Docetaxel has been the primary chemotherapy for prostate cancer, while cabazitaxel is used as second-line therapy when resistance is observed with other taxanes.

Chemotherapy is typically administered in cycles, with a combination of drugs given over a specific period, followed by a rest period. The treatment duration and frequency depend on the specific chemotherapy regimen and the individual patient's response. While chemotherapy can be effective in treating advanced prostate cancer, it can also cause side effects, including nausea, fatigue, hair loss, and increased susceptibility to infections. These side effects can vary in severity and can be managed or mitigated with supportive care measures[109].

1.6.7 Immunotherapy

Immunotherapy is a treatment that stimulates the patient's immune system to recognize and attack cancer cells. Immunotherapy has emerged as a promising approach for treating prostate cancer. One type of immunotherapy called Sipuleucel-T (Provenge), a cancer vaccine that stimulates ex vivo antigen presenting cells, is the only FDA-approved treatment for advanced prostate cancer. Sipuleucel-T is an autologous cellular immunotherapy that consists of collecting a patient's peripheral blood mononuclear cells (PBMCs), which include antigen-presenting cells such as dendritic cells, processing them in a laboratory to activate them against a fusion protein composed of PAP and granulocyte-macrophage colony-stimulating factor (GM-CSF), and then re-infusing them back into the patient[113]. It was approved by the FDA in 2010 for the treatment of asymptomatic or minimally symptomatic metastatic castrate-resistant prostate cancer but not licensed for use in Europe.

More recently, the FDA approved another immunotherapy called Pembrolizumab (Keytruda) for patients with metastatic prostate cancer with certain genetic mutations, specifically in patients with microsatellite instability-high (MSI-H) or mismatch repair deficient (dMMR) tumors. MSI-H and dMMR tumors have a high mutational burden and are associated with increased neoantigen expression, which can stimulate an immune response[114]. Pembrolizumab (Keytruda) is a monoclonal antibody that targets the programmed cell death protein 1 (PD-1) receptor on T-cells. PD-1 is a checkpoint receptor that regulates T-cell activation and function. Cancer cells can exploit this checkpoint to evade the immune system by expressing ligands that interact with PD-1 and suppress T-cell activity. By blocking the interaction between PD-1 and its ligands, Pembrolizumab releases the brake on the immune system, enhancing the ability of T-cells to recognize and attack cancer cells. The EMA approved Pembrolizumab in 2021 for the treatment of adult patients with mCRPC who have received chemotherapy in combination with docetaxel. The approval was based on results from

the KEYNOTE-199 clinical trial, which showed that Pembrolizumab improved overall survival in patients with mCRPC[115].

While there is great enthusiasm for immunotherapeutic approaches in prostate cancer, monotherapy with immunotherapeutic may not result in significant clinical benefits. However, combination strategies that leverage the synergistic effects of multiple therapies hold more promise and offer a reason for optimism. For example, a Phase II clinical trial (KEYNOTE-365) evaluated the combination of Pembrolizumab with Enzalutamide in patients with mCRPC and showed promising results, with an overall response rate of 18.5% and a median duration of response of 16.6 months[116]. Another Phase II trial (Checkmate 650) evaluated the combination of Nivolumab and Ipilimumab in mCRPC patients and showed a higher overall response rate (25%) and longer progression-free survival compared to monotherapy with Nivolumab[117].

Similarly, combining immunotherapy with radiation therapy has shown promise in preclinical and clinical studies. Radiation can induce immunogenic cell death and increase the expression of immune checkpoint molecules on cancer cells, making them more susceptible to checkpoint inhibitor therapy[118]. In a Phase I/II trial, patients with metastatic castration-resistant prostate cancer who received a combination of Ipilimumab and radiation therapy had a higher rate of PSA response and longer progression-free survival compared to those who received radiation therapy alone[119].

1.6.8 Targeted Therapy

Targeted therapy is another treatment option for prostate cancer that involves using drugs to specifically target molecules or pathways involved in the growth and spread of cancer cells. For example, drugs that target the AR pathway, such as enzalutamide[120] and abiraterone[121], have been approved for the treatment of mCRPC. This is because ADT, which reduces the levels of androgens in the body, is an effective treatment for prostate cancer, but many patients eventually develop resistance to this therapy and their cancer progresses. Targeted therapy aims to overcome this resistance by targeting the AR signaling pathway using drugs that block the AR or inhibit its activity [122]. Enzalutamide is a nonsteroidal antiandrogen that blocks the binding of androgens to the AR, while abiraterone acetate inhibits the production of androgens in the body.

Other targeted therapies being studied for prostate cancer include poly(ADP-ribose) polymerase (PARP) inhibitors, PI3K inhibitors, and CDK inhibitors. PARP inhibitors target

PARP family of enzymes, which play a critical role in DNA repair, particularly in the repair of single-strand DNA breaks (SSBs). PARP inhibitors work by inhibiting the activity of PARP enzymes, which leads to the accumulation of SSBs in cancer cells. When cancer cells with SSBs attempt to replicate their DNA, the SSBs are converted into double-strand DNA breaks (DSBs), which are more difficult to repair. This results in the accumulation of DNA damage and ultimately leads to cancer cell death[123]. PARP inhibitors are particularly effective in cancer cells with defects in homologous recombination (HR) DNA repair, such as those with BRCA1/2 mutations. In these cells, PARP inhibition leads to synthetic lethality, where the combination of PARP inhibition and HR deficiency results in cell death[123]. The FDA has approved Olaparib and Rucaparib for the treatment of mCRPC in patients with certain genetic mutations[124].

Targeted therapies can often be more effective and less toxic than traditional chemotherapy or radiation therapy by specifically targeting molecules or pathways involved in cancer growth and spread. FDA-approved targeted therapies for prostate cancer include Abiraterone[121] and Enzalutamide[120], which target the androgen receptor pathway, and Radium-223[125], which targets bone metastases. PARP inhibitors such as Olaparib and Rucaparib have also shown promise as targeted therapies for certain patients with advanced prostate cancer[124]. However, it is important to carefully consider the individual patient's specific situation and goals when deciding on a treatment approach for prostate cancer.

1.7 Limitations of Current Therapies and The Need for Personalized Medicine Approaches

Prostate cancer is a complex disease that arises from the accumulation of genetic alterations that contribute to its initiation and progression. Current treatment approaches for prostate cancer, including surgery, radiation therapy, and chemotherapy, are effective in some cases but have limitations, such as treatment resistance and toxicity. Therefore, there is a growing need for personalized medicine approaches that consider the unique genetic and molecular characteristics of each patient's cancer to develop targeted therapies that are more effective and less toxic.

One of the major limitations of current prostate cancer therapies is the development of treatment resistance. Resistance to current therapies has been linked to molecular alterations, such as mutations in the AR gene and activation of alternative signaling pathways. For example, in CRPC, which develops after ADT, the AR signaling pathway remains active

despite low levels of androgens. This is often due to overexpression of AR or mutations that make the receptor insensitive to androgens[126]. Similarly, chemotherapy and radiation therapy can also lead to treatment resistance, making it difficult to achieve long-term remission. Another limitation of current therapies for prostate cancer is the potential for off-target effects. Current therapies may not only target cancer cells but also affect normal, healthy cells, leading to adverse effects. For example, chemotherapy agents such as docetaxel and cabazitaxel have been shown to cause bone marrow suppression, leading to decreased white blood cell count and increased risk of infections[127]. Radiation therapy can also lead to damage of surrounding tissues, causing inflammation and fibrosis, which can result in urinary incontinence and rectal bleeding[128]. Furthermore, current therapies may not be effective in all patients, and predicting response to treatment can be challenging. While the response to treatment can be often similar in the majority of patients, some individuals may not respond to the prescribed treatment. This can lead to delays in finding the right treatment and may result in unnecessary side effects and decreased quality of life. Because we all exhibit genetic divergence, as evidenced by the presence of genetic predispositions in diasease, especially in cancer[129]. In addition, tumors evolve clonally and have heterogeneous intratumoral characteristics at diagnosis. Even the tumor microenvironment of a single tumor can diverge spatially[130]. As emphasized by Mel Greaves, each cancer is unique and has its own specific history[130].Therefore, there is a need to develop targeted therapies that overcome treatment resistance by targeting specific molecular alterations in individual tumors.

In line with this need, genomic testing is one approach to personalized medicine for prostate cancer that involves identifying genetic mutations and alterations that may be driving the cancer. This information can be used to develop targeted therapies that specifically target the altered genes or pathways, thereby increasing treatment efficacy and reducing side effects. For example, PARP inhibitors have shown promise in prostate cancer patients with BRCA mutations[124]. Another approach to personalized medicine is the use of biomarkers to predict response to treatment. For example, PSA levels in the blood can be used to monitor response to treatment and disease progression. Other biomarkers I mentioned in Figure 6 can also be used to monitor disease progression and response to treatment. While current therapies for prostate cancer have limitations, personalized medicine approaches offer new opportunities for improved treatment outcomes by targeting specific molecular alterations in individual tumors. By considering the individual characteristics of each patient, including their genetics, tumor biology, and biomarkers, personalized medicine approaches can improve treatment efficacy

and reduce side effects. Another advantage of personalized medicine is the ability to monitor treatment response and adapt therapies accordingly. By integrating liquid biopsy-based genomics analyses and single-cell/spatial genomics, it becomes possible to track the evolution of a patient's tumor, identify mechanisms of drug resistance, and make informed decisions about treatment adjustments. This dynamic approach allows for real-time assessment and optimization of treatment strategies,leading to the potential improvement of patient outcomes.

Chapter II. Gene Regulatory Network Reconstruction to Guide Personalized Therapy

Chapter II. Gene Regulatory Network Reconstruction to Guide Personalized Therapy

Networks provide a global view that transcends individual entities, shedding light on the very fabric of existence. From the origins of life to the intricacies of communication within living systems, networks play a role in shaping our understanding. At its core, network theory encompasses the basics of interconnectedness. Nodes represent individual entities, while edges signify the connections or relationships between them. This framework applies to diverse systems, whether biological, social, or technological. Networks serve as a fundamental construct for comprehending how information, resources, and influence flow through complex webs of interactions.

Networks can range from simple pairings to vast interconnected systems, providing a means for information flow, exchange, and cooperation in both living and non-living systems. Social networks, for instance, comprise individuals (nodes) and their social connections (edges), forming intricate webs of relationships that facilitate the spread of ideas, influence, and behaviors, shaping collective norms, cultural dynamics, and even political structures. These networks offer insights into phenomena like the spread of epidemics, information cascades, and the formation of online communities. Moreover, social networks serve various functions such as enabling social support systems, resource flow, socio-economic mobility, identity formation, social influence, and the transmission of cultural heritage.

Intriguingly, molecular systems exhibit parallels to social networks in terms of structure, function, and evolution. The human body, with its approximately 37 trillion cells, relies on dynamic and regulated information transfer reminiscent of social networks. These networks are hierarchically organized and depend on dynamic, responsive, flexible, redundant, and resilient networks of information transfer. Just as in human societies, cells within the body communicate and exchange information to coordinate and direct growth, development, survival, and homeostasis. These networks facilitate communication within and between organelles, cells, organs, and other body compartments. They ensure the coordination of various processes and responses, such as differentiation, gene regulation, and the acquisition of specific functions by cells or organelles in different body compartments. These networks comprise nodes representing proteins, genes, metabolites, and other molecular entities, connected by edges that denote physical or functional relationships. These networks regulate biological processes, including cellular signaling, gene regulation, metabolic pathways, and protein-protein interactions, analogous to the information flow and behavior propagation seen

in social networks. The importance of networks extends beyond their descriptive power. Network thinking provides a framework for analyzing complex systems, enabling predictions, and informing decision-making. It enhances our understanding of biological processes, such as protein-protein interactions, gene regulatory networks, and metabolic pathways.

As discussed in the previous chapter, personalized medicine approaches offer promising opportunities for improved treatment outcomes in prostate cancer by targeting specific molecular alterations in individual tumors. One such approach is the reconstruction of gene regulatory networks (GRNs) in individual tumors, which can identify key genes and pathways that are dysregulated in each individual tumor, and design targeted therapies to modulate those pathways. These complex systems of genes and their regulatory interactions govern cell behavior, including tumor growth and metastasis.

GRN reconstruction involves analyzing large datasets of gene expression, epigenetic modifications, and other molecular features of individual tumors, and using computational algorithms to infer the regulatory relationships between genes. This approach has shown effectiveness in predicting drug responses and identifying new therapeutic targets in various cancer types, including prostate cancer[131].

2.1 Network Inception

Recent advances in molecular biology, biotechnology, and bioinformatics have revolutionized our understanding of cellular processes by uncovering a vast array of molecular interactions that contribute to the regulation of metabolic circuits, signaling networks, and molecular machineries. These circuits do not function as isolated complexes but rather contain thousands of different interconnected components working together in a complex regulatory system. These networks represent a map of the interactions between various biological processes and can provide insight into the mechanisms that drive disease. An analysis of the global topological organization of these networks can provide a comprehensive understanding of the functional map of the entire disease state. This can help identify key nodes and pathways that are critical for disease progression, enabling the development of targeted therapies that address these specific targets. Network analysis has thus emerged as a powerful approach for elucidating disease processes and can help pave the way for more effective treatments that improve patient outcomes[131].

The study of networks, or the interconnected systems that underlie many complex phenomena, has a long and fascinating history. The origins of this field can be traced back to

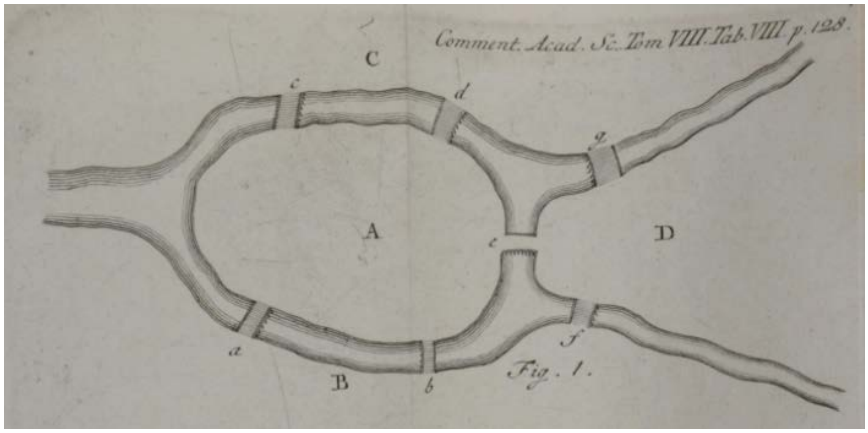


Figure 9 Problem of Seven Bridges of Konigsberg[2]

the 18th century when mathematician Leonhard Euler tackled the problem of the Seven Bridges of Königsberg[2]. This problem involved finding a route through the city of Königsberg that crossed each of its seven bridges exactly once. Euler's solution to this problem, which involved the development of new mathematical concepts related to node-edge relationships and graph theory, laid the foundation for the study of networks as we know it today. As a result of these ideas, topology and graph theories were developed, and their concepts have been applied to numerous disciplines over the past several hundred years.

Biological networks can exhibit a variety of different architectures, but two of the most common types are scale-free and modular networks. Scale-free networks are often used to model biological systems because they exhibit a power-law distribution of node degrees, which is similar to the distribution of many biological systems such as protein-protein interaction networks and gene regulatory networks[132,133]. Modular networks, on the other hand, are also commonly found in biology, where they can represent functional modules or pathways within a larger biological system[134]. For instance, in a protein-protein interaction network, a module might represent a set of proteins that work together to carry out a specific biological function. These two key observations led to a paradigm shift in the modeling of biochemical reactions in cells. Rather than interpreting reactions as straightforward interactions between enzymes and substrates or binding events, biochemical interactions were abstracted into nodes and links, forming a complex network[135]. This innovative approach, known as network or systems biology, has opened up new avenues for investigating intricate regulatory networks in cells and organisms. By leveraging this approach, researchers can delve deeper into the complexities of gene regulatory networks, which are essential for comprehending disease processes and developing personalized therapies.

2.2 Gene Regulatory Network Reconstruction

GRNs are interpretable computational models of the regulation of gene expression in the form of networks, mathematically also defined as graphs. These models incorporate various components involved in gene regulation, including TFs, splicing factors, lncRNAs, miRNAs, and metabolites. The exploration of GRNs has been a longstanding pursuit in the field of biology, exemplified by pioneering research conducted in the 1960s on the lac operon in bacteria[136]. The reconstruction of extensive GRNs has emerged as a central focus in systems biology, employing diverse high-throughput experimental techniques and computational algorithms[137,138]. Unraveling the structure and dynamics of GRNs is crucial for comprehending the mechanisms behind cellular identity establishment and maintenance, offering valuable insights for manipulating cell fate and preventing diseases[139].

There are two main approaches to using graph theory for the analysis of GRNs - the structural approach and the dynamic approach. The structural approach involves analyzing the global organization of the network by computing properties and attributes for individual nodes, links, and/or groups of nodes and links, and comparing them to random reorganized networks. This approach provides an understanding of the overall topology of the network and can identify key regulators/pathways by analyzing the attributes of nodes or edges. An important property of nodes is their degree, which is the number of direct neighbors a node has[140,141]. In real-world networks, most nodes have few neighbors, but a substantial number of nodes have a high degree, known as hubs. The identification of hubs is often of interest as they are topologically and functionally important and have been shown to be preferentially targeted by both bacterial and viral pathogens[142]. The dynamic approach, on the other hand, uses multivariate experiments such as microarray data sets to infer cause-consequence relationships and regulatory links within the context of known pathways and networks. This approach aims to model the behavior of the network over time and predict how changes in the expression of one gene will affect the expression of other genes in the network. A complete understanding of GRNs is often gained through the combination of both approaches, depending on the particular research question.

Constructing a network can be achieved through various methods, one of which is querying different interaction databases to identify the interactions of a list of genes or proteins of interest, such as those that are differentially expressed. Protein-protein databases like STRING[143] can provide an undirected network where the direction of signal flux is not

represented. Alternatively, directed networks can be obtained by using databases or tools like CellNet[144] that contain information on transcription factor-target gene interactions. By reconstructing the network, one can identify hubs and/or bottlenecks, which are nodes interconnecting highly connected nodes or hubs in the system. This approach enables the identification of a larger network for analysis beyond interactions restricted to those that occur between nodes in the gene/protein list. In addition to identifying important components of the network, network analysis can help to identify sub-networks that are enriched in co-regulated genes or identify non-differentially expressed nodes that are topologically important. These approaches can reveal key regulatory mechanisms and pathways that may be missed by traditional differential expression analyses.

Despite the challenges in analyzing complex biological networks, network reconstruction provides valuable information that can aid in the identification of potential drug targets and the development of novel therapies[131]. These networks can also help in the identification of driver mutations in cancer[145,146] or pathways associated with survival of cancer patients, leading to personalized therapies[147]. In addition, reconstructing networks can aid in understanding the mode of action of pharmacological compounds, as well as in the assessment of cells destined for regenerative medicine.

To address the challenge of analyzing large and dynamic biological networks, integrating contextual information such as gene expression data with standard network analysis can provide insight into the most relevant key factors and sub-networks in a particular context. To identify hubs in networks, a number of tools have been developed based on different node parameters, such as degree, Maximum Neighborhood Component, Density of Maximum Neighborhood Component, and other parameters. Some of these tools include Hubba[148], APID2Net[149], PinnacleZ[150], NetworkAnalyzer[151], CentiScaPe[152], and TETRAMER[153]. Our lab has contributed to this field by generating TETRAMER, a tool specifically designed to reconstruct gene-regulatory networks using temporal transcriptome data during cell fate transitions. TETRAMER employs simulations of cascades of temporal transcription-regulatory events to predict key "master" regulators involved in the process.

In addition to cancer, network reconstruction can also aid in the development of therapies for inflammation, degenerative diseases, and infectious diseases caused by emerging pathogens[154]. By employing advanced network analysis techniques to identify crucial nodes and edges, such as those with high centrality and significant pathway involvement, potential

drug targets can be efficiently pinpointed. This process of network reconstruction, in conjunction with a deep comprehension of the underlying mechanisms of diseases, could hold great promise in paving the path toward highly effective and personalized treatments.

2.3 Patient-specific Network Alterations

The study of molecular networks has revolutionized our understanding of biological systems, allowing us to investigate the complexity of cellular processes from a systems-level perspective. Network analysis has been particularly useful in shedding light on disease mechanisms and identifying potential therapeutic targets. However, while early network analyses focused on static representations of molecular interactions, recent advances in experimental techniques and computational algorithms have allowed for the construction of dynamic and patient-specific networks. By considering the variability in molecular interactions and gene expression patterns across different individuals and disease states, these approaches have the potential to provide a more nuanced understanding of disease mechanisms.

Cancer is a complex disease that is driven by a combination of genetic, epigenetic, and environmental factors[155]. Traditional cancer treatment methods, such as chemotherapy and radiation therapy, are often associated with significant side effects and limited effectiveness. The development of personalized treatment strategies that consider the molecular heterogeneity of individual tumors has become an important area of research in cancer biology. One promising approach is the analysis of patient-specific network alterations.

Conceptually, patient-specific network analysis offers several advantages over traditional methods of cancer treatment. Firstly, it allows for the identification of molecular alterations that are specific to individual patients, which can inform the development of personalized treatment strategies. This approach can help to avoid the use of ineffective therapies that may have significant side effects, as well as identify novel therapeutic targets that may not have been identified using traditional approaches. An illustrative instance of precision medicine is the drug trastuzumab (Herceptin), which has proven effective against breast cancer. This medication specifically targets tumors exhibiting human epidermal growth factor receptor 2 (HER2)-overexpression and/or -amplification[156]. By focusing on patients with this specific molecular alteration, trastuzumab demonstrates enhanced efficacy while minimizing unnecessary treatment for individuals lacking the HER2 alteration. Another example involves the use of mercaptopurine (Purinethol) for the treatment of acute lymphoblastic leukemia. Patients with specific genetic variants that hinder the drug's

processing require genetic testing to determine the appropriate therapy[156]. By identifying these variants, healthcare providers can tailor treatment options to ensure patients receive the most suitable and effective therapies for their unique genetic makeup.

Secondly, patient-specific network analysis can provide insights into the molecular mechanisms underlying cancer progression and treatment resistance. By identifying key signaling pathways and regulatory networks that are dysregulated in individual tumors, a deeper understanding of the molecular mechanisms underlying cancer and developing targeted therapies that are tailored to the specific molecular profile of each patient can be gained. This information can be used to develop combination therapies that target multiple pathways simultaneously, leading to more effective treatment outcomes. As an example of developing combination therapies, the study by Jaeger et al. proposed a computational method to simulate pathway cross-talk inhibition (PCI) using combined drugs in breast cancer[157]. This method considered the cross-talk between different pathways by analyzing shared protein interactions. The researchers developed a computational algorithm, known as the PCI index, which took into account all KEGG pathways containing any of the primary targets of FDA-approved drugs in breast cancer[157]. Furthermore, another research team developed PanDrugs, a computational framework that studies gene-drug interactions integrating genomic profiles, biological pathways and pharmacological evidence to address the needs of individual patients based on their specific genetic and molecular characteristics[158].

Thirdly, patient-specific network analysis can be used to predict treatment outcomes and monitor disease progression. By analyzing changes in the molecular network over time, biomarkers that are predictive of treatment response or disease recurrence can be identified. This information can be used to guide treatment decisions and monitor the effectiveness of therapy. These approaches offer a promising avenue for the development of personalized treatment strategies for cancer patients. However, it is important to acknowledge that cancer is a heterogeneous disease, and the dysregulated molecular networks within individual tumors can be remarkably complex and dynamic. This complexity may pose a challenge in identifying the crucial regulatory networks and signaling pathways that are pertinent for the design of targeted therapies.Nevertheless, despite these challenges, the benefits of personalized medicine in cancer treatment are substantial and hold the potential to revolutionize the field in the future[159]. By tailoring treatments to the specific molecular characteristics of each patient's tumor, we can enhance treatment efficacy and minimize adverse effects, ultimately improving patient outcomes. The ongoing advancements in understanding the intricate molecular

networks involved in cancer development and progression provide optimism for continued progress in this field.

Chapter III. Unveiling the Molecular Complexity of Prostate Cancer: Exploring the Genetic and Epigenetic Factors and the Role of Noncoding RNAs in Gene Regulatory Networks

Chapter III. Unveiling the Molecular Complexity of Prostate Cancer: Exploring the Genetic and Epigenetic Factors and the Role of Noncoding RNAs in Gene Regulatory Networks

3.1 Genetic and epigenetic alterations

Prostate cancer is a complicated condition that results from a buildup of genetic and epigenetic changes in prostate epithelial cells. These changes can disrupt critical signaling pathways that manage cell growth, differentiation, and survival, leading to unchecked cell proliferation and the development of tumors. Various molecular signaling pathways provide either positive or negative regulatory signals that control cell growth in a way that aims to maintain cell number and balance. However, this process is fundamentally altered in cancer cells [160,161].

In order for normal cells to transform into cancer cells, they must exhibit at least eight distinct attributes. These attributes include: 1) genetic instability and mutation; 2) self-sustaining growth; 3) evading growth suppressor; 4) resistance to apoptosis and other forms of induced cell death; 5) enabling replicative immortality; 6) ability to form new blood vessels (angiogenesis); 7) localized invasive behavior that differentiates benign and malignant neoplasms; and 8) evasion of the immune system.

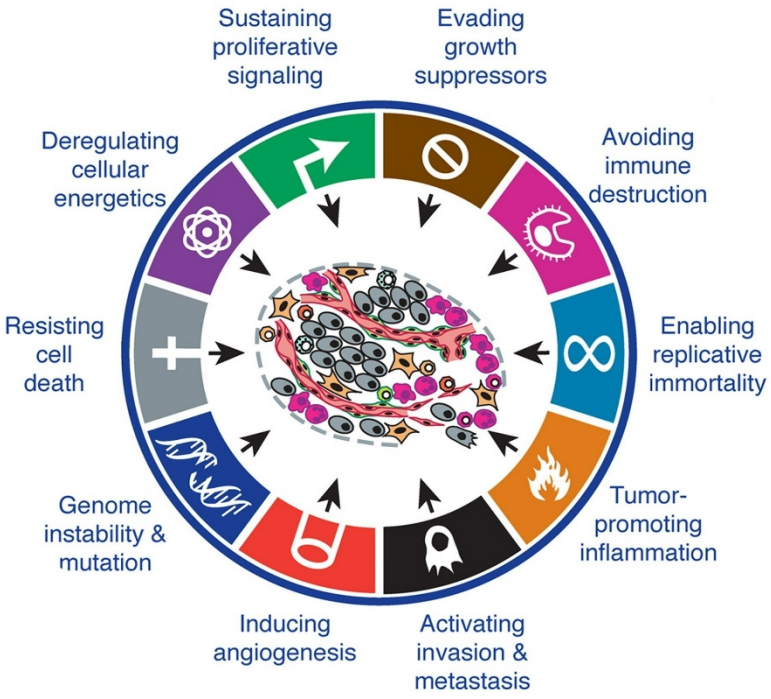


Figure 10 Hallmarks of cancer[6]

Furthermore, cancer cells require a source of energy to maintain their self-sustaining growth and unlimited replication. Tumor-associated inflammatory mediators also promote the progression of pre-cancerous cells to invasive cancer cells, and ultimately, cancer cells acquire the ability to metastasize or spread to other organs or tissues [6,162].

3.1.1 Unraveling the Significance of Tumor Suppressor Gene Alterations

Prostate cancer has shown promise in being associated with biomarkers that include the alteration of tumor suppressor genes and oncogenes. Tumor suppressor genes comprise a diverse group that shares a common characteristic: they all are responsible for negatively regulating cell growth and playing a vital role in the normal cell cycle, DNA repair, and cell signaling. For the development and survival of cancer cells, both copies of a tumor-suppressor gene must undergo inactivation. The absence or mutation-induced inactivation of these genes can lead to the onset and progression of cancer[163]. There are various pathways that can lead to cancer, including homozygous gene deletion, loss of one allele combined with mutational inactivation of the second, mutations in both alleles, or loss of one allele combined with epigenetic inactivation[160]. TP53, PTEN, and RB1 are among the most commonly mutated tumor suppressor genes in prostate cancer.

Retinoblastoma protein (RB1) is a key regulator of the G1/S transition of the cell cycle, and its inactivation can cause increased cell cycle progression and tumor growth. Rb mutations frequently occur in both local and advanced prostate cancer, with almost 50% of them observed in mCRPCs[164,165]. Therefore, the detection of RB1 mutations can be useful in predicting the risk of metastasis and disease progression in prostate cancer. Interestingly, Rb loss does not impact PCa growth; however, upon castration, Rb-null cells display increased proliferation, possibly due to elevated AR expression[166]. In addition, Rb gene alterations have been shown to cause cytoskeleton reorganization, epithelial-to-mesenchymal transition (EMT), and migration in various in vitro and in vivo prostate cancer models, promoting tumor cell spread induced by PTEN mutation[164,167,168]. Moreover, the loss of p53 has been found to confer anti-androgen resistance[164,168].

The phosphatase and tensin homolog (PTEN) is a frequently mutated gene in PCa, with mutations being detected in 5-30% of localized cases and 30-60% of advanced carcinomas[169]. The frequency of PTEN gene loss has been found to be positively correlated with an increase in the Gleason score, a widely used grading system for PCa[170]. Loss of PTEN is more frequent in Gleason score 7 than in Gleason score 6, which makes it an early marker for the detection of PCa. PTEN is a lipid phosphatase that inhibits the activity of phosphatidylinositol 3-kinase (PI3K), which plays an important role in cell growth, proliferation and oncogenesis in PCa[171-173] and suppresses the downstream signaling of the mechanistic target of rapamycin, Akt. Consequently, loss of PTEN function leads to an increase in Akt-pathway activity, which

promotes cell survival in cancer cells. In contrast to other solid tumors where PTEN is inactivated by heterogeneous deletion, homozygous deletion of the PTEN gene is a unique characteristic observed in PCa[174]. PTEN deletions have been found in up to 20% of patients with PCa and have been associated with earlier biochemical relapse, metastasis, resistance to castration, presence of ERG gene fusions and the accumulation of nuclear TP53[175]. Particularly, loss of PTEN, together with TMPRSS2:ERG fusion, c-myc upregulation and NKX3.1 alteration, promotes prostate tumorigenesis[176-179]. Studies reported that PTEN loss occurs after TMPRSS2-ETS rearrangement and correlates with ERG expression[180]. PTEN loss is enriched two to five-fold higher in localized prostate tumors with ERG rearrangement, indicating a potential mechanism for the re-expression of androgen-responsive genes in castration-resistant PCa development. However, the loss of PTEN could also account for the loss of Tp53. Concomitant loss of tumor suppressor PTEN and Tp53 results in metastatic progression of PCa.

In terms of prognosis, PTEN deletion has been shown to correlate with reduced response to anti-androgens, limited progression-free survival, and high risk of relapse and metastases[169]. Studies in genetically engineered mouse models have revealed that monoallelic deletion of PTEN leads to the formation of prostatic intraepithelial neoplasia (PIN), whereas biallelic deletion of PTEN results in invasive and metastatic prostatic adenocarcinoma[170,181]. It should be also noted that other members of the PI3K cascade, including PI3KCA (13%), PIK3R1 (6%), NF2 (3%), AKT1 (1.5%), and NF1 (1.5%), are commonly found in PCa[182]. The PI3K signaling mutations in advanced tumors are associated with p53 and AR alterations, which results in castration resistance after androgen deprivation therapy[183,184]. PI3K/Akt/mTOR inhibitors, therefore, are under extensive study for the treatment of CRPC[185].

3.1.2 Decoding the Significance of Oncogene Alterations

Within normal cells, there are proto-oncogenes which play crucial roles as key regulators of various biological processes. They can function as growth factors, transducers of cellular signals, and nuclear transcription factors. These proto-oncogenes are present in mammalian genomes and are responsible for controlling the normal differentiation and proliferation of cells[186]. However, alterations to these genes, either in terms of their regulatory control or the structural conformation of their encoded proteins, may occur as activated oncogenes in cancer cells. When oncogenes are formed, they drive the uncontrolled multiplication of cells and assume a pivotal role in the development of cancer. The physical

mutations that result in the activation of proto-oncogenes can be categorized into two types. The first type involves mutations that lead to changes in the structure of the encoded protein. These alterations can impact the protein's function and contribute to the acquisition of oncogenic properties. The second type of mutation causes deregulation of protein expression, leading to abnormal levels of the protein within the cell[187,188]. Both types of mutations can ultimately lead to the activation of proto-oncogenes and the subsequent development of cancer. Indeed, the activation of an oncogene or inactivation of a tumor suppressor gene does not automatically result in every cell becoming a tumor cell. Even though the development of cancer can occur in various cell types within the body, the genetic causes of tumorigenesis are generally selective to specific cell types. A notable experiment involving the expression of *K-Ras*^{G12V} demonstrated the tissue-specific susceptibility of cells to oncogenes. Mice carrying this oncogene in all cells only developed lung adenomas and adenocarcinomas[189]. In a follow-up study, the Barbacid team demonstrated that a specific lung cell type integrates the *K-Ras*^{G12V} perturbation into its signaling network, enabling tumorigenesis. They identified SPC+ alveolar type II cells as the cancer-initiating cells, as these cells were capable of forming hyperplastic lesions that later progressed to adenocarcinomas[190]. The findings indicate that the function of an oncogene is not universal, and cancer development depends on specific perturbations occurring within defined cell types following a specific developmental trajectory. Mutational analyses across various cancer types have further confirmed the presence of tumor type-specific mutations[191].

In prostate cancer, there are three most frequently mutated oncogenes which are MYC, ERG, and AR. MYC is a transcription factor that regulates genes involved in cell growth, differentiation, and metabolism, and its amplification and overexpression are observed in approximately 30-40% of primary prostate cancers. Such amplification and overexpression are associated with higher Gleason scores, advanced disease stage, and poorer overall survival rates. In addition, MYC dysregulation is often co-amplified with other oncogenes in prostate cancer, such as AR and MYCN, and their combined overexpression is associated with poorer survival rates[192]. There are three different genes in the myc family: c-myc, l-myc (MYCL), and n-myc (MYCN). In prostate cancer, c-myc amplification and overexpression have been found in both early and metastatic tumors, which generally correlates with high Gleason grade and poor prognosis[192]. MX11, a c-myc upstream inhibitor, is also frequently mutated in prostate cancer[193]. MYCL amplification is usually detected in premalignant lesions and primary tumors,

whereas MYCN is overexpressed in 40% of aggressive castration-resistant prostate cancers[194,195].

ERG is another important oncogene that encodes a transcription factor responsible for regulating genes involved in cell migration, invasion, and angiogenesis. ERG is frequently activated in prostate cancer through chromosomal rearrangements that fuse the ERG gene with the androgen-responsive TMPRSS2 promoter. This fusion results in androgen-responsiveness and deregulated expression of the ERG gene, which promotes the growth and survival of cancer cells. ERG overexpression, observed in 50% of primary prostate cancers, is associated with more aggressive disease and poorer survival rates[196]. Mutations and amplifications of the AR gene are also frequently observed in advanced prostate cancers, with 20-30% of cases exhibiting such alterations[197]. The dysregulation of the AR signaling pathway plays a pivotal role in the initiation and progression of PCa. Activation of the AR by androgens triggers a cascade of physiological effects, including the binding of AR dimers to androgen response elements in the nucleus, leading to the transcription of target genes involved in the growth, development, and maturation of the normal prostate.

Despite extensive research and therapeutic interventions targeting the AR, accumulating evidence suggests that AR inhibition often fails to completely suppress AR signaling. Mechanisms underlying this phenomenon remain unclear, but reactivation of AR can occur in response to steroidogenic agents or alterations in the AR gene. Notably, various alterations in the AR gene, such as mutations, overexpression, and splicing, have gained substantial clinical significance in PCa. Generally, AR is expressed at a low level in normal prostate. However, when patients undergo ADT, AR expression is significantly increased (20%–30%) higher than normal prostate. This enhanced AR expression renders the prostate highly sensitive to low levels of androgens, promoting continuous prostate growth even under conditions of androgen deprivation[198].

Point mutations in the AR gene, including T877A, L701H, T878A, L702H, and V715M, have been extensively investigated and confirmed in several studies, particularly in patients treated with anti-androgens[180,199]. The most recent advancement in understanding resistance mechanisms in PCa involves the splicing of the AR gene. This splicing leads to the generation of AR splice variants characterized by rearrangement of exons or the complete removal of the ligand-binding domain or DNA-binding domain of AR. Notably, elevated messenger RNA levels of splice variants such as AR-V3 and AR-V9 have been implicated in

the metastatic progression of PCa[200]. Furthermore, increased expression of AR-V7 has been associated with a higher likelihood of developing castration-resistant PCa[201]. Although AR is not specific to the prostate, alterations in AR function play a major role in the malignant progression of PCa. Recent studies have provided further validation, demonstrating that AR is upregulated in over 60% of metastatic PCas[202]. Intriguingly, co-amplification of an enhancer region of the AR gene is frequently observed in castration-resistant PCas, highlighting the complexity of AR dysregulation in advanced disease[202]. Targeting the AR signaling pathway remains a cornerstone of therapeutic strategies for PCa. Various AR-targeted therapies, including abiraterone and enzalutamide, have been developed to inhibit AR signaling and combat disease progression. However, the multifaceted nature of AR alterations and the maintenance of AR signaling despite inhibition pose ongoing challenges in the effective management of PCa.

In addition, deregulation of downstream effectors of the AR signaling pathway, such as the PI3K/AKT and Wnt/β-catenin pathways, has also been observed in prostate cancer. Among these, the Wnt/β-catenin pathway holds significance as it plays a crucial role in regulating cell proliferation and differentiation, with frequent deregulation observed in cancer, including prostate cancer. Dysregulation of the Wnt/β-catenin pathway in prostate cancer can arise from mutations or amplifications affecting key components such as β-catenin, AXIN1/2, GSK3β or APC[203]. These alterations lead to the aberrant activation of the pathway, contributing to disease aggressiveness and resistance to androgen deprivation therapy. Understanding the molecular mechanisms underlying Wnt/β-catenin pathway dysregulation provides valuable insights into prostate cancer progression and therapeutic interventions.

To counteract the adverse effects of dysregulated Wnt/β-catenin signaling, targeted therapies have been developed to inhibit this pathway. Promising agents, including PRI-724[204,205], LGK974[206,207] and CWP232291[208], have emerged as potential therapeutic options. These compounds aim to selectively block key molecular components within the Wnt/β-catenin pathway and hold promise in mitigating tumor growth and overcoming treatment resistance. Ongoing clinical trials are currently evaluating the efficacy and safety of these targeted therapies in prostate cancer patients. These trials can shed light on the therapeutic potential of inhibiting the Wnt/β-catenin pathway and its impact on patient outcomes. Targeting the Wnt pathway as a potential treatment for prostate cancer holds promise, but it also presents several limitations and challenges that demand attention[209]. One of the primary concerns is the uncertainty surrounding the functional implications of activating the Wnt

pathway at different stages of prostate cancer progression. As the Wnt pathway is intricately involved in various normal physiological processes, therapeutically targeting it may result in unintended side effects and toxicity. Furthermore, the complexity of the Wnt pathway, which comprises multiple components, means that focusing on a single component might not be adequate to achieve the desired therapeutic benefits. Another significant challenge is the development of resistance to Wnt pathway inhibitors, emphasizing the necessity to identify biomarkers that can predict treatment response and aid in making informed decisions about patient care. Moreover, the efficacy of Wnt pathway inhibitors may depend on the genetic characteristics of the tumor, necessitating the identification of specific patient subgroups that are more likely to derive optimal benefits from this therapy. Addressing these limitations and challenges is crucial to maximize the potential of Wnt pathway targeting as a viable treatment approach for prostate cancer[209].

3.1.3 Changes in DNA methylation and histone modifications

In all human cancers, epigenetic changes cooperate with genetic alterations to drive the cancer phenotype[210]. Cancer genetics and epigenetics are inextricably linked in generating the malignant phenotype.Epigenetic alterations have the potential to induce mutations in genes, and, conversely, genetic mutations are frequently observed in genes that regulate the epigenome[210]. Epigenetic changes, such as DNA methylation, histone modifications and non-coding ribonucleic acids play crucial roles in prostate cancer pathogenesis. These patterns, including specific DNA hypo- or hypermethylated sites and expression of key regulators, have the potential to serve as diagnostic or predictive biomarkers. Aberrant DNA methylation patterns have been extensively observed in prostate cancer, involving both hypermethylation of tumor suppressor genes, including GSTP1 and APC, and hypomethylation of oncogenes like MYC and AR.

The most prevalent somatic alteration reported in prostate cancer is the aberrant methylation of the CpG island at the GSTP1 locus, which is detected in up to 90% of prostate cancer cases and 70% of PIN[211]. GSTP1 methylation is prevalent throughout all stages of prostate cancer and its overexpression in prostatic epithelial columnar cells due to aberrant methylation is linked to worse clinical outcomes. In addition to GSTP1, numerous genes, including adenomatous polyposis coli (APC), Ras-associated domain family 1A (RASSF1), O6-methylguanine DNA methyltransferase (MGMT), cyclin-dependent kinase inhibitor 2A (CDKN2A), death-associated protein kinase (DAPK), and tissue inhibitors of metalloproteinases (TIMPS), are hypermethylated in the promoter regions in prostate cancer

RASSF1A[212-214]. APC is known to suppress the Wnt/β-catenin signaling pathway, which plays a crucial role in tumorigenesis, embryonic development, and the maintenance of tissue homeostasis[215]. Hypermethylation of APC has been described in several studies, and abnormal methylation in its promoter region has been associated with poor prognosis[216,217]. RASSF1A methylation is observed in 60-70% of prostate carcinomas, particularly in high-grade tumors, suggesting its potential as a marker of aggressive disease[218,219]. The p16/CDKN2/MTS1 gene, a tumor suppressor gene, is frequently inactivated in prostate cancer through hypermethylation of its DNA sequence[220]. The methylation of these genes and others can lead to the silencing of multiple genes, including those crucial for regulating cell growth and differentiation. Consequently, this loss of gene function contributes to uncontrolled cell growth and the development of prostate cancer.

DNA methylation serves as a ‘defense’ mechanism in mammalian genomes to silence transposable elements[221] and also repetitive DNA, which constitutes a significant portion of the genome (at least 50%)[222], and prevent its propagation. Demethylation of normally methylated DNA, also known as hypomethylation, can disrupt such a ‘defense' mechanism. DNA hypomethylation refers to the demethylation of normally methylated CpG sites and contributes to upregulation of gene expression. DNA hypomethylation could contribute to oncogenesis by the activation of tumor-promoting genes. Although the precise mechanism underlying hypomethylation in cancer has not been fully elucidated, accumulating evidence over the past few decades shows that DNA hypomethylation is frequently observed in highly repeated DNA sequences or long interspersed transposable elements (LINE 1 or L1) in cancer. L1 promoter hypomethylation[223], L1 ORF1p protein expression[224], and somatic L1 retrotransposition[225] have been associated with numerous cancers, including lung, colon, pancreatic, and ovarian cancers. Global DNA hypomethylation is correlated with genomic instability, and contributes to cancer development and progression. However, DNA hypomethylation is more frequently observed in the late phase, such as metastasis, rather than the early stage of tumorigenesis in PCa[226,227]. While global hypomethylation is frequent in primary tumors, it is even more evident in metastatic prostate cancers[226]. DNA methylation may also play a role in the regulation of the PLAU gene in prostate cancer[228], with hypomethylation of the PLAU promoter being associated with its increased expression in hormone-independent prostate cancer cells, higher invasive capacity in vitro, and increased tumorigenesis in vivo[228]. However, in normal prostate epithelial cells and in hormone-dependent LNCaP cells, the PLAU promoter is methylated, resulting in lower expression of

the gene[228]. Other hypomethylated gene in prostate cancer include CAGE, a novel cancer/testis antigen gene[229]. Hypomethylation of CAGE, which occurs at a frequency of approximately 40% in prostate cancers, is responsible for its exclusive expression in cancer tissues[230]. The hypomethylation of specific genes, such as PLAU, CAGE, and heparanase, may have implications for the development and progression of prostate cancer.

Epigenetic modifications extend beyond DNA methylation to include histone modifications, such as acetylation and methylation, which can influence gene expression and contribute to prostate cancer progression. Histone modifications can alter the accessibility of DNA to transcription factors and RNA polymerase, resulting in altered gene expression patterns that promote cancer progression. Histone modifications, such as histone H3 lysine 27 trimethylation (H3K27me3) and histone H3 lysine 9 acetylation (H3K9ac), impact chromatin structure and gene accessibility. Aberrant histone modifications have been observed in prostate cancer, with decreased H3K27me3 associated with the upregulation of genes involved in cell cycle progression, proliferation, and metastasis, including EZH2[231]. Metastatic prostate cancer exhibits a higher enrichment of H3K27me3 marks at the promoter regions of tumor suppressor genes compared to localized PC and normal prostate tissues, thereby promoting increased cell proliferation. The increased genomic distribution of H3K27me3 in metastatic PCa is largely attributed to the overexpression of EZH2, a histone methyltransferase that catalyzes the addition of methyl groups to H3K27[232]. EZH2 overexpression is commonly observed in various cancers, including PCa, and represents a promising therapeutic target for PCa. Several reports have demonstrated the antiproliferative effects of targeting EZH2 in PCa[232,233]. Conversely, increased H3K9ac is linked to elevated expression of genes promoting cell growth and survival. Specifically, increased histone acetylation at the AR gene locus has been associated with androgen-independent growth and resistance to therapy, implicating the role of histone modifications in therapeutic resistance mechanisms[234]. The process of histone acetylation, specifically at lysine residues, is tightly regulated by the interplay between histone acetyltransferases (HATs) and histone deacetylases (HDACs)[235,236]. In the context of prostate cancer, the transcription of AR target genes is controlled by the assembly of a complex of transcription factors. Activation of the AR by agonists promotes the recruitment of the AR and coactivators possessing histone acetyltransferase activity to the promoters of AR target genes, resulting in histone acetylation and active transcription[237]. Also, studies have shed light on the role of super-enhancers, which are clusters of enhancers characterized by high H3K27 acetylation and strong enrichment of transcriptional coactivators such as BRD4 and mediator

of RNA Pol-II transcription subunit 1 (MED1), in driving oncogenic processes in various tumor cells[238]. These super-enhancers, marked by their high density of transcription factors and regulatory elements, share the common mechanism with "normal" enhancers of facilitating transcriptional activity through enhancer-promoter looping interactions. In TMPRSS2-ERG-positive prostate cancer, the overexpression of ERG leads to the formation of super-enhancers, enhancing the transcriptional activity of genes associated with prostate cancer development[239].

Similarly to DNA, RNA can also be modified. These modifications can be found in various types of RNA, ranging from mRNA to non-coding RNAs like ribosomal RNA (rRNA), transfer RNA (tRNA), miRNAs and lncRNAs[240]. Unlike DNA modifications, which mainly regulate gene expression[241], RNA modifications have diverse functions beyond transcription, such as controlling RNA stability, location, splicing, degradation, and translation efficiency[242,243]. The field of RNA epitranscriptomics, which explores RNA modifications, represents the cutting edge of cancer research. Numerous changes in RNA modifications have been associated with different types of tumors, including PCa[244-249], suggesting their potential as tumor biomarkers. For example, several studies showed that increased m^6A deposition (methylation of position N6 of adenosine in mRNA) has an oncogenic effect in prostate cancer cells and targeting METTL3 (methyltransferase-like 3) could have clinical benefits for PCa patients[244-246]. Studies on m^5C deposition in RNA and its role in PCa showed that increased expression of NOP2 (also known as nucleolar antigen p120 - is the only m^5C methyltransferase that has been demonstrated to be associated with PCa) has been considered a marker of bad prognosis for years, correlating with Gleason score, PSA serum levels and recurrence after radical prostatectomy[250,251]. RNA editing is also a very prevalent post-transcriptional event in human transcriptomes, mainly involving the conversion of adenosine (A) to inosine (I) by the ADAR enzyme family[252]. A-to-I editing is prevalent in various regions, including Alu elements, introns, untranslated regions (UTRs), and coding transcripts. In a comprehensive study of prostate cancer patients, over a hundred thousand putative RNA editing events were identified, affecting introns, UTRs, and coding regions, potentially leading to deleterious amino acid alterations[253]. Subsequent transcriptome-wide research revealed increased RNA-editing levels and higher expression of ADAR enzymes in various cancer tissues, including the prostate[254]. Additionally, rare germline heterozygous variants in ADAR were found to be linked to an increased susceptibility to prostate cancer[255].

In sum, epigenetic dysregulation, including changes in DNA methylation, histone modifications and nucleosome remodeling occurs at every phase of prostate cancer

development and progression. Understanding the complex interplay between genetic, epigenetic and also epitranscriptomic alterations in prostate cancer is important for the development of effective diagnostic and therapeutic approaches. Identification of specific methylation patterns and other epigenetic modifications holds great promise as potential biomarkers for early detection, prognosis, and novel therapeutic targets in prostate cancer. Further research is needed to unravel the precise mechanisms underlying these epigenetic alterations and their impact on prostate cancer progression.

3.1.4 The Role of ncRNAs in Prostate Cancer Development and Progression

The noncoding portion of the human transcriptome is increasingly recognized as playing a crucial role in the development and progression of cancer and boasts many attractive features for both biomarker and therapeutic research[256]. Genetic linkage studies show that many noncoding RNAs (ncRNAs) are located in cancer-associated genomic regions that are frequently deleted or amplified in prostate cancer, whereas aberrant ncRNA expression patterns have well-established links with prostate tumor cell proliferation and survival[257]. The dysregulation of pathways controlled by ncRNAs results in a cascade of multicellular events, ultimately contributing to carcinogenesis and tumor progression. The characterization of RNA species, their functions, and their clinical applicability is highly relevant area of both biological and clinical research[256].

Long noncoding RNAs (lncRNAs) and microRNAs (miRNAs) are two classes of regulatory ncRNAs that have garnered significant attention in cancer research. Mutations in lncRNAs and miRNAs can disrupt their normal regulatory functions, leading to aberrant gene expression patterns and contributing to cancer development and progression. Unlike protein-coding genes, the mutations in ncRNAs do not directly affect protein sequences but instead, influence the expression and function of mRNAs and proteins indirectly. These mutations can occur in the genomic regions encoding lncRNAs or miRNAs themselves or in the regulatory elements that control their expression.

Aberrant expression patterns of lncRNAs and miRNAs are frequently observed in prostate cancer and have been associated with disease progression and patient outcomes. These ncRNAs can act as oncogenes or tumor suppressors, modulating the expression of target genes involved in various cellular processes. For example, certain lncRNAs have been shown to interact with chromatin-modifying complexes, leading to changes in the epigenetic landscape and altering the expression of genes involved in cell cycle regulation and apoptosis[258].

Moreover, ncRNAs can regulate gene expression post-transcriptionally through interactions with miRNAs. Some lncRNAs can act as competing endogenous RNAs (ceRNAs), sequestering miRNAs and preventing them from binding to their target mRNAs. This ceRNA cross-talk can have profound effects on gene expression networks and contribute to prostate cancer progression[259]. Understanding the specific roles of ncRNAs and their target genes in these processes can help identify potential diagnostic biomarkers and therapeutic targets for prostate cancer.

Given the part of my thesis on miRNAs, I will provide additional details regarding miRNA biogenesis and its significance in the context of prostate cancer.

i. Unveiling the Mechanisms of miRNA Biogenesis: A Concise Overview

miRNAs are small noncoding regulatory RNAs (19–25 nucleotides) that play a major role in regulation of gene expression. They are responsible for the control of fundamental cellular processes that has been reported to be involved in human tumorigenesis. The characterization of miRNA profiles in human tumors is crucial for the understanding of

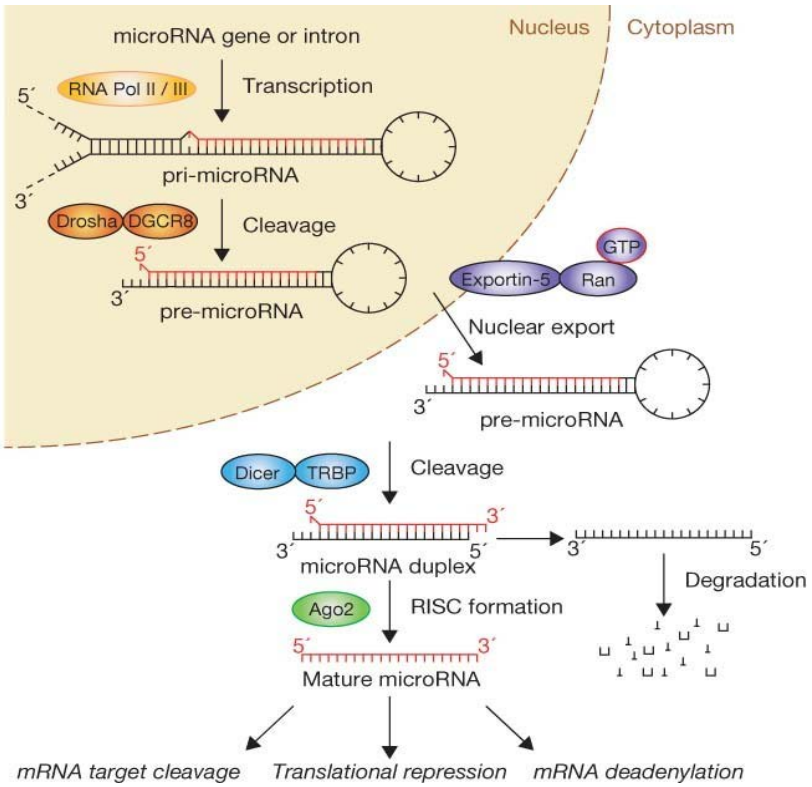


Figure 11 MicroRNA biogenesis[3].

carcinogenesis processes, finding of new tumor markers, and discovering of specific targets for the development of innovative therapies.

The biogenesis of miRNAs involves a series of well-coordinated steps, starting from their transcription to their incorporation into the RNA-induced silencing complex (RISC). miRNAs are transcribed from specific genomic loci by RNA polymerase II as long primary miRNA transcripts (pri-miRNAs). Pri-miRNAs can be independent transcriptional units or can be embedded within introns of protein-coding genes. In the nucleus, the pri-miRNA is recognized and cleaved by the Microprocessor complex, composed of the RNase III enzyme Drosha and its cofactor DGCR8[260] (DiGeorge syndrome critical region 8). The microprocessor complex recognizes the stem-loop structure of the pri-miRNA and cleaves it near the base of the stem, producing a hairpin-shaped precursor miRNA (pre-miRNA). The pre-miRNA is then transported from the nucleus to the cytoplasm by the Ran/GTP/Exportin 5 complex. Exportin-5 recognizes and binds to the pre-miRNA hairpin structure, facilitating its translocation through the nuclear pores. In the cytoplasm, the pre-miRNA is further processed by another RNase III enzyme called Dicer. Dicer recognizes the double-stranded hairpin structure of the pre-miRNA and cleaves it near the loop region, generating a short RNA duplex (~ 20–22-nucleotide miRNA/miRNA* duplex). From the RNA duplex, one strand, known as the guide strand, is selected to be incorporated into the RISC, while the other strand, known as the passenger strand or miRNA*, is usually degraded. The selection of the guide strand is determined by the thermodynamic stability and structural features of the duplex ends[261]. The guide strand within the RISC complex binds to the target mRNA molecules through partial complementarity between the miRNA seed region (a 6- to 8-nucleotide-long fragment at the 5′-end of the miRNA) and the target mRNA sequence[262]. The binding of the miRNA to the target mRNA leads to translational repression and/or mRNA degradation, resulting in post-transcriptional gene silencing[263]. It's important to note that miRNA biogenesis is tightly regulated at each step to ensure the proper production and function of mature miRNAs. Dysregulation of miRNA biogenesis can occur in various diseases, including cancer, leading to aberrant miRNA expression and disruption of gene regulatory networks.

ii. miRNA deregulation in cancer

The discovery of miRNAs has led to a global research effort to reveal their significance in various diseases. The number of human miRNAs reported so far (the June 2019 release of miRbase V22) is in excess of 2650[264] and extensive studies have revealed the critical roles miRNAs play in diverse biological processes such as proliferation, cell cycle regulation, apoptosis, differentiation, development, metabolism, neuronal patterning, and aging[265-267]. Moreover, their involvement in cancer formation and metastasis has been extensively

demonstrated. These small ncRNAs are typically found in specific regions associated with gene amplification, loss of heterozygosity, fragile sites, and common breakpoint regions close to oncogenes or tumor suppressor genes. Remarkably, each member of the vast family of ncRNAs can target hundreds of different genes, thereby exerting substantial influence over gene regulation and at least one miRNA is responsible for regulating nearly 30% of all genes[265].

miRNA dysregulation has been implicated in the development of each of the hallmark features of cancer, and restoration of expression of some of these critical downregulated miRNAs has been studied as a potential treatment for several different cancers. Understanding the mechanisms is crucial for uncovering the underlying molecular events that drive cancer development and progression. Key mechanisms of miRNA deregulation in cancer include genetic alterations, epigenetic modifications, dysregulation of miRNA biogenesis machinery, dysregulated miRNA targeting and post-transcriptional regulation, and cross-talk with signaling pathways and transcription factors. Genetic abnormalities such as chromosomal deletions, amplifications, and mutations can directly impact the genomic regions encoding miRNAs or their target sites in the 3' untranslated regions (UTRs) of target mRNAs. These genetic changes lead to the dysregulation of specific miRNAs and their target genes. For example, loss of heterozygosity (LOH) or deletions in chromosomal regions harboring miRNA genes can result in the downregulation of tumor-suppressive miRNAs, while amplifications or mutations can lead to the overexpression of oncogenic miRNAs. For instance, the loss of miR-15a/16-1 cluster gene located at chromosome 13q14, frequently observed in B-cell chronic lymphocytic leukemia patients, represents an early discovery of miRNA gene location alteration[268]. In lung cancer, deletion of the 5q33 region harboring miR-143 and miR-145 is common, leading to decreased expression of both miRNAs[269]. Conversely, amplification of the miR-17–92 cluster gene has been observed in B-cell lymphomas[270] and lung cancers[271]. Translocation of this cluster gene has also been observed in T-cell acute lymphoblastic leukemia, resulting in the overexpression of these miRNAs in these malignancies[272]. The epigenetic alterations found to influence the activity of miRNAs are the same as those previously described for protein-coding genes, such as DNA methylation and histone modifications.

Another important aspect is the dysregulation of miRNA biogenesis machinery. Perturbations in the components involved in miRNA biogenesis and processing can disrupt miRNA production in cancer. Key proteins such as Drosha, Dicer, and Exportin-5 are crucial for miRNA processing and maturation. Downregulation or functional impairment of these

proteins can result in decreased production of mature miRNAs, leading to dysregulated gene expression patterns associated with cancer[273-276]. Furthermore, dysregulated miRNA targeting and post-transcriptional regulation contribute to cancer development and progression. Aberrant expression of miRNAs in cancer cells can affect the targeting and post-transcriptional regulation of their target mRNAs. This dysregulation influences mRNA stability and translation, ultimately impacting the expression of genes involved in critical cellular processes. Dysregulated miRNA-target interactions can promote oncogenic pathways or suppress tumor-suppressive mechanisms, further driving prostate cancer progression.

Changes in the expression level of miRNAs have subsequently been detected by different groups in many types of human tumors. miRNAs have been proposed to contribute to oncogenesis because they can function either as tumor suppressors or oncogenes. OncomiRs, acting as oncogenes, are frequently overexpressed in cancer, promoting tumor growth and facilitating metastasis. They tend to be located within amplified regions of the genome. In contrast, tumor-suppressor miRs, which impair tumor growth and promote apoptosis while blocking cell migration, are often down-regulated in cancer. Their repression can occur through various mechanisms, such as mutations, promoter methylation, or chromosomal rearrangements[277]. Some microRNAs can act as both oncogene or tumor-suppressor gene depending on the cellular context[277]. Several well-described examples, such as miR-17, miR-21, miR155, miR-221/222, let-7, miR-15/16, miR-200, miR-34 indicate that they have important roles as oncogenic or tumor suppressor molecules in cancer biology.

A single miRNA possesses the capability to target multiple mRNAs, enabling it to regulate numerous target genes simultaneously, within a single pathway or even across different pathways[278], thereby exerting substantial influence over gene regulation and at least one miRNA is responsible for regulating nearly 30% of all genes[265]. For instance, while the miR-17 family targets various components of the transforming growth factor-beta (TGF-β) signaling pathway, including TGF-β receptors and down-stream transducers such as SMADs and cyclin dependent kinase inhibitor 1A[279], miR-21 suppresses key genes associated with major tumor suppressor functions, including Bcl2, PTEN, programmed cell death 4 (PDCD4), sprouty, Fas ligand, tissue inhibitor of metalloproteinase 3 (TIMP3), and tropomyosin 1 (TPM1)[280-282]. Increased expression of miR-21 leads to the downregulation of its multiple

target genes, promoting cell proliferation, inhibits apoptosis, enhancing angiogenesis, and facilitating metastasis, thereby contributing to cancer progression.

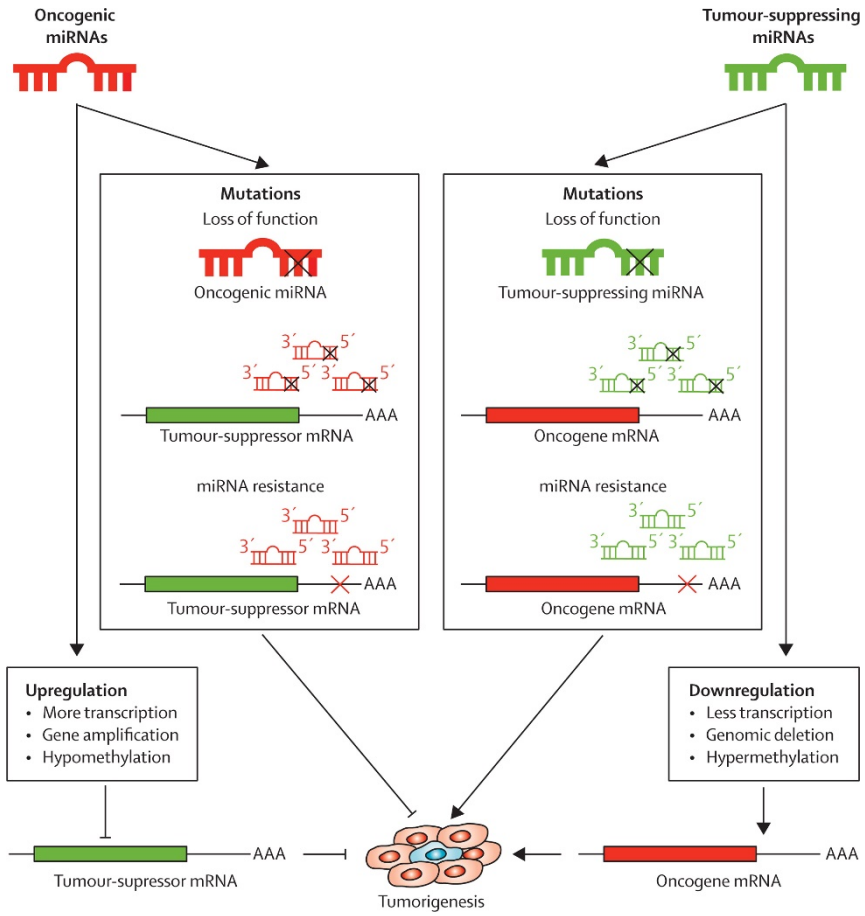


Figure 12 miRNAs play a crucial role in regulating tumorigenesis[8]. Oncogenic miRNAs, when upregulated, decrease the expression of tumor-suppressor proteins, while downregulation of tumor-suppressing miRNAs leads to increased production of oncogenic proteins. Tumorigenesis can occur due to loss-of-function mutations in tumor-suppressing miRNAs, as well as mutations in the target region of oncogene mRNA, which disrupt the regulation of oncogenic protein expression. Conversely, loss-of-function mutations in oncogenic miRNAs and mutations in tumor-suppressor mRNA would increase the expression of tumor-suppressor proteins, thereby reducing the likelihood of tumorigenesis.

One such highly conserved miRNA, miR-205, exhibits dual roles as both an oncomiRNA and a tumor suppressor. However, extensive research suggests its predominant function as a tumor suppressor in various cancers. Numerous studies have demonstrated that miR-205 expression is frequently downregulated in PCa, and it exerts its tumor suppressive effects by impacting migration, invasion, and growth through modulation of the AR and MAPK signaling pathway. Luciferase reporter assays demonstrated that miR-205 binds directly to the 3′ UTR of AR, and that ectopic expression of miR-205 has been shown to decrease AR expression at the mRNA level[283]. Other studies have revealed decreased expression of miR-

205 in PCa[284,285], partly explained by hypermethylation of the CpG islands in the miR-205 promoter, and to act as a tumour suppressor by affecting migration, invasion and growth[286,287].

MiR-27a and miR-27b have been demonstrated to function as putative tumor suppressors in prostate and bladder cancer by modulating proliferation and epithelial-mesenchymal transition. The study showed that miR-27a, as a tumor suppressor, decreased PCa cells proliferation and migration by suppressing MAP2K4, identified as a direct target of miR-27a, was an oncogene in prostate cancer[288]. On the other hand, miR-23a has been observed to be downregulated in polycystic ovary syndrome (PCOS) patients, indicating its inhibitory role in this condition. In cellular experiments, miR-23a mimic demonstrated the ability to inhibit cell proliferation and promote apoptosis and the study confirmed that miR-23a directly binds to the 3'UTR of FGD4, leading to the suppression of FGD4 expression and activates the CDC42/PAK1 pathway, leading to cell cycle arrest[289]. Another experiment showed that miR-23a downregulated the expression of IRF-1, which is involved in apoptosis suppression in hepatocellular carcinoma (HCC)[290]. In addition to individual miRNAs, miRNA clusters also play a significant role in cancer development. miRNA clusters consist of multiple miRNAs located in close proximity on the genome, and their coordinated expression can have a collective impact on cellular processes. For instance, the miR-17-92 cluster, known as oncomiR-1, is frequently amplified or overexpressed in several cancers, including lymphomas and lung cancer. This cluster targets multiple tumor suppressors, such as PTEN, E2F1, and BIM, promoting cell proliferation, inhibiting apoptosis, and facilitating angiogenesis[291].

Furthermore, the expression patterns of miRNAs can serve as valuable indicators for cancer classification and the identification of subtypes. Each cancer type exhibits unique miRNA profiles that reflect the tissue of origin and underlying molecular alterations. Through miRNA profiling, it has become possible to subclassify different cancers, including breast, lung, and prostate cancers, into distinct molecular subtypes that display diverse clinical behaviors[292,293]. This subclassification has implications for treatment decisions and patient management, as certain subtypes may respond differently to specific therapies. Therefore, the analysis of miRNA expression provides critical insights into cancer heterogeneity and aids in tailoring personalized treatment strategies based on the molecular characteristics of individual tumors.

iii. The potential use of miRNA-based therapies

The potential use of miRNA-based therapies is based on the ability of miRNAs to target multiple mRNAs that are altered in disease conditions, making them interesting candidates of therapeutics[278]. There are two main approaches to miRNA-based treatments: miRNA reduction (miRNA inhibitors) and miRNA replacement (miRNA mimics). In miRNA reduction treatment, inactivating miRNAs that are upregulated or overexpressed in tumors are targeted, while miRNA replacement treatment involves reintroducing miRNAs that are downregulated or deleted in tumors[294]. Anti-miRNAs (miRNA inhibitors) are synthetic single-stranded RNA molecules that bind and sequester mature miRNAs through sequence complementarity, counteracting the effects of upregulated miRNAs[295]. Conversely, miRNA mimics are synthetic oligonucleotide duplexes that, upon introduction into cells, function similarly to endogenous miRNAs and can restore the expression of downregulated miRNAs[294]. However, there are several challenges that remain, including the identification of the best miRNA candidates or miRNA targets, the design of effective delivery vehicles, and the avoidance of potential toxicities and off-target effects. One limitation is the ability to develop mechanisms to overcome miRNA-mediated repression, especially in cancer cells. This can occur through mutations within miRNA response elements or the expression of mRNA isoforms with shorter 3'UTR lacking miRNA-binding sites, achieved by alternative cleavage and polyadenylation[296,297]. In such cases, the absence of repressive elements in mRNA isoforms prevents miRNA mimics from binding and can result in the failure of therapy.

Another major challenge in miRNA-based drug therapy is the design of miRNA delivery vehicles that provide higher stability to the therapeutic candidate and enable tissue-specific targeting, while also avoiding potential toxicities and off-target effects. Off-target effects occur when a miRNA inadvertently regulates genes other than its intended targets. This can happen because a single miRNA can potentially bind to and regulate multiple mRNAs with similar or partially complementary sequences. For instance, if a therapeutic miRNA is designed to inhibit a specific oncogene, it might also unintentionally downregulate other non-cancer-related genes, causing adverse effects on normal cellular functions. These off-target effects can contribute to unwanted toxicity and limit the clinical application of miRNA-based therapies[298]. To mitigate off-target effects and improve the delivery of miRNA-based therapeutics several encapsulation methods have been developed. One of the most commonly used delivery systems is lipid nanoparticles, such as neutral lipid emulsions (NLEs) or dendrimer complexes with a targeting moiety attached[299]. These delivery systems can protect the miRNA from degradation

by nucleases and facilitate its uptake by target cells. Another approach to miRNA-based drug therapy is the use of chemically modified miRNAs. Chemical modifications, such as the addition of a 2′-O-methyl group or locked nucleic acids (LNAs), can increase the thermostability of LNA-RNA duplexes, increase target specificity, and are resistant to exonucleases and endonucleases, thereby improving the stability of miRNAs in vitro and in vivo[300]. In a study using mice, an anti-miRNA, miR-122, targeting a liver-specific miRNA successfully inhibited its function. The treatment resulted in the upregulation of specific genes and demonstrated a long-lasting effect[301]. Similar approaches using LNA-modified oligonucleotides showed potential for inhibiting endogenous miRNAs (such as oncomiR-155) and upregulating target proteins[300].

In addition to delivery and chemical modifications, the identification of the tumor-specific delivery and retention of miRNAs is crucial. Because many mRNAs are targeted by one miRNA, off-target effects are likely to be significant. Targeted delivery to specific tissues can be achieved by linking tumor-specific ligands to nanoparticles, which can be directed to tumor cells through active or passive targeting. Passive targeting utilizes the size of nanoparticles and tumor vasculature properties to selectively deliver the load to specific cells[302]. Active targeting of nanoparticles requires their conjugation with different compounds that have a specific affinity for tumors. Various cancer-associated cell-surface proteins (eg, HER2[303], EGFR[304], and CA-125[305]) and hyaluronic acid[306] could potentially be used for this conjugation.

In conclusion, the potential use of miRNA-based therapies in cancer treatment is an exciting area of research. While there are still challenges to overcome, the development of safe and specific methods for delivering of miRNA-based treatments to the cancer cells could allow modulation of miRNAs to become a central feature of cancer treatment and management.

Drug Name	Therapeutic agent	Background disease	Delivery system	Clinical trial number(s)	Phase
AMT-130	Artificial miRNA	Huntington disease	Viral transfer (adeno-associated vector	NCT04120493 NCT05243017	Stage Phase I/II, ongoing
CDR132L	AntimiR-132	Myocardial Infarction, Heart failure	LNA-modified antisense inhibitor	NCT04045405	Stage Phase I, completed
				NCT05350969	Stage Phase II, ongoing
RG-012	AntimiR-21	Alport Syndrome	Chemical Modification (Phosphorothioate)	NCT03373786N CT02136862	Stage Phase I, completed
MRG-110	AntimiR-92a	Skin Excisional Wound	LNA-modified antisense inhibitor	NCT03603431	Stage Phase I, completed
EXTH-61	AntimiR-21 AntimiR-10b	Glioblastoma	c-RGD peptide surface tagged PLGA nanoparticles (NPs)	NCT01849952	Stage Phase I, ongoing
MRG-201/ Remlarsen	miR-29 (mimic)	Keloid	Biomolecule conjugation (cholesterol)	NCT03601052	Stage Phase II, completed
				NCT02603224	Stage Phase I, completed
RG-125/ AZD4076	AntimiR-103 AntimiR-107	Non-alcoholic Steatohepatitis (NASH)	GalNAc-conjugated antimiR	NCT02612662	Stage Phase I, completed
MRG-106/ Cobomarsen	AntimiR-155	Cutaneous T-Cell Lymphoma/Mycosis Fungoides	LNA-modified antisense inhibitor	NCT03713320 NCT03837457	Stage Phase II, terminated
				NCT02580552	Stage Phase I, completed
MRX34	miR-34a (mimic)	Multiple solid tumours	Liposomal Injection	NCT01829971	Stage Phase I, terminated
Mesomir-1	miR-16 (mimic)	Malignant pleural mesothelioma, non-small cell lung cancer	Drug delivery vehicle - EDVs (nanoparticles)	NCT02369198	Stage Phase I, completed
RG-101	AntimiR-122	Chronic hepatitis C virus	GalNAc-conjugated antimiR	-	Stage Phase II, completed
Miravirsen	AntimiR-122	Chronic hepatitis C virus	LNA-modified antisense inhibitor	NCT02452814 NCT02508090 NCT01200420	Stage Phase II, completed
				NCT01727934 NCT01872936	Stage Phase II, unknown
				NCT01646489	Stage Phase I, completed

Figure 13 Clinical trials of selected miRNA therapeutics

THESIS OBJECTIVES

Study I : Patient-matched analysis identifies deregulated networks in prostate cancer to guide personalized therapeutic intervention

The objective of this study is to investigate the inter-individual heterogeneity observed within the stratified class of TMPRSS2-ERG-positive patients and propose a personalized therapy approach using network analysis. Despite the extensive genomics analyses conducted in cancer research, the identification and effective utilization of targets for pharmacological treatments have been disappointingly limited. This limitation can be attributed to the gene-centric views and analyses traditionally employed by molecular biologists.

While these approaches have resulted in significant discoveries, such as (proto)oncogenes, tumor suppressor genes, and therapeutics like Herceptin (Trastuzumab), there has been a lack of substantial conceptual progress in cancer therapy. Therefore, this project aims to shift the focus from analyzing a large number of patients and individual genes to studying individual patients and establishing a network of Differentially Expressed Genes (DEGs) to identify key nodes that may be targeted for therapy.

To achieve this objective, patient-matched normal and tumor samples were obtained from primary ERG-positive prostate cancers. Somatic variations specific to each patient were identified through exome sequencing (exome-seq), while differentially expressed genes (DEGs) were pinpointed through RNA sequencing (RNA-seq) in a patient-matched manner. To illustrate the complex interactions between genes and their physiological context, a strategy was generated to construct patient-specific networks since it is well-known that genes do not operate in isolation, but rather in coordination with other genes and cellular components. Master networks were reconstructed using all DEGs, integrating the connectivities provided by the validated STRING and CellNet databases. The resulting networks were visualized using Cytoscape software. Subsequently, the connectivity between these DEGs and mutated genes was investigated to understand how these alterations impact regulatory networks and pathways and enable the development of personalized therapeutic strategies.

Study II : Role of miRNAs in the differential inter-individual gene-regulatory networks in prostate cancer

Following the first study and given that we conducted a comprehensive ssRNA-seq analysis in which we identified also deregulated ncRNAs in each patient, we decided to include not only deregulated mRNAs but also DEmiRs and their target genes to enhance the complexity of our DEG network. The reason was that over the past two decades, the field of miRNA biology has significantly expanded. MiRNAs have been recognized as valuable tools and targets for novel therapeutic approaches due to their roles in development, disease, and particularly cancer. Extensive research has revealed that miRNA dysregulation plays a causal role in many cancer cases, where miRNAs can function as either tumor suppressors or oncogenes (oncomiRs) and clinical development of miRNA-targeted therapeutics has progressed, including a mimic of miR-34, as well as inhibitors of miR-122.

The primary questions we sought to address in our study were: Are there any causal links between miRs and their targets, and what are the impacts of these miRs on cancer deregulation? To answer these questions, we first analyzed the DEmiRs and identified their target genes using bioinformatic tools. Subsequently, we integrated these DEmiRs and their target genes into the previously established personalized networks. We further conducted experimental validations to assess the connectivity between the miRs and their target genes in prostate cancer cells. Additionally, we manipulated miRNAs using agomirs and antagomirs to observe the consequences of these alterations.

Overall, by employing a combination of bioinformatics analyses and experimental validations, we generated novel personalized multi-layer deregulated cancer networks with the aim of enhancing our understanding of the impact of miRNA deregulation in individual patients and elucidating the potential of targeting miRs for the development of novel therapeutic strategies.

RESULTS

PUBLICATION N°1

Patient-matched analysis identifies deregulated networks in prostate cancer to guide personalized therapeutic intervention

Akinchan Kumar, Yasenya Kasikci, Alaa Badredine , Karim Azzag, Marie L Quintyn Ranty, Falek Zaidi, Nathalie Aragou, Catherine Mazerolles, Bernard Malavaud, Marco A Mendoza-Parra, Laurence Vandel, Hinrich Gronemeyer

American Journal of Cancer Research, 2021

Am J Cancer Res 2021;11(11):5299-5318
www.ajcr.us /ISSN:2156-6976/ajcr0133575

Original Article

Patient-matched analysis identifies deregulated networks in prostate cancer to guide personalized therapeutic intervention

Akinchan Kumar[1,2,3,4,5], Yasenya Kasikci[1,2,3,4,5], Alaa Badredine[1,2,3,4,5,8], Karim Azzag[6,9], Marie L Quintyn Ranty[7,10], Falek Zaidi[7], Nathalie Aragou[7], Catherine Mazerolles[7#], Bernard Malavaud[7], Marco A Mendoza-Parra[1,2,3,4,5,11], Laurence Vandel[6,12], Hinrich Gronemeyer[1,2,3,4,5]

[1]Institut de Génétique et de Biologie Moléculaire et Cellulaire (IGBMC), Department of Functional Genomics and Cancer, Illkirch, France; [2]Centre National de la Recherche Scientifique, UMR7104, Illkirch, France; [3]Institut National de la Santé et de la Recherche Médicale, U1258, Illkirch, France; [4]Université de Strasbourg, Illkirch, France; [5]Equipe Labellisée Ligue Contre le Cancer; [6]Centre de Biologie du Développement (CBD), Centre de Biologie Intégrative (CBI), Université de Toulouse, CNRS, UPS, France; [7]Institut Universitaire du Cancer Toulouse-Oncopole (IUCT-O), Toulouse, France; [8]CNRS UMR8199-EGID Building, Lille University-Faculty of Medicine Henri-Warembourg, Lille, France; [9]Lillehei Heart Institute, Department of Medicine, University of Minnesota, Minneapolis, MN, USA; [10]Pathology Department, CHU, Caen, France; [11]UMR 8030 Génomique Métabolique, Genoscope, Institut François Jacob, CEA, CNRS, Université Evry-val-d'Essonne, University Paris-Saclay, Évry, France; [12]Université Clermont Auvergne, CNRS, Inserm, GReD, Clermont-Ferrand, France. [#]Deceased in 2015.

Received March 26, 2021; Accepted June 12, 2021; Epub November 15, 2021; Published November 30, 2021

Abstract: Prostate cancer (PrCa) is the second most common malignancy in men. More than 50% of advanced prostate cancers display the TMPRSS2-ERG fusion. Despite extensive cancer genome/transcriptome data, little is known about the impact of mutations and altered transcription on regulatory networks in the PrCa of individual patients. Using patient-matched normal and tumor samples, we established somatic variations and differential transcriptome profiles of primary ERG-positive prostate cancers. Integration of protein-protein interaction and gene-regulatory network databases defined highly diverse patient-specific network alterations. Different components of a given regulatory pathway were altered by novel and known mutations and/or aberrant gene expression, including deregulated ERG targets, and were validated by using a novel *in silico* methodology. Consequently, different sets of pathways were altered in each individual PrCa. In a given PrCa, several deregulated pathways share common factors, predicting synergistic effects on cancer progression. Our integrated analysis provides a paradigm to identify druggable key deregulated factors within regulatory networks to guide personalized therapies.

Keywords: Cancer systems biology, prostate cancer, personalized therapy, patient-matched deregulated networks

Introduction

The mutational landscapes of primary and advanced/metastatic PrCa have been extensively analyzed [1-5], as has been the prevalence of the androgen-sensitive TMPRSS2 promoter fusion with ETS transcription factors [6], which endows ETS with responsiveness to the androgen receptor (AR) that is frequently overexpressed in antiandrogen-resistant PrCa [7]. In these studies, recurrent mutations have been found in genes coding for factors regulating a plethora of pathways and key cellular functions, such as the androgen receptor signaling, PI3K/RAS/RAF/WNT pathways, and factors involved in DNA repair and chromatin methylation, or cell cycle control. One of the caveats in all these studies was that, with a few exceptions [2, 8], information was generally compiled from large numbers of tumors from different patients. Thus, while enabling identification of predominant mutations, these studies did not reveal the spectrum of aberrations that existed in individual patients' prostates at diagnosis. All these aberrations may affect different regulatory pathways and their added, possibly synergistic action may be critical for malignancy and tumor progression. Indeed, restoring a normal state would require the correction of a highly complex and dynamically

regulated system of interactive multi-component networks which are deregulated in disease [9]. Towards this goal, the identification of aberrant networks and their inherent hierarchies is essential to design patient-selective therapeutic interventions through generic or key factor-specific modulation of the affected pathways.

Material and methods

Patient sample collection

All samples were collected within 15 minutes after radical prostatectomy to shorten the delay between de-vascularization and freezing, and to ensure preservation of labile molecules. Immediately following prostatectomy, punch biopsies ("carrots") of 8 mm diameter were taken from tumor and adjacent normal tissue, snap-frozen in liquid nitrogen and stored at -80°C. Carrots used for genomic and transcriptomic studies were cut into sequential tissue sections and the tumor cellularity was monitored at regular intervals by histological staining to ensure homogeneity of tumor and normal sections stored in LoBind tubes at -80°C.

Tissue microarrays

Tissues microarrays (TMA) were made from paraffin-embedded tissue cores of histopathology-confirmed prostate cancer and patient-matched tumor-adjacent normal tissue. For each tumor, two representative tumor areas were selected and two cores of 2 mm in diameter were punched and included in paraffin recipient blocs. Two adjacent normal tissues of each selected prostatic sample were arrayed on TMAs and constituted the "normal" counterparts of each tumor sample. The TMAs were performed on the histopathology platform of the Biological Resource Center (CRB) of the Toulouse University Hospital, in a semi-automated way using the EZ-TMA™ Manual Tissue Microarray Kit (IHC World). The slides were examined by HE coloration and immunohistochemical studies were performed on TMA tissues. Immunohistochemistry was done using an automated Dako Autostainer. The following antibodies were used: ERG, EZH2, Androgen Receptor. Slides were digitalized using a Hamamatsu NanoZoomer slide scanner (Japan) at 20× magnification with a resolution of 0.46 microns per pixel. The results were interpreted under an optical microscope by two pathologists (CM and M-LQ), blinded to the clinical data.

Whole exome sequencing (WES) and analysis pipeline

For WES, DNA was isolated from frozen tumor and matched normal tissue (10 sections of 10 µm for each tumor and normal sample) using QIAamp DNA micro kit according to manufacturer's instructions. DNA was processed by GATC Biotech for exome capture, library preparation and sequencing. Briefly, SureSelectXT Human all exon V6 kit was used to capture exons, libraries were prepared using TruSeq DNA library preparation kit according to manufacturer's instruction and Paired-End 125-base sequencing was performed on Illumina HiSeq 2000. FastQ files provided by GATC Biotech were processed for variant discovery with Genome Analysis Toolkit (GATK, 3.7) [10] using default parameters. To assist in WES analysis we developed a WES Analysis Pipeline [written in Python3 with the Snakemake (3.13.3) [11] management tool] and used the Genome Analysis Toolkit (GATK, 3.7) [10] according to the authors' instructions. The following tools were used for each step in the pipeline.

Pre-processing of the samples: FastQ files were first aligned to hg19 using BWA-mem (0.7.17) [12] using standard parameters. The output SAM files were then converted to BAM files using SAM tools (1.6) [13]. BAM files were processed using Picard tools (2.14) to sort by coordinates, remove duplicates and add read group tags (essential to differentiate between Normal and Tumor samples) to samples before indexing them with SAM tools. Then BAM files were recalibrated (BQSR) using GATK, as recommended for enhancing variant calling by providing databases of known polymorphic sites: a set of curated INDEL entries, a Single Nucleotide Polymorphism database dbSNP, the COSMIC database of somatic cancer mutations.

Creating the Panel of Normals (PON): The 'Panel of Normals' is created from the normal samples using GATK. This method is used as a filter to reject artifacts and germline variants that are present in at least two normal samples (-minN 2). It uses as input the hg19, the dbSNP, the COSMIC and the intervals of the genome to analyze only the exons of all genes captured. It generates a new file that will be used when calling the variants.

Variant calling with MuTect2: After pre-processing and alignment, MuTect2 (GATK) was used to call somatic variants [14]. Inputs are the following files: hg19, PON, HumanAllExonV6r2, COSMIC and dbSNP. The normal and tumor samples were compared using the following parameters: pir_mad_threshold: 6; max_alt_alleles_in_normal_count: 5; pir_median_threshold: 35; standard_min_confidence_threshold_for_calling: 30.

Annotation of VCF: The annotations of the VCF files were done using SnpEFF [15] and SnpSift [15] with hg19 as reference genome.

Validation of mutations

Target regions were amplified by PCR. PCR products were purified using Qiagen gel extraction kit and sequenced by Eurofins Genomics using the BigDye Terminator Cycle Sequencing Kit and an ABI 3730xl automated sequencer (Applied Biosystems). The sequencing primers were the same as those used for PCR amplification. Variants were confirmed using SNAP gene viewer.

RNA-sequencing

RNA was isolated from frozen tumor and matched normal tissue using Trizol reagent (Invitrogen). In all cases, two independent sets of adjacent 10 µm sections (N=10) were processed for RNA isolation to generate biological duplicate RNA-seq data. For subsequent interpretations only data that were consistent between the biological duplicates were retained. RNA was further cleaned up using RNeasy MinElute Cleanup Kit. RNA was then sent to GATC Biotech (Konstanz, Germany) for strand-specific, paired-end and Ribo-minus total RNA-seq. Briefly, ribosomal RNA depletion was done using Ribo Zero gold kit (Illumina Inc); libraries were prepared using TruSeq stranded total RNA library prep kit (Illumina Inc.). Paired-end 125 base or 150 base sequencing was performed using Illumina HiSeq 2000. FastQ files received from GATC Biotech were used for further analysis.

RNA-seq analysis pipeline

The analysis pipeline consists of the following steps.

Pre-processing, alignment and counting raw reads: FastQ files were assessed for quality using FastQC. FastQ files were aligned to reference genome (human genome hg19) using the Hisat2 [16] aligner. Aligned SAM files were converted to BAM files and sorted using SAMtools [13]. The R package Summarized Experiment [17] was used for counting raw reads per exon/gene.

Differential gene expression analysis: The patient-specific differential gene expression analysis was done using DESeq2 (1.20.0) [18] according to the general steps described with the parameters given below. The samples have been analyzed by giving the matched raw read counts normal/tumor duplicates as input.

- Removing sum of row counts: 0;

- CooksCutoff: False;

- Alpha: 0.01;

- Subset genes with Adjusted *P*-value ≤0.01;

- Subset genes with Log2FC ≤-1 or Log2FC ≥+1.

The corresponding list of Differentially Expressed Genes (DEGs) for each patient was used for further analysis.

Pathway enrichment analysis

To interpret the gene expression data, the DEG list was loaded into GeneCodis [19] and the Panther pathway analysis function was used to retrieve enriched pathways. Hypergeometric correction of *P*-values was applied and pathways displaying a corrected *P*-value <0.01 were considered enriched. We clustered all pathways of all samples using Plotly in R. For the datasets obtained from TCGA (54 PrCa patient data along with matched normal), HT-seq counts were downloaded for each patient corresponding to tumor and matched normal. DEseq2 was used to identify the DEGs for each patient. Pathway enrichment analysis was performed as described before.

Patient-specific network generation and visualization

To generate the gene networks for individual patients we extracted the list of mutated genes from WES and differentially expressed genes

(DEGs) from RNA-seq of tumor *vs* normal samples for each patient. These lists of genes were queried against two known databases of network interactions, STRING [20], a Protein-Protein Interaction (PPI) database, and CellNet [21], a gene regulatory network (GRN) database. For STRING, we merged the list of genes (DEGs and Mutation, keeping the information whether the gene is a DEG or a mutated gene as attributes), removed any duplicated genes and queried them using an in-house script. As for the parameters, we only chose edge interactions that have been experimentally validated (exp_score ≠ 0). For CellNet, we queried only the differentially expressed genes on the target genes and retrieved along the cognate transcription factors. We chose interactions that had only a z-score ≥5. After obtaining networks from both databases, we proceeded to add the information from WES and RNA-seq whether the genes were mutated, differentially expressed or both, in addition to the information obtained from the databases.

Network visualization and merging using Cytoscape

Individual networks, created by using Cellnet and String for each patient, were visualized using Cytoscape [22]. Finally, CellNet and STRING networks for each patient were merged using the Cytoscape merge function to obtain master networks for each patient. Sub-networks were then extracted for further visualization and analysis.

Identification of putative AR and ERG target genes

A two-step approach was used. First, we collected sequenced read files (bed format) associated to public ChIP-seq assays targeting ERG in TMPRSS2-ERG positive human VCaP prostate cancer (GSM1328978, GSM1328979) and RWPE-1 normal prostate epithelium cells (GSM2195103, GSM2195106). BED Replicate files per cell-type were merged together prior performing peak calling (MACS 1.4; no model, shiftsize =150 nts, *P*-value threshold: 1×10^{-5}), followed by their genomic annotation to the closest transcription start sites (annoPeakR). This analysis allowed to pair the characterized DEGs and mutated genes within the patient-derived networks with genes presenting proximal AR binding sites (<10 kb distance) on VCaP ChIP-seq profiles. This primary analysis has been validated in a second step by comparative visual inspection of ChIP-seq profiles. For this we used the qcGenomics platform, in which the dedicated genome browser NAVi allows to visualize any publicly available ChIP-seq profile. Specifically, we used NAVi to extract all AR and ERG ChIP-seq profiles for TMPRSS2-ERG positive human VCaP prostate cancer and RWPE1 normal prostate epithelium cells. The pre-computed datasets were displayed simultaneously in the NAVi browser for comparative visualization. Only tracks with an apparent high signal-to-noise ratio were retained (VCaP-ERG: GSM2058880, GSM1328978, GSM1378979, GSM1328980, GSM1328981; VCaP-AR: GSM-1410768, RWEP1-ERG: GSM927071, GSM-2195110, GSM2195103; VCaP-GROseq: GSM-2235682). Promoter-proximal ERG binding was scored positive in this visual 'validation' (attributing a yellow color to the respective nodes) only when there was a clearly visible peak above the background at a scale of 30 to 300 (read count intensity; depending on the signal and noise intensities of each profile), provided that there was no other known TSS closer (see Supplementary Figure 6 for examples of gain of ERG binding).

Data availability

The RNA-seq data sets generated in the context of this study from 15 patient-matched tumor and normal prostate tissue are available in the Gene Expression Omnibus (GEO) repository under the accession number GSE133626. The corresponding Exome-seq data sets from the prostates of the same 15 patients are available from the SRA database under the accession number PRJNA555457. Data sets for the pathway analysis shown in Supplementary Figure 3 were downloaded from The Cancer Genome Atlas (TGCA) as described in the methods section. The sequencing statistics of RNA-seq and Exome-seq experiments are specified in Supplementary Table 4.

Results

Overview of the approach

We chose prostate cancer as a solid tumor paradigm to integrate patient-specific differentially expressed (DEGs) and mutated genes, using information from protein-protein interaction

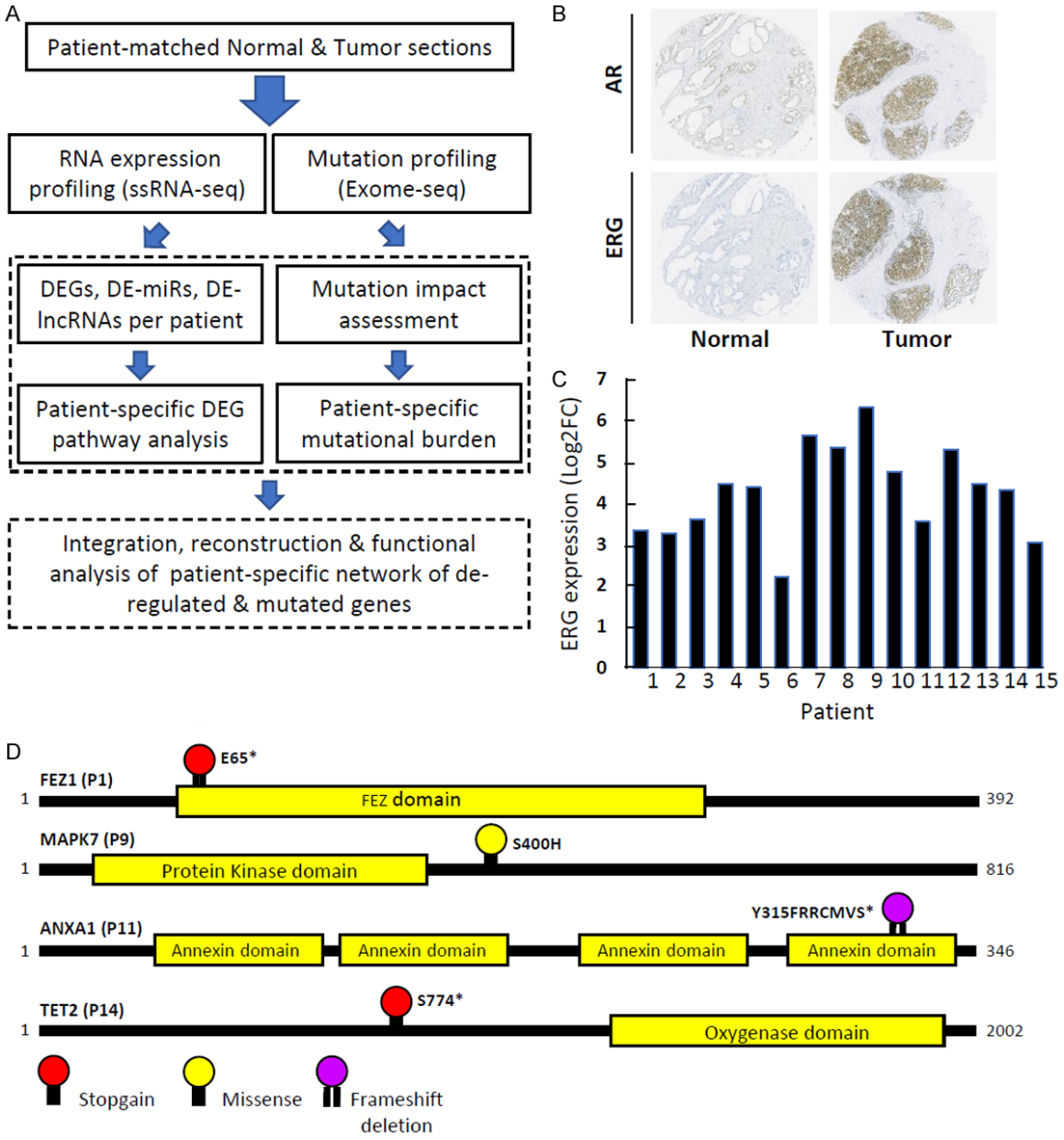


Figure 1. Analysis strategy and characterization of patient-matched samples. A. Sketch of the workflow of this study. B. Representative IHC images of cancer and corresponding matched normal samples from patient 15 (P15) stained with an anti-AR antibody (top panel) or an anti-ERG antibody (bottom panel). C. RNA-seq data of ERG expression in all the patients' tumors relative to their matched normal. Differential ERG expression for all patient-matched duplicate samples was supported by q value $<10^{-69}$ using DEseq (see Methods for details). D. Schematic illustration of 4 novel mutations in P1, P9, P11 and P14, which are predicted to have a high or moderate impact.

and gene-regulatory network databases [21, 23] (**Figure 1A**) to generate patient-specific cancer-modified networks. Extensively characterized normal and tumor frozen punch biopsies from the same prostate were obtained from radical prostatectomy specimens of non-treated patients. 15 primary ERG-positive tumors (T) and matched normal tissue (N) were selected by expert pathologists on the basis that consecutive sections of the same biopsy differed only minimally in tumor cellularity (>80% tumor cells), while the sections of N biopsies from the same prostate had 0% tumor cells. With one exception of patient 14 (P14), the proportion of infiltrating lymphocytes relative to tumor cells was close to 0%, only occasionally rare scattered lymphocytes were observed in the stroma. The tumor sections of

P14 showed up to 25% (area-based) mononuclear immune cells. Immunohistochemistry and RNA-seq (biological duplicates) confirmed ERG overexpression relative to the matched N samples and all samples revealed increased androgen receptor (AR) levels (**Figure 1B**, **1C**).

Identification of large-scale patient-specific genomic changes using exome sequencing analysis

To identify somatic variation, we performed whole-exome sequencing (WES, Supplementary Table 1). Variants were called using MuTect2, which revealed between 49 and 114 mutations in each cancer relative to the corresponding normal prostate tissue; only mutations predicted to have high or moderate impact were considered subsequently (Supplementary File 1). Intriguingly, in addition to classical mutations, for example in *MYC*, *TP53, PTEN* or components of the PI3K and WNT pathways [1, 2, 24], unreported patient-specific mutations were observed in all samples (Supplementary Table 2*;* for validations see Supplementary Figure 1). In P1 three hitherto unreported somatic mutations affected the putative tumor suppressors *BANP* [25], *FEZ1* [26, 27] (**Figure 1D**) and *TINAGL1*, which interferes with both integrin and EGFR signaling [28]. We also found novel mutations in *MAPK7* (R400H) in P9, Annexin A1 (*ANXA1*, frameshift deletion; P11) and a *TET2* mutation that truncates the protein and renders it non-functional (**Figure 1D**; P14). These novel somatic mutations were seen only in single patients. However, the nature of the mutations, often truncating proteins of functional importance, is likely to have a significant impact in the individual case. Moreover, several of those genes were found mutated in other cancer types, supporting their functional impact (Supplementary Figure 2). The ability of TINAGL1 to inhibit progression and metastasis of triple-negative breast cancer [28], provides strong rational for such personalized genomic analysis. Our data underscores the recent notion that "significantly mutated genes" in PrCa may occur at frequencies of only a few percent [29].

RNA-seq analysis revealed patient-specific altered tumor transcriptome

Mutations in regulatory elements (e.g., enhancers) and factors (e.g., transcription factors, epigenetic modulators, enzymes) can affect global gene expression. To integrate these effects in the network analysis, we performed high-throughput strand-specific paired-end total RNA sequencing after ribosomal RNA depletion from matched T and N biopsy sections as biological duplicates. As expected, T *vs*. N analysis of the RNA-seq datasets identified tumor-specific differentially expressed genes (TS-DEGs; Supplementary File 2) with diverse functionalities, comprising (i) cancer-specific deregulated proto-oncogenes like c-*MYC* (all except P6, P11, P13) but also (ii) pleiotropic factors like the serine protease KLK4 (P4, P10, P14, P15), a regulator of AR and the PI3K/AKT/mTOR pathway [30] and of protease-activated receptors [31]. Notably, deletion of *KLK4* impairs PrCa growth [30]. Moreover, (iii) epigenetic modifiers like *JMJD6* (P14), *KDM4B* (P5), *KDM6A* (P2, P8), *KDM6B* (P6, P9), *TET3* (P12, P14), *KAT2A* (P3), *KAT6A* (P2) or *HDAC9* (P2-5, P7-11, P13-15) were differentially expressed in certain tumors. In addition to protein-coding genes, expression of (iv) certain regulatory RNAs was altered in tumors [micro-RNAs (miRs)], as well as long non-coding RNA (lncRNAs); for annotated miRs and lncRNAs, see Supplementary Table 3. Of note, the p53-inducible lncRNA NEAT, a promising therapeutic target whose ablation generates synthetic lethality with chemotherapy and p53 reactivation therapy [32, 33], was over-expressed in 7/15 PrCa samples. A prominent ERG binding site in VCaP and in normal prostate epithelial RWPE-1 cells about 4.6 kb upstream of the NEAT transcriptional start site may account for this deregulation (e.g., GSM2086313, using the qcGenomics browser). The androgen-responsive lncRNA ARLNC1 [34] was up-regulated in 9/15 paired samples but down-regulated in P5. HOTTIP, a component of H3K4 methyltransferase complexes [35] that can act as AR co-activator [36] and was reported as negatively androgen-regulated lncRNA [34] in prostate cancer cells, was downregulated in 9/15 PrCa samples. This included several, but not all of those with up-regulated ARLNC1. A similar divergence was seen with putative tumor suppressor and oncogenic miRNAs that are actively considered for clinical development [37]. For example, the RNA levels of tumor-suppressor miR34a were decreased in three samples (P10, P13 and P15) but increased in P4 and not affected in 11 other samples. MiR222, which displays targe-

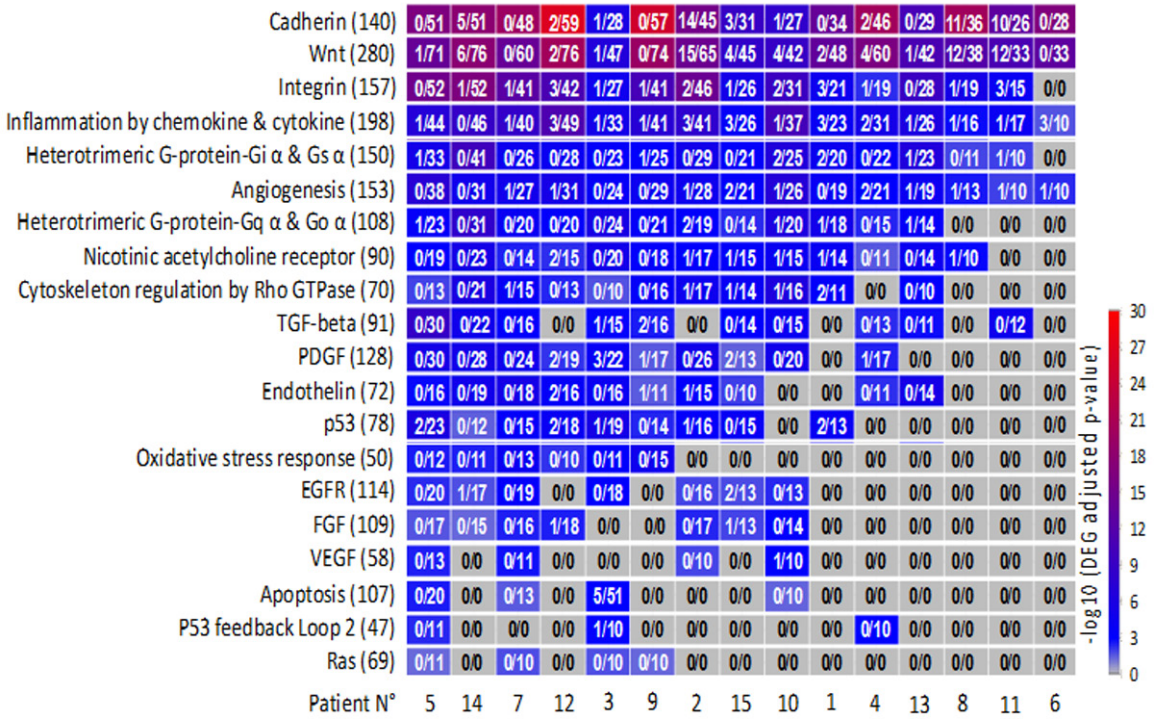


Figure 2. Divergence of pathways and severity of pathway alteration in individual prostate cancer patients. A selection of pathways predicted by Panther to be significantly enriched in patient-specific DEGs are shown on the left of the table while patient identification numbers are given at the bottom. Panther-computed *P*-values for the deregulation of a given pathway are illustrated as blue-to-red color-coded rectangles; grey color indicated no significant alteration. In the rectangles, the numbers of mutated and deregulated genes are given for each pathway and patient (mutated/DEG; see Methods for details).

table oncomiR characteristics in liver, pancreas and lung tumors [12, 37], was unexpectedly down-regulated in 8/15 PrCa samples. Together, these vastly divergent genetic mutations and altered, often counter-intuitive gene expression patterns revealed the need to decipher for each individual patient the complexity of the deregulated systems to identify key targets in critical signaling pathways and/or key nodes in (sub)networks for concomitant intervention at several functionally different levels to generate synergistic effects.

Pathway analysis revealed deregulated pathways in a patient-specific manner

As first step towards the integration of the various deregulated functions within each tumor, we performed patient-centered pathways enrichment analyses for TS-DEGs using Panther in the GenCodis3 environment. While several pathways were commonly deregulated in PrCa of several patients-particularly cadherin, Wnt and integrin signaling-these analyses also demonstrated that in each patient different sets of pathways were deregulated. Indeed, P5 and P6 had, respectively, the most and least severely affected PrCa in terms of numbers of deregulated pathways (**Figure 2**). This finding was further supported by analysis of additional 52 patients from the TCGA repository (Supplementary Figure 3). Moreover, different numbers and components of a commonly deregulated pathway were altered in different patients. As pointed out previously [24], genetic mutations of core Wnt pathway components are rare in PrCa, while abnormal expression of β-catenin is frequent, suggesting that this deregulation occurs indirectly.

Integration of genomic and transcriptomic datasets to generate patient-specific networks

Genes never function in isolation but rather in a highly complex physiological context, which can be illustrated by their communication with other cellular components. To gain a more precise insight into the altered communication by

patient-specific gene deregulations and mutations, we reconstructed master networks from all deregulated genes for each prostate cancer by integrating the connectivity provided by the validated STRING protein-protein and CellNet transcription factor-target gene interaction databases; in addition, we integrated all mutated genes and putative ERG and AR target genes identified by cognate binding sites in the vicinity of the transcriptional start site (TSS) [see Supplementary File 3 (cytoscape master-file for each patient)]. Within these master networks, we studied first the components of the canonical and non-canonical Wnt pathways by merging all 183 deregulated/mutated genes of 15 patients (Supplementary Figure 4A). Displaying the affected components in color in the context of the entire Wnt pathway connectivity revealed an unexpected heterogeneity (Supplementary Figures 4B-L and 5). In P2 (**Figure 3A**) an important signaling factor (phospholipase PLCB1) for the production of second messenger molecules (DAG, IP3) is mutated in the phospholipase domain (V571M) and the expression of multiple other master genes is deregulated, including *PPP3CA*, *GSK3B*, *MYC*, *TP53*, *HDAC1* in addition to several *WNT* and *Frizzled (FZD)* receptor genes. For several of the upregulated DEGs (*GSK3B*, *HDAC1, FZD8*; red arrows) drugs exist which have been approved or are tested in clinical trials [39]. Several druggable genes *(HDAC1, FZD5, FZD8, and MAP3K9)* are also upregulated in P9 but not in P6 while P10 showed overexpressed *FZD8* (**Figure 3B-D**). Notably, ChIP-seq data of TMPRSS2-ERG-positive VCaP and normal RWPE1 prostate epithelial cells indicate that *WNT7B* and *HDAC1* are putative dual AR and ERG target genes, most likely affected by deregulated ERG and possibly, AR signaling (Supplementary Figure 6A, 6B). Even more strikingly, the genes of several key signaling factors (PAK1, CREM) and of the epigenetic modulator SMARCD3 have apparently acquired ERG binding capability in their promoter regions during tumorigenesis (for *SMARCC1,* Supplementary Figure 6C), as it was reported for the ERG-mediated repression of checkpoint kinase 1 [40]. In contrast, P6 showed a very small number of deregulated components of the core Wnt pathway (**Figure 3B**), comprising three upregulated *FZD* receptors along with the cognate *WNT2* ligand acquired ERG binding near the TSS in VCaP cells (Supplementary Figure 6D, 6E). Such a scenario may be addressed with WNT inhibitor-based therapeutics [39]. Patient-specific network alteration was also seen for less frequently affected signaling pathways. The PDGF and EGFR pathways were affected seriously in 10 and 7 patients, respectively (**Figure 4A**, **4E**; merged networks of alterations). However, the scenarios were completely different across individual patients (**Figure 4**; Supplementary Figures 7 and 8). Important changes were seen in P4 and P5 (**Figure 4B**, **4C**) but hardly any in P13 (**Figure 4D**). Note that P13, in contrast to the other patients, did not reveal any upregulated druggable target (red arrow in **Figure 4**). A similar scenario of alterations was found for the EGFR pathway in P5 and P15 (**Figure 4F**, **4G**), while much less nodes were affected in P8 (**Figure 4H**).

Crosstalk among different deregulated pathways in same patient

Finally, given that pathways do not act in isolation, we extracted the affected components of several pathways from the “master networks”. This analysis showed very clearly that, for P2 and P5 several genes of the Wnt, cadherin and integrin pathways are shared between two or even three pathways (**Figure 5**); the same was observed for other combinations of pathways (Supplementary Figure 9). The functional consequence of deregulation/mutation of such genes is predicted to be serious and such nodes may be candidates for therapeutic targeting. It is worth pointing out that also genes at the nexus of several pathways diverged from one patient to another, as shown for P2 and P5 (**Figure 5A**, **5B**). Indeed, hypothetical treatment of these two patients would have to consider different scenarios. In the PrCa of P2, common to two pathways, there is a strong upregulation of the expression of several *WNT* and *FZD* genes, as well as *GSK3B* and *LEF1*. *GSK3B, FZD8, FRK* and *HDAC1* are druggable targets and several compounds have been approved. These genes are functionally connected with important other upregulated genes of the Wnt-pathway, such as *TP53*, *MYC* or *PPP3CA*. For P5, only two *FZD* genes are overexpressed in cancer, including ERG-induced FZD8. Moreover, *HDAC1* is mutated and *MYC* is repressed. On the other hand, *RANBP2* is uniquely overexpressed in P5. Given its multi-functional role in scaffolding for the Ran-GTPase cycle and nuclear pore complex binding, its overexpres-

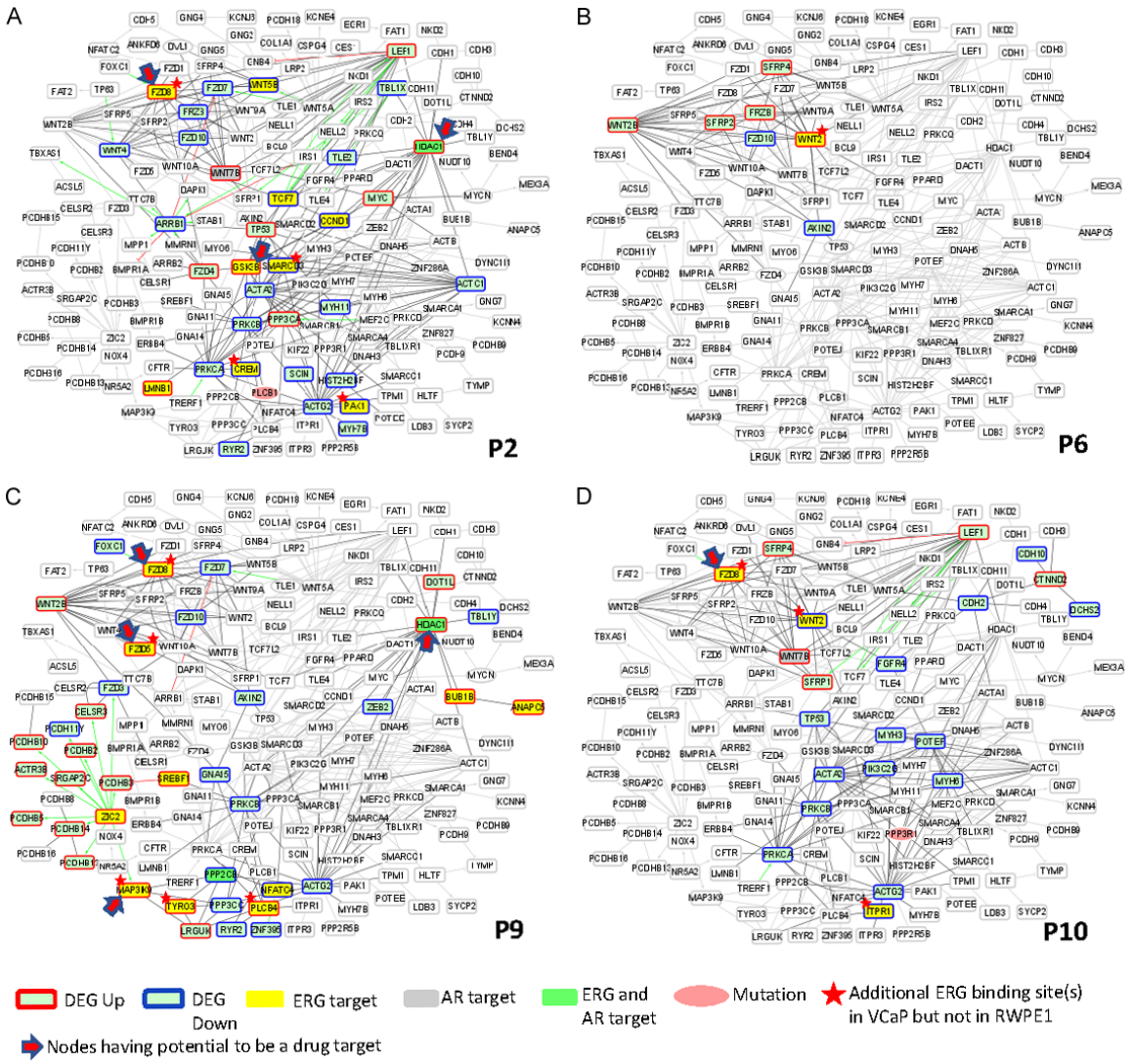


Figure 3. Patient-specific aberrations of the Wnt network are highly divergent. Global networks were established from differentially expressed genes (DEGs, >2-fold) in duplicate patient-matched tumor *vs.* normal samples by integrating the connectivities provided by STRING (protein-protein interaction database) and CellNet (transcription factor-target gene interactions); this integration yielded a 'master network' of DEGs for each patient, revealing the connectivities between deregulated genes. The DEG master network of each patient was complemented by the mutations of predicted high and moderate impact, and the components of the canonical and non-canonical Wnt pathways were extracted. DEGs and mutated genes are depicted in color for (A), P2, (B), P6, (C), P9 and (D), P10 in the background (grey nodes and connectivities) of all merged components of the Wnt pathways that are deregulated or mutated in all 15 patients (Supplementary Figure 4A). Genes for which approved drugs exist or are in clinical trials were identified in drug databases and are indicated by arrows. The corresponding deregulated networks of the other patients are shown in Supplementary Figure 4B-L. When known, connectivities are displayed as green (activation) or red (inhibition) lines; unknown connectivities and protein-protein interactions are displayed as grey lines. DEG specifics and mutations are color-coded as described below the figure.

sion may be an important component of the deregulated network.

Novel approach for in silico validation of deregulated transcription factor cistromes

Even though the above patient-specific DEG networks were derived from true RNA-seq duplicates and several mutations confirmed by Sanger sequencing, we sought additional evidence for validating the network. Given the personalized nature of our study, we excluded cell line or animal studies as proxy. Organoids could not be established either, as all prostates had been deep-frozen. However, we decided to exploit the knowledge existing in databases to

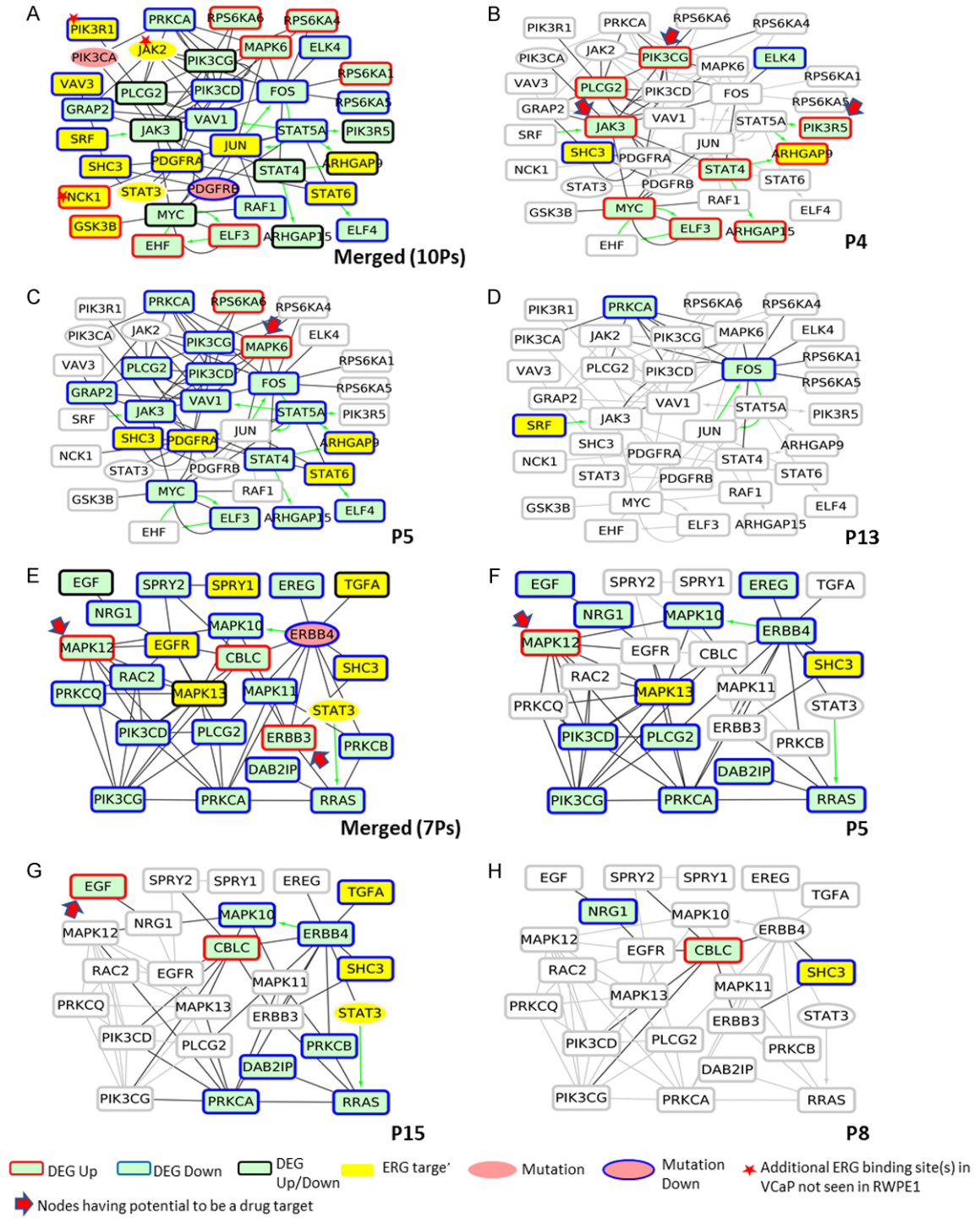


Figure 4. Divergent alterations of the PDGF and EGFR signaling networks in each patient. Ten patients exhibited serious aberrations in the PDGF and 7 patients in the EGFR signaling networks. (A) Merged network of alterations (DEG, mutation) in the PDGF and (E) EGFR networks. Using this merged network as background (grey nodes and connectivities) the aberrations in each individual patient are depicted in color. (B, C) Patients with heavily (P4, P5) or (D), minimally (P13) affected PDGF networks. (F, G) Patients with heavily (P5, P15) or (H) minimally (P8) affected EGFR networks.

provide external supportive evidence for the accuracy of the DEG networks. The rational was as follows: If a network contains upregulated TFs, most if not all of the cognate targets have

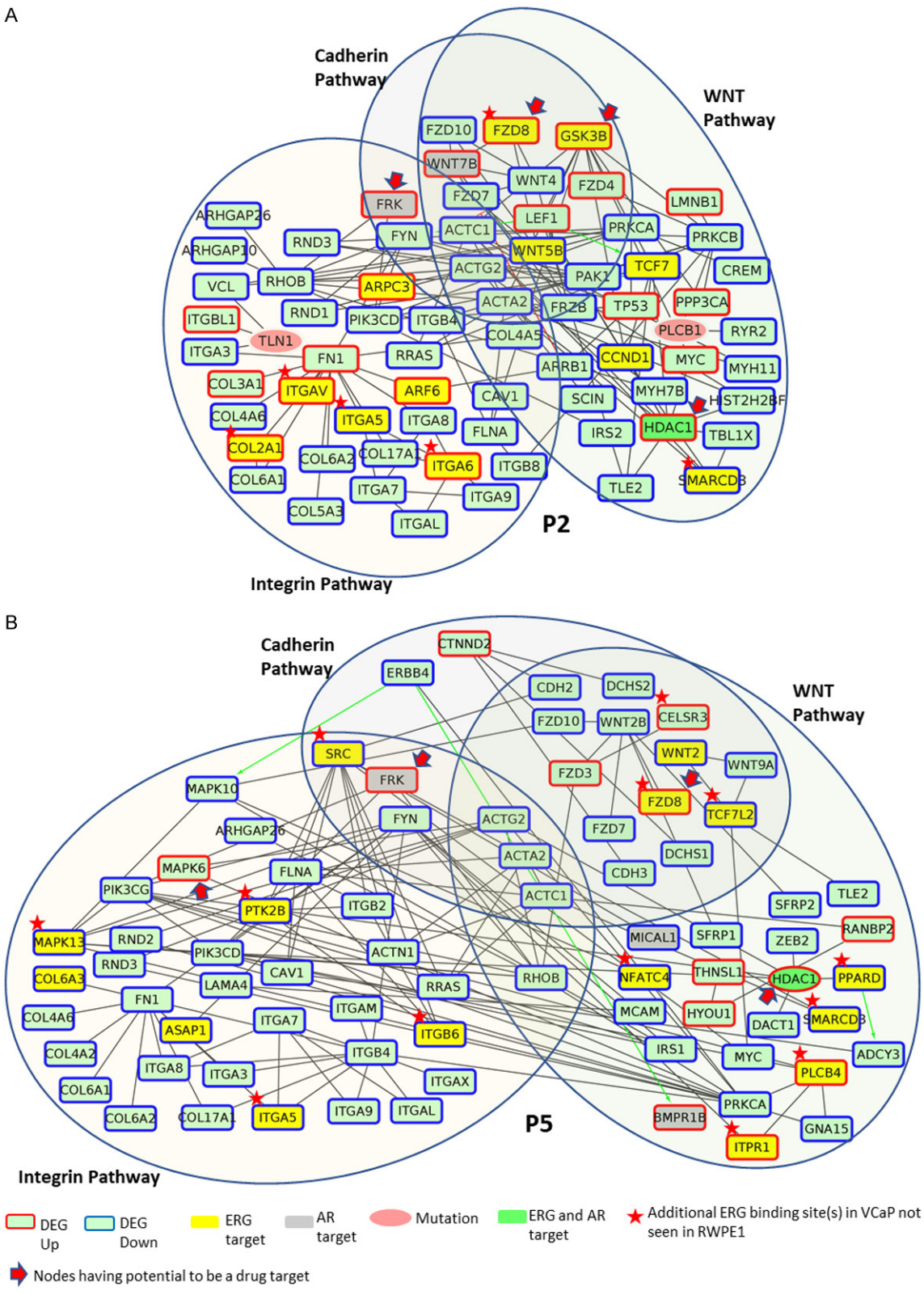


Figure 5. Affected signaling network crosstalk divergently between each other in each patient. (A) Illustration of the merged networks of the affected genes from the Wnt, Cadherin and Integrin signaling pathways in the prostate of P2 revealing that several of the DEGs are common to different pathways. (B) Illustration as in (A) but for P5. Color codes are displayed below the figure.

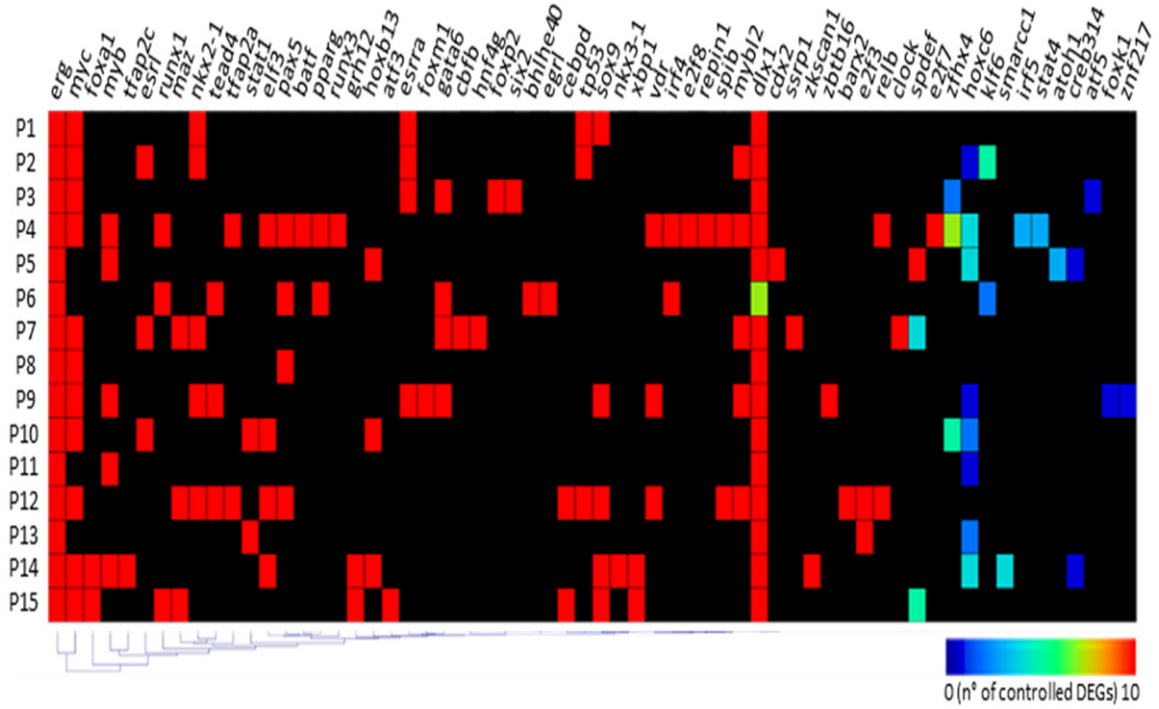


Figure 6. Upregulated TFs associated to DEGs identified for each prostate cancer patient (P1-P15). A large collection of qualified (QC quality A to C) public ChIP-seq datasets for TFs was used for identifying their binding sites (MACS peak calling, pval $<10^{-50}$). Each binding site has been annotated to its most proximal gene promoter (10 kb distance). Only TF-TG (target gene) associations for DEGs retrieved within the 15 patients have been retained. The matrix presents upregulated TFs per patient, the heatmap corresponds to the number of associated DEGs.

been already identified by ChIP-seq. As tumor cells may have TF targets beyond those of the cognate tissue, we collected high confidence targets ($P<10^{-50}$) from all tissues available in the qcGenomics database [41]. We then asked, if the personalized DEG network of a patient contained also the expected cognate TGs of a given TF. In case a significant number of TGs was detected, this provided additional evidence for the accuracy of the network (Supplementary File 4). Indeed, monitoring more than 60 deregulated TFs, the large majority of the corresponding experimentally validated TGs were co-deregulated (**Figure 6**, Supplementary Table 5).

In silico validation of deregulated ERG reveals ribosomal protein genes as frequent targets in addition to patient-specific ones

We tested the validation approach using the TF ERG which is overexpressed in all patient samples due to the TMPRSS2-ERG fusion. In P1, 462 genes were deregulated of which 150 corresponded to ERG TGs. Notably, these ERG target genes contained a strong cluster of ribosomal protein and translation regulatory genes, like *EEF1B2*, *UBA52* or *NCL* (**Figure 7A**). That these genes are *bona fide* ERG target genes was confirmed by qcGenomics profiling of ERG ChIP-seq data sets revealing ERG binding in VCaP (GSM2086309 to GSM2086314) but not in RPWE (GSM2195103, GSM2195106, GSM2195110) cells. While similar results concerning the ERG activation of ribosomal genes were seen for other samples (e.g., P8) with generally less ERG targets being deregulated, several cancers did not reveal such effects despite the overexpression of ERG (e.g., P10, P11, P13, P15; **Figure 7B**, Supplementary Figure 10) but in all cases a network of deregulated ERG target genes was noted. This variability may reflect patient-specific alterations of the chromatin landscape. Together these results show that our novel *in silico* approach is a valuable method to validate patient-specific deregulated TF-TG networks.

Several patient-specific overexpressed TFs deregulate their cognate cistromes

To test if the above approach is also valid for TFs other than ERG we established patient-

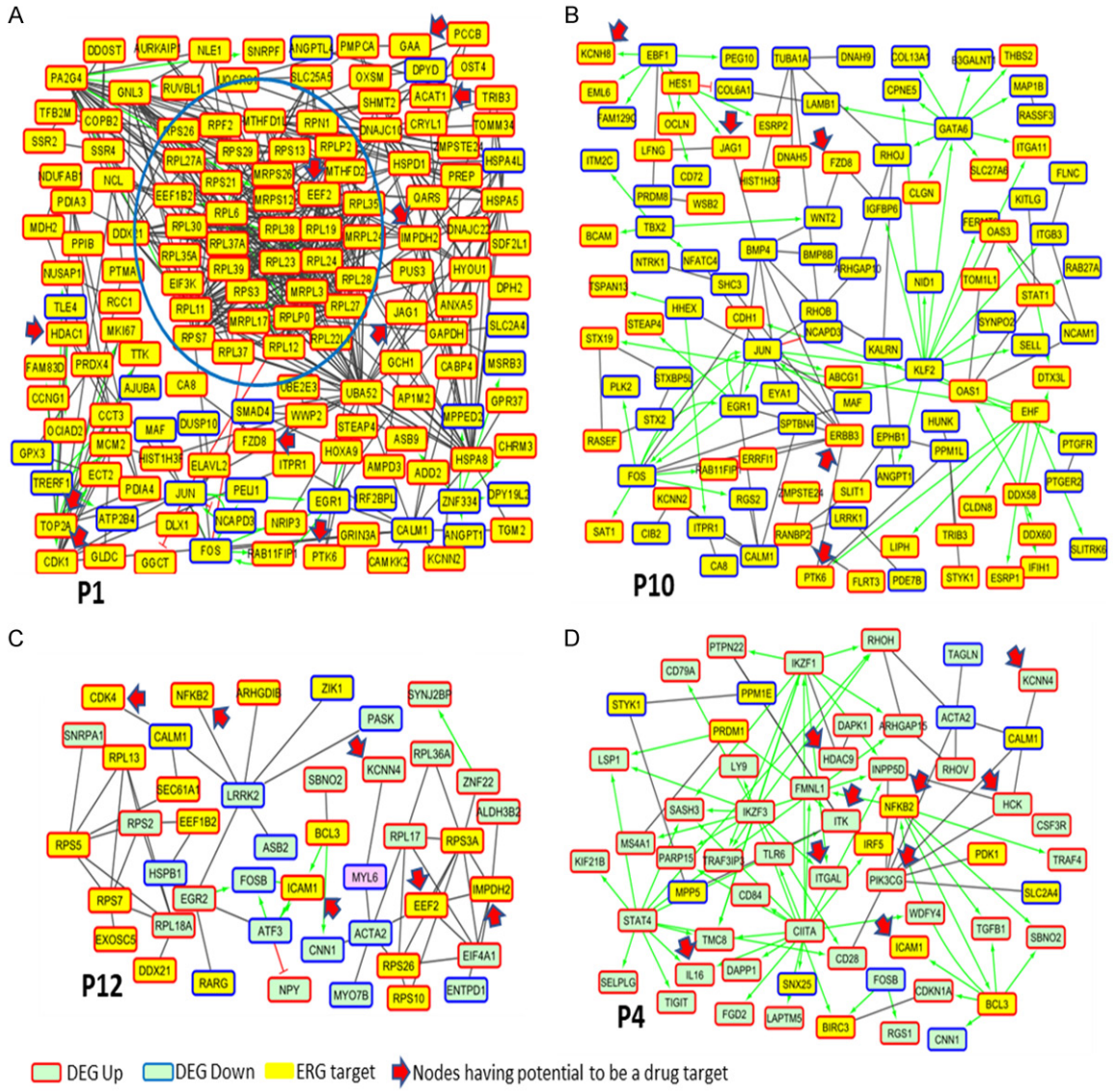


Figure 7. Patient-specific networks of upregulated transcription factor target genes are highly dissimilar. (A) Network of qcGenomics-predicted ERG target genes in P1 and (B) P10. ERG target genes were identified (see text) and the ERG-regulated gene networks were extracted from the master network for each patient. (C) Networks of predicted PAX5 target genes in P12 and (D) P4. Note the strong divergence between the ERG and PAX5 target gene networks between different patient-matched samples. When known, connectivities are displayed as green (activation) or red (inhibition) lines; unknown connectivities and protein-protein interactions are shown as black lines. DEGs and gene mutations are color-coded as described below the figure. The blue circle in (A) reveals tumor-selectively ERG-upregulated ribosomal and translation-associated genes in P1.

specific DEG TG networks for PAX5 (**Figure 7C**, **7D**) and MAZ (Supplementary Figure 11). In all cases the DEG networks are significantly populated by PAX5 and MAZ target genes. Note however, that the repertoire of target genes for a given TF can significantly differ from patient to patient, despite its common overexpression. Interestingly, MAZ can also bind to ribosomal protein genes. Indeed this is fully supported by the corresponding ChIP-seq data using qcGenomics (GSM93-5337, GSM1003613, GSM935272, GSM935-335), even though data are only available for K562, IMR90, Hela and HepG2 cells. The above integration of TF cistrome data further supports our overall notion that in individual tumors very different gene networks can be deregulated with different sets of overexpressed potentially druggable targets.

Discussion

Subsequent to the discoveries of oncogenes and tumor suppressors, and the concept of drivers and passengers of tumorigenesis [42, 43], the enormous progress in genome-wide sequencing together with a plethora of functional genomics applications has raised hopes that cancer genomics will rapidly reveal genome alterations causal to the disease and provide novel targets for therapy. However, like in human genetics, we face a scenario in which the origin of monogenic diseases has been largely deciphered, while the deregulated networks underlying multigenic diseases remain unknown. Indeed, despite several success stories, the anticipated rapid translation from cancer genomics to therapies did not occur. Rather, these studies revealed an ever-increasing complexity of multiple deregulated systems in tumors, intra-tumor and inter-patient heterogeneity and the incomplete understanding of the affected regulatory pathways operating within and between cells and tissues. This complexity is likely the major caveat for translating cancer genomics towards therapy. Moreover, the common approach of comparing hundreds to thousands of patients to identify individual targets may be conceptually problematic, as (i) new single denominators, in addition to those already discovered, may not exist and the deregulation/mutation of multiple interacting genes/pathways may be critically involved in the origin/evolution of the disease. Moreover, (ii) inter-tumoral/inter-individual variation will be disregarded by this approach, as well as (iii) the altered cross-talks between pathways. Finally, (iv) genes may exert distinct functions in different pathways/communication networks and (v) act as functionally divergent paralogues, like the multiple Wnt genes.

Previously, numerous studies were performed to unravel novel potential therapeutic targets using large-scale genomic [2, 44, 45] and integrative transcriptome-genome analyses [1, 8, 46, 47]. Multiple genome-wide association studies led to the identification of more than 160 disease-susceptible loci, most of which have unknown clinical implications [reviewed by 3]. A recent study involving large numbers of PrCa transcriptome profiles from 38 cohorts, developed a classification system based on pathway activation signals and presented a 37-gene signature which can classify PrCa into 3 subtypes [48]. More recently, the analysis of 18 recurrent DNA- and RNA-based genomic alterations, including androgen receptor variant expression and transcriptional output, and neuroendocrine expression signatures revealed *RB1* as the only gene associated with clinical outcome [4]. All these studies focused on common high-frequency targets. Somewhat disturbingly, each study revealed different (sets of) targets and a potential therapeutic application was postulated in several cases. In addition, as pointed out in a recent report that highlighted the role of low-frequency mutations in cancer progression [29], these approaches will miss low-frequency targets. That all the above-mentioned studies do not focus on individual patients builds strong case for a thorough analysis of mutations/aberrations at the individual level. In this respect, a recent network-based integrative study used genomic, transcriptomic and phosphoproteomic datasets to compare treatment-naive and metastatic PrCa and suggested personalized signatures in individual patients [49].

To provide insight into the various aspects of heterogeneity and with the aim of developing a pathway/network-centric rather than a gene-centric approach, we assessed the complexity of alteration occurring during tumorigenesis in the well-defined main class of prostate tumors with the TMPRSS2-ERG fusion [6] by comparing the tumor and its adjacent apparently normal tissue from the same untreated patients. Indeed, a recent report revealed that the use of paired tumor-normal samples improved mutation identification and decreased false-positive rates [50].

This patient-centered network analysis revealed highly divergent patient-specific deregulated and mutated genomic landscapes. All 15 patients except P6 revealed aberrations (mutation, expression) of one or several components of the WNT, Cadherin and Integrin pathways with large differences between deregulated pathway components in each of the patients, thus advocating the need for a patient-centered analysis. As previously pointed out, targeting the WNT pathway is challenging due to its complex nature driving diverse biological processes and cross-talk with multiple other pathways [51]. Therefore, a thorough understanding

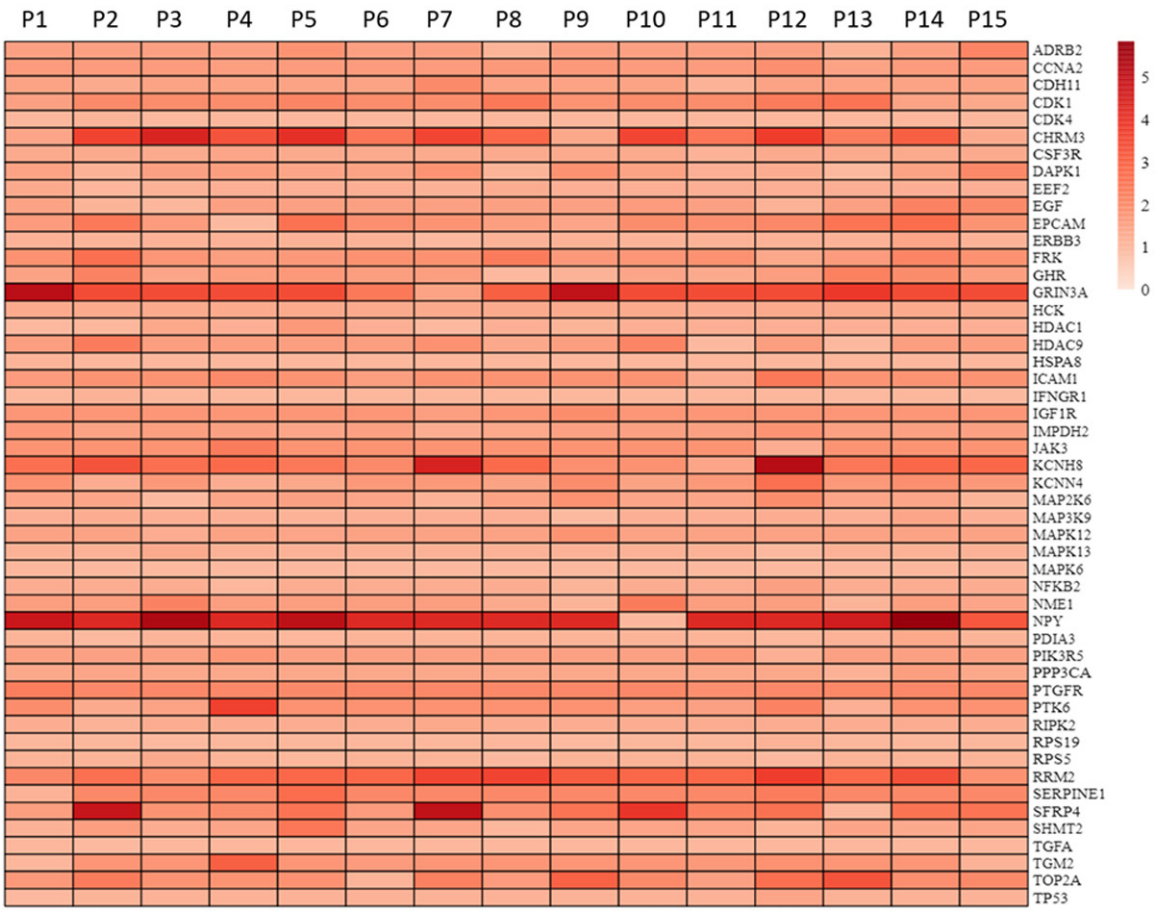


Figure 8. Heat map revealing the extent of upregulation of genes whose products correspond to validated drug targets. As discussed in the text, only gene products for which drugs have been approved or have been/are enrolled in clinical trials are considered, irrespective of the disease for which they have been/are being developed. In the network figures, these genes are marked with a red asterisk. Color codes represent log2-fold changes. P1 to P15, patients 1 to 15. Color codes represent log2 fold change.

about each component of these interacting pathways and key nodes as potential therapeutic target, based on a systems medicine approach [52] with patient-specific networks, is the ideal way to move forward towards targeted therapy. Indeed, patients who received a personalized therapy adapted to their genomic aberrations, as recommended by a multidisciplinary molecular tumor board, had improved oncological outcomes including survival [53]. Our patient-specific network analysis facilitated the identification of key nodes involved in the cross-talk between different deregulated pathways in the same patient. In this respect, we have identified a significant number of genes which correspond to validated drug targets (**Figure 8**). The extent of these deregulated genes varies largely between patients and several highly upregulated genes occur in individual patients (e.g. P1, P7, P9, P12, P14). This information will be useful to decide about combinatorial therapies for individual patients by targeting key nodes of different deregulated pathways. For example, in P2, the upregulated direct ERG targets *ITGAV* and *ITGA6* (Integrin pathway) along with *FZD8* and *GSK3B* (Wnt and Cadherin pathway) could be used as potential drug targets for combinatorial therapy. Similarly, ETS family members ERG and ETV1, overexpressed due to the fusion with TMPRSS2, were reported to directly suppress CHK1 promoting tumorigenesis bypassing DNA damage response [40]. Our patient-specific network analysis in the present study revealed several deregulated genes that acquired additional ERG binding in their respective promoters. These direct targets of ERG, which are involved in different cooperating pathways, correspond

to potential therapeutic targets. Based on the above results and reflections, we propose to develop therapeutic options in the context of a personalized integrative functional genomics analysis rather than trying to identify common single targets from the analysis of large numbers of patients.

Around 97 percent of potential drugs undergoing clinical trials fail to get FDA approval [54]. The off-target toxicity of cancer drugs undergoing clinical trials [55] suggests the need for more robust genetic analysis when predicting the potential drug target. Multilayered patient-specific network analysis will be useful to identify not only putative drug targets but also predict potential off-target effects resulting from pathway cross-talks. Apart from identifying targets of potential therapeutic use, there is strong need for development of novel drugs which can target components of complex pathways like WNT signaling. In this respect, patient-derived organoid cultures [56] which recapitulate the diversity of primary tumors may facilitate screening of novel molecules against these putative therapeutic targets.

In addition, there is a growing importance of single-cell functional genomics done with circulating tumor cells for diagnosis. Recent study involving pan-cancer analysis of chromatin accessibility revealed novel protein-DNA interactions in primary cancer tissues [57]. Integrating chromatin accessibility data from individual patients can identify cancer-specific novel regulatory connections which can be used as potential drug target. Ultimately, additional dimensions like RNA regulators, such as the newly described circular RNAs [58, 59], as well as metabolomics changes, may be integrated in this analysis to reveal what communication networks are at the origin, maintenance and progression of the disease and which regulatory circuits can be modulated for therapeutic purposes, including escape from resistance to therapy.

Acknowledgements

We dedicate this work to the memory of Dr. Catherine Mazerolles who sadly passed away when this study was gaining full speed. We thank all members of our labs for discussions and suggestions. These studies were supported by funds from the Plan Cancer, AVIESAN-ITMO Cancer to HG and LV, the Ligue National Contre le Cancer (HG; Equipe Labellisée); and the Institut National du Cancer (INCa) to HG and LV. Support of the Agence Nationale de la Recherche (ANRT-07-PCVI-0031-01, ANR-10-LABX-0030-INRT and ANR-10-IDEX-0002-02) is acknowledged. Samples from non-treated patients were obtained after written informed consent in accordance with the Declaration of Helsinki and stored at the « CRB Cancer des Hôpitaux de Toulouse (BB-0033-00014) » collection. According to the French law, the CRB Cancer collection has been declared to the Ministry of Higher Education and Research (DC-2008-463) and obtained a transfer agreement (AC-2013-1955) after approbation by ethical committees. Clinical and biological annotations of the samples have been declared to the CNIL (Commission Nationale de l'Informatique et des Libertés).

Disclosure of conflict of interest

None.

Address correspondence to: Hinrich Gronemeyer, Institut de Génétique et de Biologie Moléculaire et Cellulaire (IGBMC), Department of Functional Genomics and Cancer, Illkirch, France. E-mail: hg@igbmc.fr; Laurence Vandel, Université Clermont Auvergne, CNRS, Inserm, GReD, Clermont-Ferrand, France. E-mail: laurence.vandel@uca.fr

References

[1] Robinson D, Van Allen EM, Wu YM, Schultz N, Lonigro RJ, Mosquera JM, Montgomery B, Taplin ME, Pritchard CC, Attard G, Beltran H, Abida W, Bradley RK, Vinson J, Cao X, Vats P, Kunju LP, Hussain M, Feng FY, Tomlins SA, Cooney KA, Smith DC, Brennan C, Siddiqui J, Mehra R, Chen Y, Rathkopf DE, Morris MJ, Solomon SB, Durack JC, Reuter VE, Gopalan A, Gao J, Loda M, Lis RT, Bowden M, Balk SP, Gaviola G, Sougnez C, Gupta M, Yu EY, Mostaghel EA, Cheng HH, Mulcahy H, True LD, Plymate SR, Dvinge H, Ferraldeschi R, Flohr P, Miranda S, Zafeiriou Z, Tunariu N, Mateo J, Perez-Lopez R, Demichelis F, Robinson BD, Schiffman M, Nanus DM, Tagawa ST, Sigaras A, Eng KW, Elemento O, Sboner A, Heath EI, Scher HI, Pienta KJ, Kantoff P, de Bono JS, Rubin MA, Nelson PS, Garraway LA, Sawyers CL and Chinnaiyan AM. Integrative clinical genomics of advanced prostate cancer. Cell 2015; 161: 1215-1228.

[2] Baca SC, Prandi D, Lawrence MS, Mosquera JM, Romanel A, Drier Y, Park K, Kitabayashi N, MacDonald TY, Ghandi M, Van Allen E, Kryukov GV, Sboner A, Theurillat JP, Soong TD, Nicker-

son E, Auclair D, Tewari A, Beltran H, Onofrio RC, Boysen G, Guiducci C, Barbieri CE, Cibulskis K, Sivachenko A, Carter SL, Saksena G, Voet D, Ramos AH, Winckler W, Cipicchio M, Ardlie K, Kantoff PW, Berger MF, Gabriel SB, Golub TR, Meyerson M, Lander ES, Elemento O, Getz G, Demichelis F, Rubin MA and Garraway LA. Punctuated evolution of prostate cancer genomes. Cell 2013; 153: 666-677.

[3] Farashi S, Kryza T, Clements J and Batra J. Post-GWAS in prostate cancer: from genetic association to biological contribution. Nat Rev Cancer 2019; 19: 46-59.

[4] Abida W, Cyrta J, Heller G, Prandi D, Armenia J, Coleman I, Cieslik M, Benelli M, Robinson D, Van Allen EM, Sboner A, Fedrizzi T, Mosquera JM, Robinson BD, De Sarkar N, Kunju LP, Tomlins S, Wu YM, Nava Rodrigues D, Loda M, Gopalan A, Reuter VE, Pritchard CC, Mateo J, Bianchini D, Miranda S, Carreira S, Rescigno P, Filipenko J, Vinson J, Montgomery RB, Beltran H, Heath EI, Scher HI, Kantoff PW, Taplin ME, Schultz N, deBono JS, Demichelis F, Nelson PS, Rubin MA, Chinnaiyan AM and Sawyers CL. Genomic correlates of clinical outcome in advanced prostate cancer. Proc Natl Acad Sci U S A 2019; 116: 11428-11436.

[5] Cancer Genome Atlas Research Network. The molecular taxonomy of primary prostate cancer. Cell 2015; 163: 1011-1025.

[6] Tomlins SA, Rhodes DR, Perner S, Dhanasekaran SM, Mehra R, Sun XW, Varambally S, Cao X, Tchinda J, Kuefer R, Lee C, Montie JE, Shah RB, Pienta KJ, Rubin MA and Chinnaiyan AM. Recurrent fusion of TMPRSS2 and ETS transcription factor genes in prostate cancer. Science 2005; 310: 644-648.

[7] Chen CD, Welsbie DS, Tran C, Baek SH, Chen R, Vessella R, Rosenfeld MG and Sawyers CL. Molecular determinants of resistance to antiandrogen therapy. Nat Med 2004; 10: 33-39.

[8] Haffner MC, Mosbruger T, Esopi DM, Fedor H, Heaphy CM, Walker DA, Adejola N, Gürel M, Hicks J, Meeker AK, Halushka MK, Simons JW, Isaacs WB, De Marzo AM, Nelson WG and Yegnasubramanian S. Tracking the clonal origin of lethal prostate cancer. J Clin Invest 2013; 123: 4918-4922.

[9] Yi S, Lin S, Li Y, Zhao W, Mills GB and Sahni N. Functional variomics and network perturbation: connecting genotype to phenotype in cancer. Nat Rev Genet 2017; 18: 395-410.

[10] McKenna A, Hanna M, Banks E, Sivachenko A, Cibulskis K, Kernytsky A, Garimella K, Altshuler D, Gabriel S, Daly M and DePristo MA. The genome analysis toolkit: a mapreduce framework for analyzing next-generation DNA sequencing data. Genome Res 2010; 20: 1297-1303.

[11] Koster J and Rahmann S. Snakemake--a scalable bioinformatics workflow engine. Bioinformatics 2012; 28: 2520-2522.

[12] Garofalo M, Di Leva G, Romano G, Nuovo G, Suh SS, Ngankeu A, Taccioli C, Pichiorri F, Alder H, Secchiero P, Gasparini P, Gonelli A, Costinean S, Acunzo M, Condorelli G and Croce CM. miR-221&222 regulate TRAIL resistance and enhance tumorigenicity through PTEN and TIMP3 downregulation. Cancer Cell 2009; 16: 498-509.

[13] Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, Marth G, Abecasis G and Durbin R; Genome Project Data Processing Subgroup. The sequence alignment/map format and SAMtools. Bioinformatics 2009; 25: 2078-2079.

[14] Cibulskis K, Lawrence MS, Carter SL, Sivachenko A, Jaffe D, Sougnez C, Gabriel S, Meyerson M, Lander ES and Getz G. Sensitive detection of somatic point mutations in impure and heterogeneous cancer samples. Nat Biotechnol 2013; 31: 213-219.

[15] Cingolani P, Platts A, Wang le L, Coon M, Nguyen T, Wang L, Land SJ, Lu X and Ruden DM. A program for annotating and predicting the effects of single nucleotide polymorphisms, SnpEff: SNPs in the genome of Drosophila melanogaster strain w1118; iso-2; iso-3. Fly (Austin) 2012; 6: 80-92.

[16] Kim D, Langmead B and Salzberg SL. HISAT: a fast spliced aligner with low memory requirements. Nat Methods 2015; 12: 357-360.

[17] Lawrence M, Huber W, Pages H, Aboyoun P, Carlson M, Gentleman R, Morgan MT and Carey VJ. Software for computing and annotating genomic ranges. PLoS Comput Biol 2013; 9: e1003118.

[18] Love MI, Huber W and Anders S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biol 2014; 15: 550.

[19] Tabas-Madrid D, Nogales-Cadenas R and Pascual-Montano A. GeneCodis3: a non-redundant and modular enrichment analysis tool for functional genomics. Nucleic Acids Res 2012; 40: W478-483.

[20] Szklarczyk D, Morris JH, Cook H, Kuhn M, Wyder S, Simonovic M, Santos A, Doncheva NT, Roth A, Bork P, Jensen LJ and von Mering C. The STRING database in 2017: quality-controlled protein-protein association networks, made broadly accessible. Nucleic Acids Res 2017; 45: D362-D368.

[21] Cahan P, Li H, Morris SA, Lummertz da Rocha E, Daley GQ and Collins JJ. CellNet: network biology applied to stem cell engineering. Cell 2014; 158: 903-915.

[22] Shannon P, Markiel A, Ozier O, Baliga NS, Wang JT, Ramage D, Amin N, Schwikowski B and Ideker T. Cytoscape: a software environment for integrated models of biomolecular interaction networks. Genome Res 2003; 13: 2498-2504.

[23] Szklarczyk D, Gable AL, Lyon D, Junge A, Wyder S, Huerta-Cepas J, Simonovic M, Doncheva NT, Morris JH, Bork P, Jensen LJ and von Mering C. STRING v11: protein-protein association networks with increased coverage, supporting functional discovery in genome-wide experimental datasets. Nucleic Acids Res 2019; 47: D607-D613.

[24] Kypta RM and Waxman J. Wnt/β-catenin signalling in prostate cancer. Nat Rev Urol 2012; 9: 418-428.

[25] Jalota A, Singh K, Pavithra L, Kaul-Ghanekar R, Jameel S and Chattopadhyay S. Tumor suppressor SMAR1 activates and stabilizes p53 through its arginine-serine-rich motif. J Biol Chem 2005; 280: 16019-16029.

[26] Ishii H, Vecchione A, Murakumo Y, Baldassarre G, Numata S, Trapasso F, Alder H, Baffa R and Croce CM. FEZ1/LZTS1 gene at 8p22 suppresses cancer cell growth and regulates mitosis. Proc Natl Acad Sci U S A 2001; 98: 10374-10379.

[27] Vecchione A, Baldassarre G, Ishii H, Nicoloso MS, Belletti B, Petrocca F, Zanesi N, Fong LY, Battista S, Guarnieri D, Baffa R, Alder H, Farber JL, Donovan PJ and Croce CM. Fez1/Lzts1 absence impairs Cdk1/Cdc25C interaction during mitosis and predisposes mice to cancer development. Cancer Cell 2007; 11: 275-289.

[28] Shen M, Jiang YZ, Wei Y, Ell B, Sheng X, Esposito M, Kang J, Hang X, Zheng H, Rowicki M, Zhang L, Shih WJ, Celià-Terrassa T, Liu Y, Cristea I, Shao ZM and Kang Y. Tinagl1 suppresses triple-negative breast cancer progression and metastasis by simultaneously inhibiting integrin/FAK and EGFR signaling. Cancer Cell 2019; 35: 64-80, e7.

[29] Armenia J, Wankowicz SAM, Liu D, Gao J, Kundra R, Reznik E, Chatila WK, Chakravarty D, Han GC, Coleman I, Montgomery B, Pritchard C, Morrissey C, Barbieri CE, Beltran H, Sboner A, Zafeiriou Z, Miranda S, Bielski CM, Penson AV, Tolonen C, Huang FW, Robinson D, Wu YM, Lonigro R, Garraway LA, Demichelis F, Kantoff PW, Taplin ME, Abida W, Taylor BS, Scher HI, Nelson PS, de Bono JS, Rubin MA, Sawyers CL and Chinnaiyan AM; PCF/SU2C International Prostate Cancer Dream Team, Schultz N, Van Allen EM. The long tail of oncogenic drivers in prostate cancer. Nat Genet 2018; 50: 645-651.

[30] Jin Y, Qu S, Tesikova M, Wang L, Kristian A, Mælandsmo GM, Kong H, Zhang T, Jerónimo C, Teixeira MR, Yuca E, Tekedereli I, Gorgulu K, Alpay N, Sood AK, Lopez-Berestein G, Danielsen HE, Ozpolat B and Saatcioglu F. Molecular circuit involving KLK4 integrates androgen and mTOR signaling in prostate cancer. Proc Natl Acad Sci U S A 2013; 110: E2572-2581.

[31] Ramsay AJ, Dong Y, Hunt ML, Linn M, Samaratunga H, Clements JA and Hooper JD. Kallikrein-related peptidase 4 (KLK4) initiates intracellular signaling via protease-activated receptors (PARs). KLK4 and PAR-2 are co-expressed during prostate cancer progression. J Biol Chem 2008; 283: 12293-12304.

[32] Adriaens C, Standaert L, Barra J, Latil M, Verfaillie A, Kalev P, Boeckx B, Wijnhoven PW, Radaelli E, Vermi W, Leucci E, Lapouge G, Beck B, van den Oord J, Nakagawa S, Hirose T, Sablina AA, Lambrechts D, Aerts S, Blanpain C and Marine JC. p53 induces formation of NEAT1 lncRNA-containing paraspeckles that modulate replication stress response and chemosensitivity. Nat Med 2016; 22: 861-868.

[33] Mello SS, Sinow C, Raj N, Mazur PK, Bieging-Rolett K, Broz DK, Imam JFC, Vogel H, Wood LD, Sage J, Hirose T, Nakagawa S, Rinn J and Attardi LD. Neat1 is a p53-inducible lincRNA essential for transformation suppression. Genes Dev 2017; 31: 1095-1108.

[34] Zhang Y, Pitchiaya S, Cieślik M, Niknafs YS, Tien JC, Hosono Y, Iyer MK, Yazdani S, Subramaniam S, Shukla SK, Jiang X, Wang L, Liu TY, Uhl M, Gawronski AR, Qiao Y, Xiao L, Dhanasekaran SM, Juckette KM, Kunju LP, Cao X, Patel U, Batish M, Shukla GC, Paulsen MT, Ljungman M, Jiang H, Mehra R, Backofen R, Sahinalp CS, Freier SM, Watt AT, Guo S, Wei JT, Feng FY, Malik R and Chinnaiyan AM. Analysis of the androgen receptor-regulated lncRNA landscape identifies a role for ARLNC1 in prostate cancer progression. Nat Genet 2018; 50: 814-824.

[35] Malek R, Gajula RP, Williams RD, Nghiem B, Simons BW, Nugent K, Wang H, Taparra K, Lemtiri-Chlieh G, Yoon AR, True L, An SS, De-Weese TL, Ross AE, Schaeffer EM, Pienta KJ, Hurley PJ, Morrissey C and Tran PT. TWIST1-WDR5-Hottip regulates Hoxa9 chromatin to facilitate prostate cancer metastasis. Cancer Res 2017; 77: 3181-3193.

[36] Malik R, Khan AP, Asangani IA, Cieślik M, Prensner JR, Wang X, Iyer MK, Jiang X, Borkin D, Escara-Wilke J, Stender R, Wu YM, Niknafs YS, Jing X, Qiao Y, Palanisamy N, Kunju LP, Krishnamurthy PM, Yocum AK, Mellacheruvu D, Nesvizhskii AI, Cao X, Dhanasekaran SM, Feng FY, Grembecka J, Cierpicki T and Chinnaiyan

AM. Targeting the MLL complex in castration-resistant prostate cancer. Nat Med 2015; 21: 344-352.
[37] Rupaimoole R and Slack FJ. MicroRNA therapeutics: towards a new era for the management of cancer and other diseases. Nat Rev Drug Discov 2017; 16: 203-222.
[38] Mi H, Muruganujan A, Ebert D, Huang X and Thomas PD. PANTHER version 14: more genomes, a new PANTHER GO-slim and improvements in enrichment analysis tools. Nucleic Acids Res 2019; 47: D419-D426.
[39] Clara JA, Monge C, Yang Y and Takebe N. Targeting signalling pathways and the immune microenvironment of cancer stem cells - a clinical update. Nat Rev Clin Oncol 2020; 17: 204-232.
[40] Lunardi A, Varmeh S, Chen M, Taulli R, Guarnerio J, Ala U, Seitzer N, Ishikawa T, Carver BS, Hobbs RM, Quarantotti V, Ng C, Berger AH, Nardella C, Poliseno L, Montironi R, Castillo-Martin M, Cordon-Cardo C, Signoretti S and Pandolfi PP. Suppression of CHK1 by ETS family members promotes DNA damage response bypass and tumorigenesis. Cancer Discov 2015; 5: 550-563.
[41] Blum M, Cholley PE, Malysheva V, Nicaise S, Moehlin J, Gronemeyer H and Mendoza-Parra MA. A comprehensive resource for retrieving, visualizing, and integrating functional genomics data. Life Science Alliance 2019; 3: e201900546.
[42] Bozic I, Antal T, Ohtsuki H, Carter H, Kim D, Chen S, Karchin R, Kinzler KW, Vogelstein B and Nowak MA. Accumulation of driver and passenger mutations during tumor progression. Proc Natl Acad Sci U S A 2010; 107: 18545-18550.
[43] Haber DA and Settleman J. Cancer: drivers and passengers. Nature 2007; 446: 145-146.
[44] Barbieri CE, Baca SC, Lawrence MS, Demichelis F, Blattner M, Theurillat JP, White TA, Stojanov P, Van Allen E, Stransky N, Nickerson E, Chae SS, Boysen G, Auclair D, Onofrio RC, Park K, Kitabayashi N, MacDonald TY, Sheikh K, Vuong T, Guiducci C, Cibulskis K, Sivachenko A, Carter SL, Saksena G, Voet D, Hussain WM, Ramos AH, Winckler W, Redman MC, Ardlie K, Tewari AK, Mosquera JM, Rupp N, Wild PJ, Moch H, Morrissey C, Nelson PS, Kantoff PW, Gabriel SB, Golub TR, Meyerson M, Lander ES, Getz G, Rubin MA and Garraway LA. Exome sequencing identifies recurrent SPOP, FOXA1 and MED12 mutations in prostate cancer. Nat Genet 2012; 44: 685-689.
[45] Grasso CS, Wu YM, Robinson DR, Cao X, Dhanasekaran SM, Khan AP, Quist MJ, Jing X, Lonigro RJ, Brenner JC, Asangani IA, Ateeq B, Chun SY, Siddiqui J, Sam L, Anstett M, Mehra R, Prensner JR, Palanisamy N, Ryslik GA, Vandin F, Raphael BJ, Kunju LP, Rhodes DR, Pienta KJ, Chinnaiyan AM and Tomlins SA. The mutational landscape of lethal castration-resistant prostate cancer. Nature 2012; 487: 239-243.
[46] Kron KJ, Murison A, Zhou S, Huang V, Yamaguchi TN, Shiah YJ, Fraser M, van der Kwast T, Boutros PC, Bristow RG and Lupien M. TMPRSS2-ERG fusion co-opts master transcription factors and activates NOTCH signaling in primary prostate cancer. Nat Genet 2017; 49: 1336-1345.
[47] Mancuso N, Gayther S, Gusev A, Zheng W, Penney KL, Kote-Jarai Z, Eeles R, Freedman M, Haiman C, Pasaniuc B and consortium P. Large-scale transcriptome-wide association study identifies new prostate cancer risk regions. Nat Commun 2018; 9: 4079.
[48] You S, Knudsen BS, Erho N, Alshalalfa M, Takhar M, Al-Deen Ashab H, Davicioni E, Karnes RJ, Klein EA, Den RB, Ross AE, Schaeffer EM, Garraway IP, Kim J and Freeman MR. Integrated classification of prostate cancer reveals a novel luminal subtype with poor outcome. Cancer Res 2016; 76: 4948-4958.
[49] Drake JM, Paull EO, Graham NA, Lee JK, Smith BA, Titz B, Stoyanova T, Faltermeier CM, Uzunangelov V, Carlin DE, Fleming DT, Wong CK, Newton Y, Sudha S, Vashisht AA, Huang J, Wohlschlegel JA, Graeber TG, Witte ON and Stuart JM. Phosphoproteome integration reveals patient-specific networks in prostate cancer. Cell 2016; 166: 1041-1054.
[50] Beaubier N, Bontrager M, Huether R, Igartua C, Lau D, Tell R, Bobe AM, Bush S, Chang AL, Hoskinson DC, Khan AA, Kudalkar E, Leibowitz BD, Lozachmeur A, Michuda J, Parsons J, Perera JF, Salahudeen A, Shah KP, Taxter T, Zhu W and White KP. Integrated genomic profiling expands clinical options for patients with cancer. Nat Biotechnol 2019; 37: 1351-1360.
[51] Kahn M. Can we safely target the WNT pathway? Nat Rev Drug Discov 2014; 13: 513-532.
[52] Sonawane AR, Weiss ST, Glass K and Sharma A. Network medicine in the age of biomedical big data. Front Genet 2019; 10: 294.
[53] Kato S, Kim KH, Lim HJ, Boichard A, Nikanjam M, Weihe E, Kuo DJ, Eskander RN, Goodman A, Galanina N, Fanta PT, Schwab RB, Shatsky R, Plaxe SC, Sharabi A, Stites E, Adashek JJ, Okamura R, Lee S, Lippman SM, Sicklick JK and Kurzrock R. Real-world data from a molecular tumor board demonstrates improved outcomes with a precision N-of-One strategy. Nat Commun 2020; 11: 4965.
[54] Wong CH, Siah KW and Lo AW. Estimation of clinical trial success rates and related parameters (vol 20, pg 273, 2019). Biostatistics 2019; 20: 366-366.

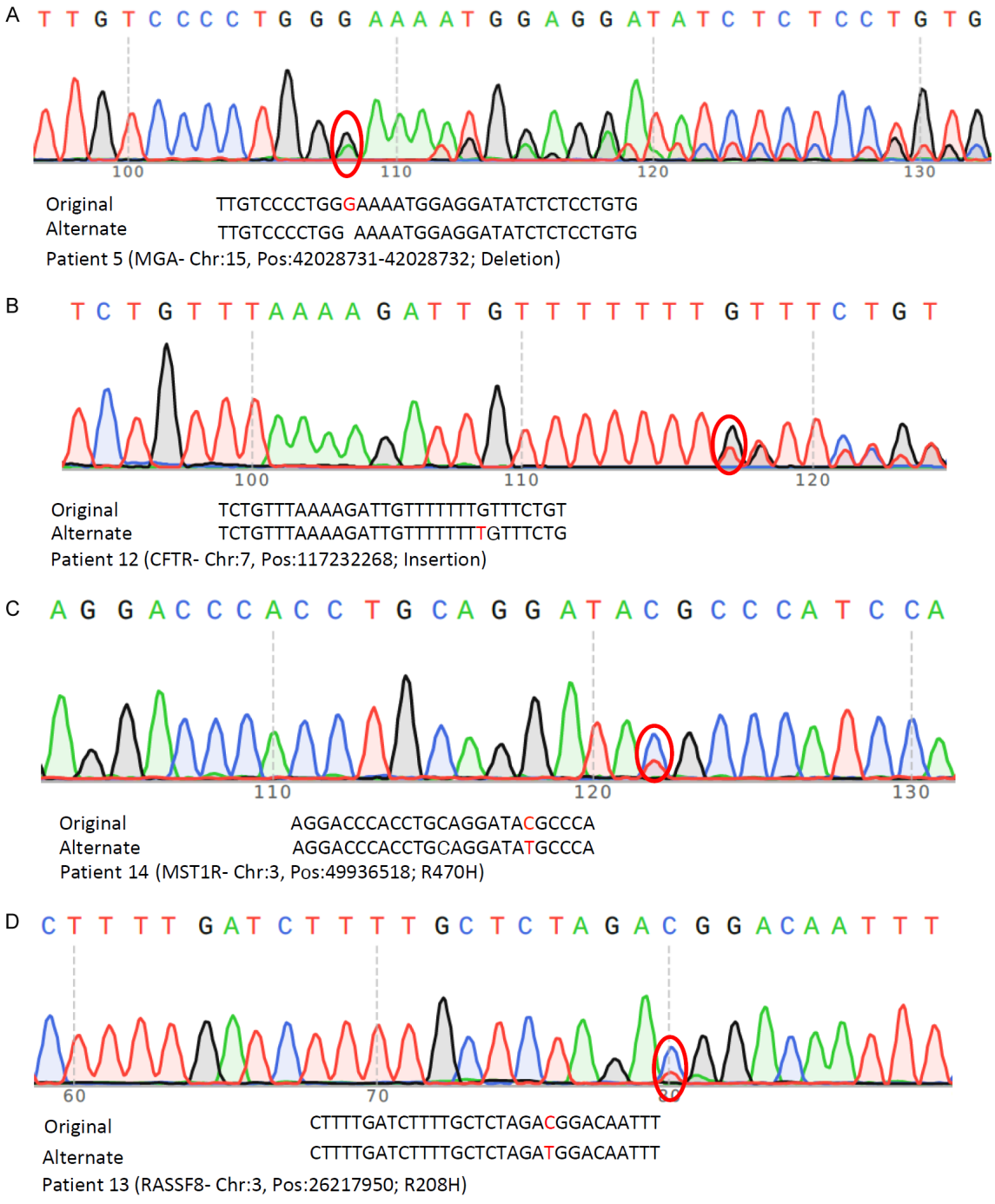
[55] Lin A, Giuliano CJ, Palladino A, John KM, Abramowicz C, Yuan ML, Sausville EL, Lukow DA, Liu L, Chait AR, Galluzzo ZC, Tucker C and Sheltzer JM. Off-target toxicity is a common mechanism of action of cancer drugs undergoing clinical trials. Sci Transl Med 2019; 11: eaaw8412.

[56] Gao D and Chen Y. Organoid development in cancer genome discovery. Curr Opin Genet Dev 2015; 30: 42-48.

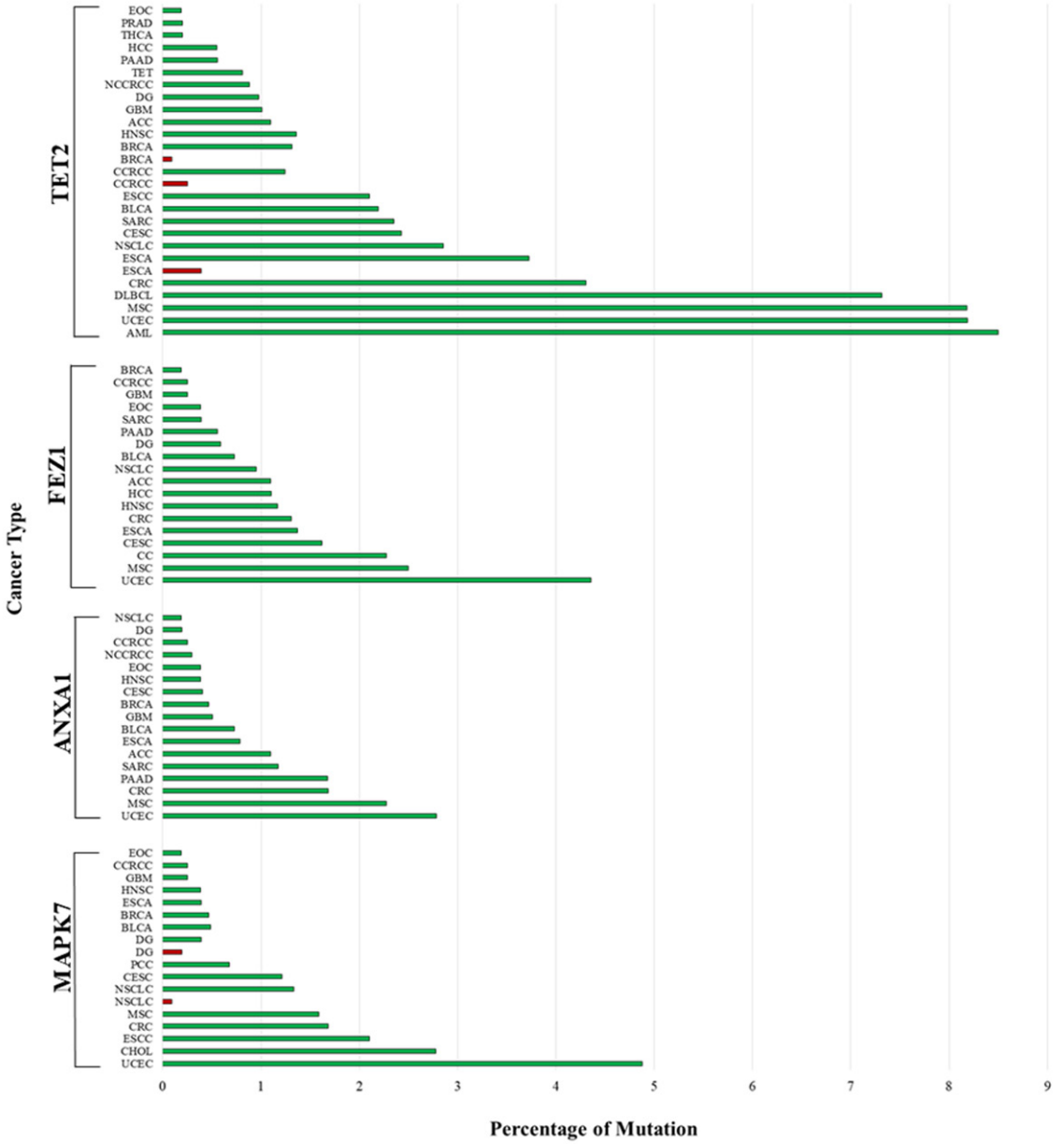
[57] Corces MR, Granja JM, Shams S, Louie BH, Seoane JA, Zhou W, Silva TC, Groeneveld C, Wong CK, Cho SW, Satpathy AT, Mumbach MR, Hoadley KA, Robertson AG, Sheffield NC, Felau I, Castro MAA, Berman BP, Staudt LM, Zenklusen JC, Laird PW and Curtis C; Cancer Genome Atlas Analysis Network; Greenleaf WJ and Chang HY. The chromatin accessibility landscape of primary human cancers. Science 2018; 362: eaav1898.

[58] Chen S, Huang V, Xu X, Livingstone J, Soares F, Jeon J, Zeng Y, Hua JT, Petricca J, Guo H, Wang M, Yousif F, Zhang Y, Donmez N, Ahmed M, Volik S, Lapuk A, Chua MLK, Heisler LE, Foucal A, Fox NS, Fraser M, Bhandari V, Shiah YJ, Guan J, Li J, Orain M, Picard V, Hovington H, Bergeron A, Lacombe L, Fradet Y, Tetu B, Liu S, Feng F, Wu X, Shao YW, Komor MA, Sahinalp C, Collins C, Hoogstrate Y, de Jong M, Fijneman RJA, Fei T, Jenster G, van der Kwast T, Bristow RG, Boutros PC and He HH. Widespread and functional RNA circularization in localized prostate cancer. Cell 2019; 176: 831-843, e22.

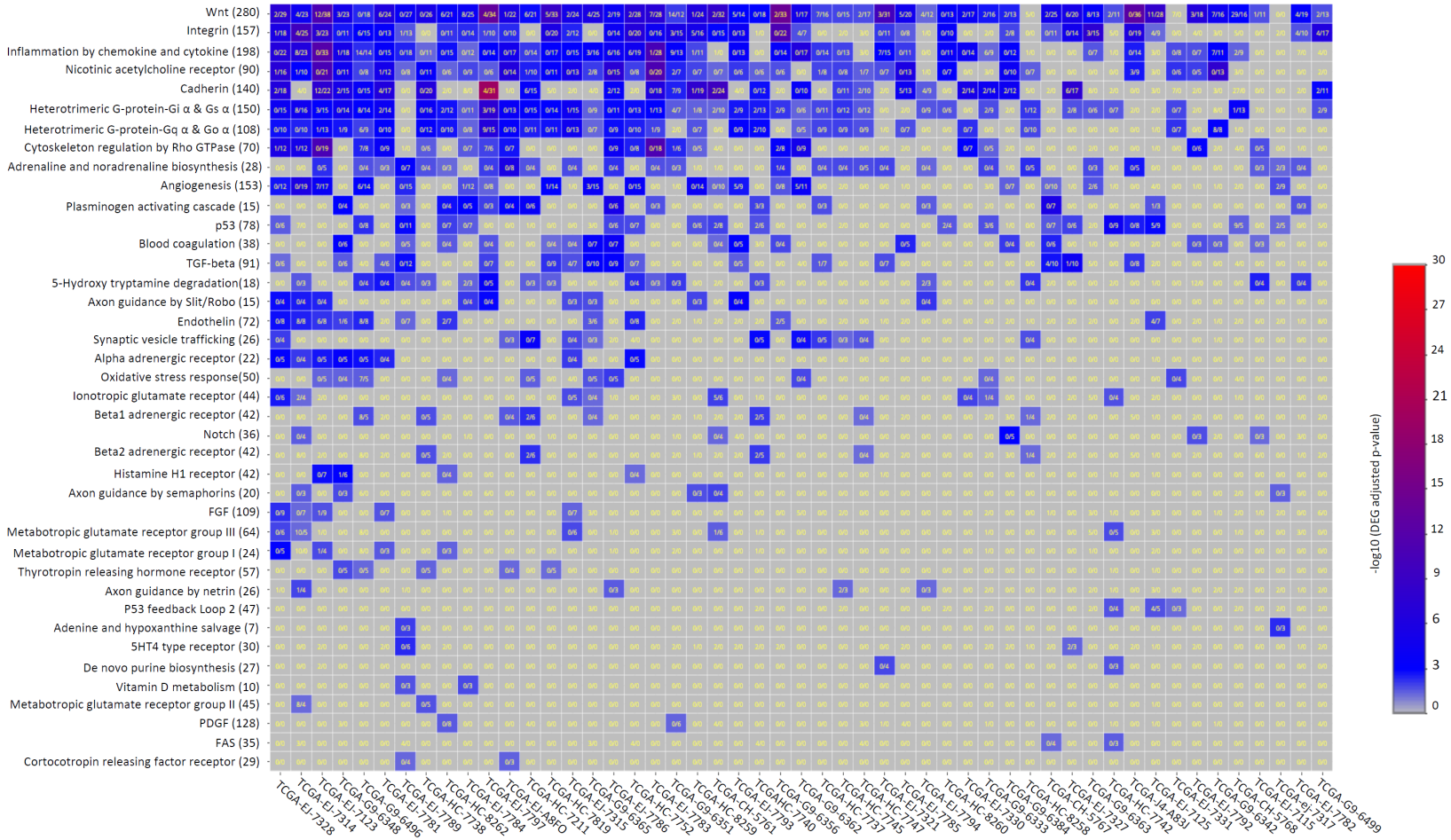
[59] Vo JN, Cieslik M, Zhang Y, Shukla S, Xiao L, Zhang Y, Wu YM, Dhanasekaran SM, Engelke CG, Cao X, Robinson DR, Nesvizhskii AI and Chinnaiyan AM. The landscape of circular RNA in cancer. Cell 2019; 176: 869-881, e13.



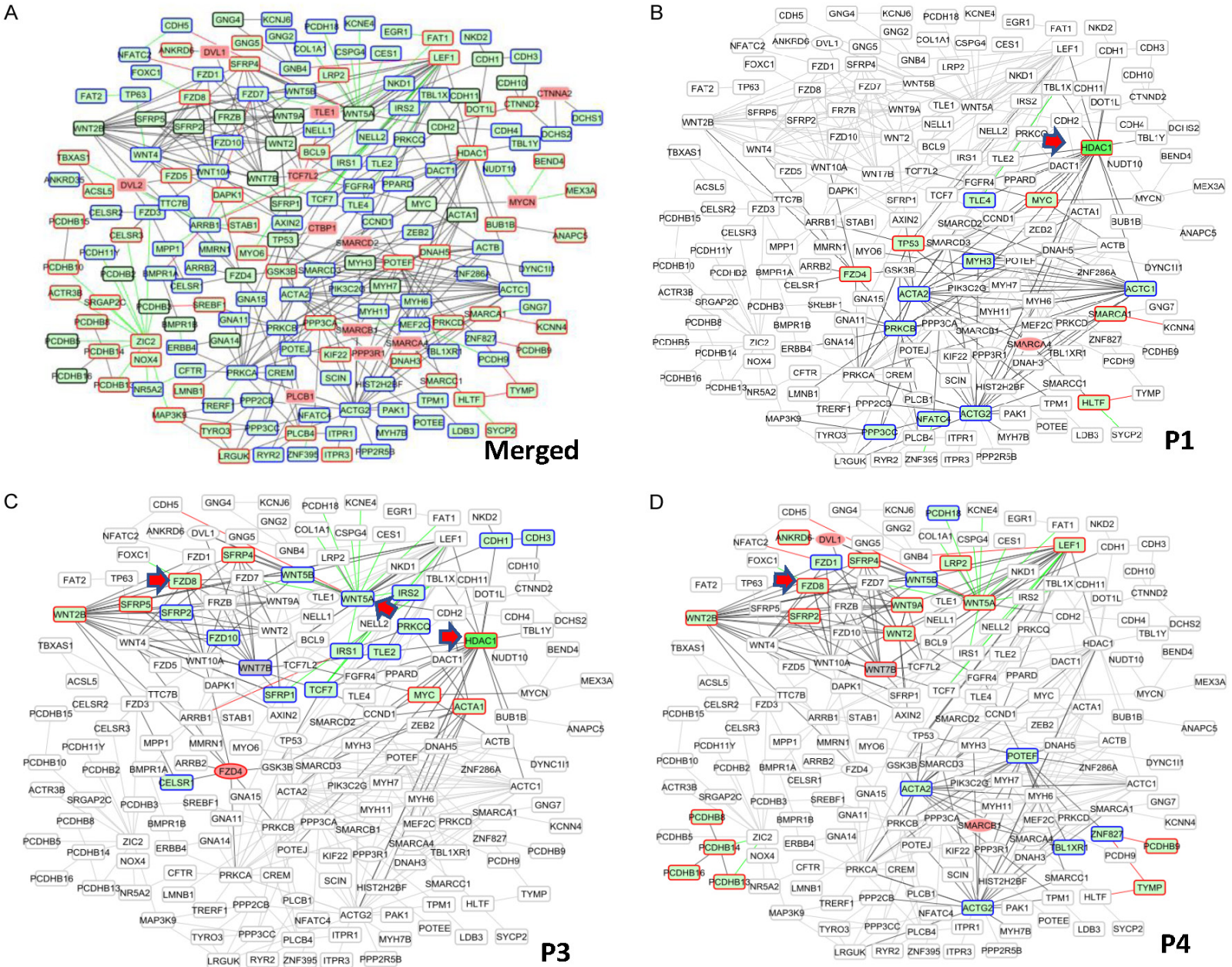
Supplementary Figure 1. Validation of four selected different types of mutations from Exome-seq analysis using PCR coupled to Sanger sequencing. A. Single base deletion in *MGA* gene in patient 5. B. Single base insertion in *CFTR* in patient 12. C. Missense mutation in *MST1R* in patient 14. D. Point mutation in *RASSF8* in patient 13. Red ovals highlight the regions of the mutations. Original (color-coded at the top) and mutated sequences are depicted below each sequence for comparison.

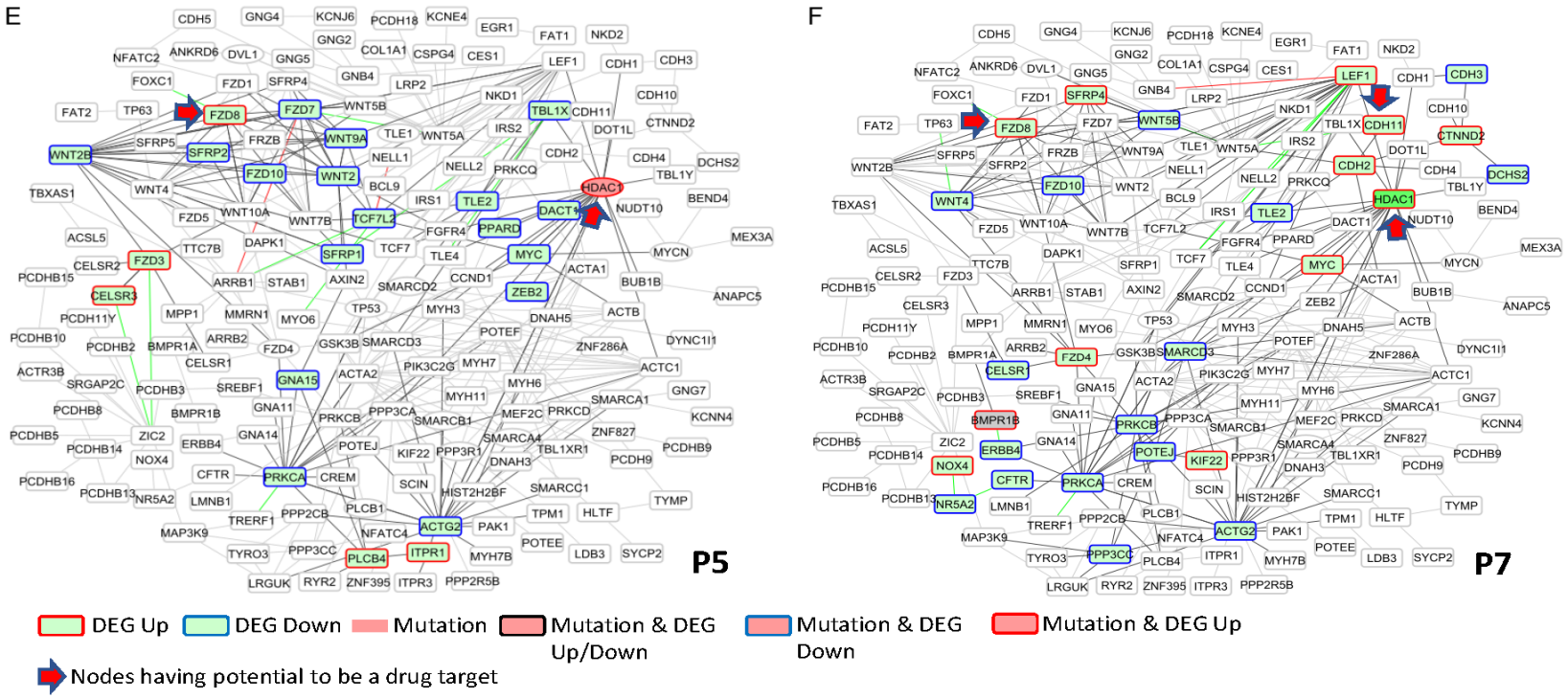


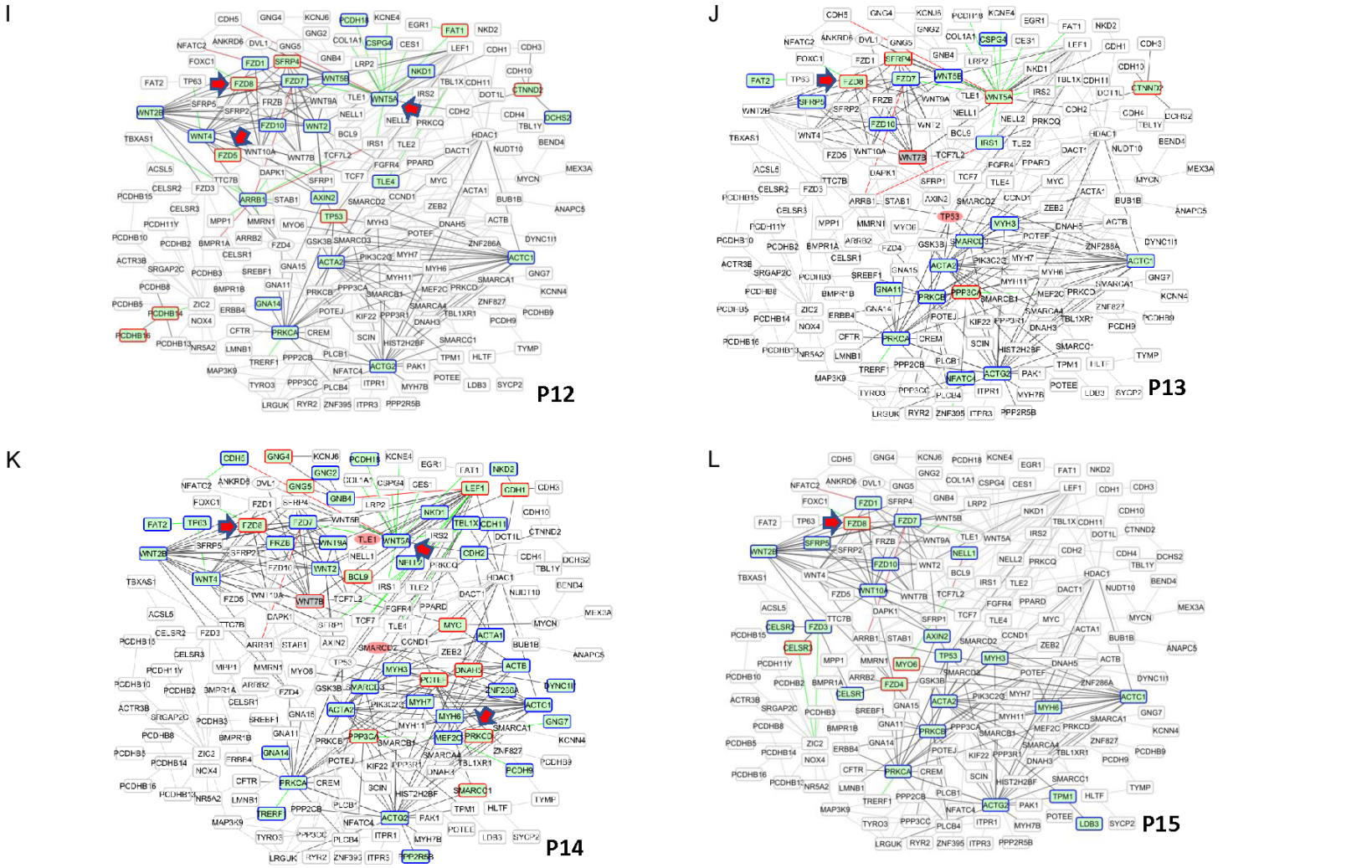
Supplementary Figure 2. Frequency of mutations in the 4 genes of **Figure 1D** in other cancer types (ACC, Adrenocortical Carcinoma; AML, Acute Myeloid Leukemia; BLCA, Bladder Urothelial Carcinoma; BRCA, Breast Invasive Carcinoma; CC, Cervical Cancer; CCRCC, Clear Cell Renal Cell Carcinoma; CESC, Cervical Squamous Cell Carcinoma; CHOL, Cholangiocarcinoma; CRC, Colorectal Cancer; DG, Diffuse Glioma; DLBCL, Diffuse Large B-Cell Lymphoma; EOC, Epithelial Ovarian Cancer; ESCA, Esophagogastric Adenocarcinoma; ESCC, Esophageal Squamous Cell Carcinoma; GBM, Glioblastoma; HCC, Hepatocellular carcinoma; HNSC, Head and Neck Squamous Cell Carcinoma; LGG, Low Grade Glioma; LUAD, Lung Adenocarcinoma; LUSC, Lung Squamous Cell Carcinoma; MSC, Melanoma Skin Cancer; NCCRCC, Non-clear Cell Renal Cell Carcinoma; NSCLC, Non-Small Cell Lung Cancer; PAAD, Pancreatic Adenocarcinoma; PCC, Pheochromocytoma; PRAD, Prostate Adenocarcinoma; SARC, Sarcoma; TET, Thymic Epithelial Tumor; THCA, Thyroid Carcinoma; UCEC, Uterine Corpus Endometrial Carcinoma; Green, mutation; Red, fusion).



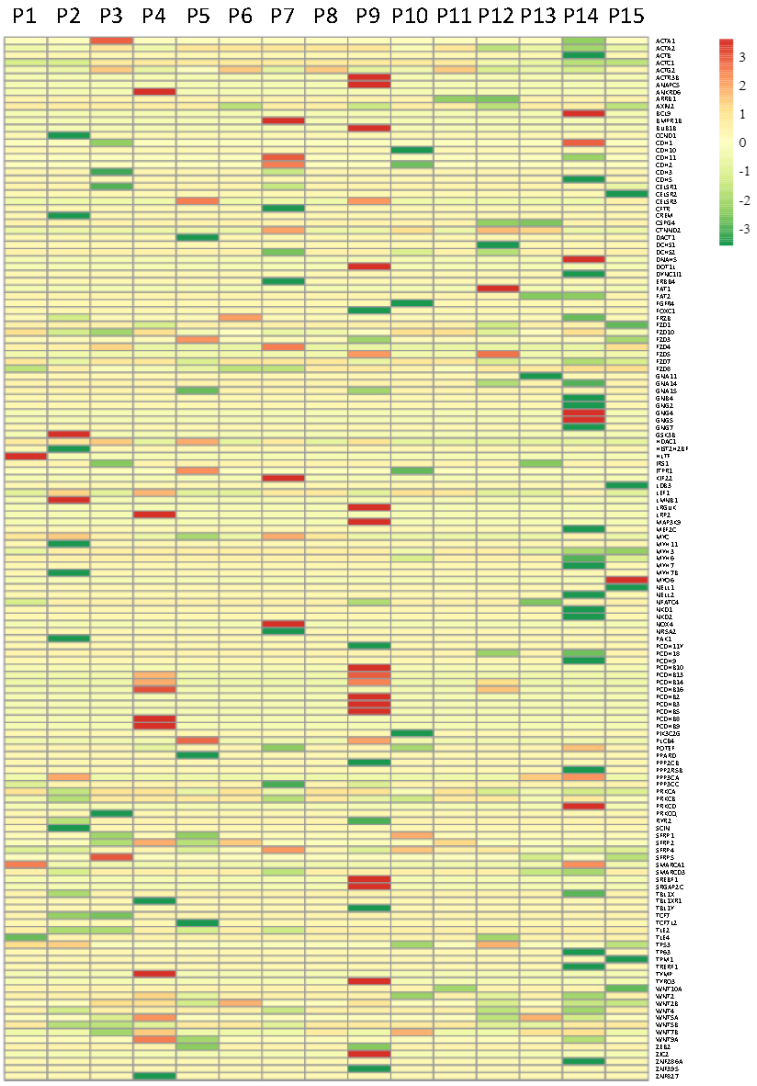
Supplementary Figure 3. Heatmap showing enriched pathways in 51 individual patients using gene expression data from TCGA for prostate cancer. Note that we selected only datasets for which patient-matched normal and prostate cancer tissue was available. The Y-axis specifies the pathways predicted by PANTHER to be enriched in the DEGs of each individual patient; numbers in parentheses indicate the number of total genes known to constitute the pathways. In the X-axis, TCGA codes for patients are given. Each box provides the following information: total number of mutated genes/total number of DEGs affected in each patient for the concerned pathway. *P*-values corresponding to the PANTHER-predicted alteration of a given pathway are shown in color code (scale is shown on the right of the table).


A
Merged
B
P1
C
P3
D
P4

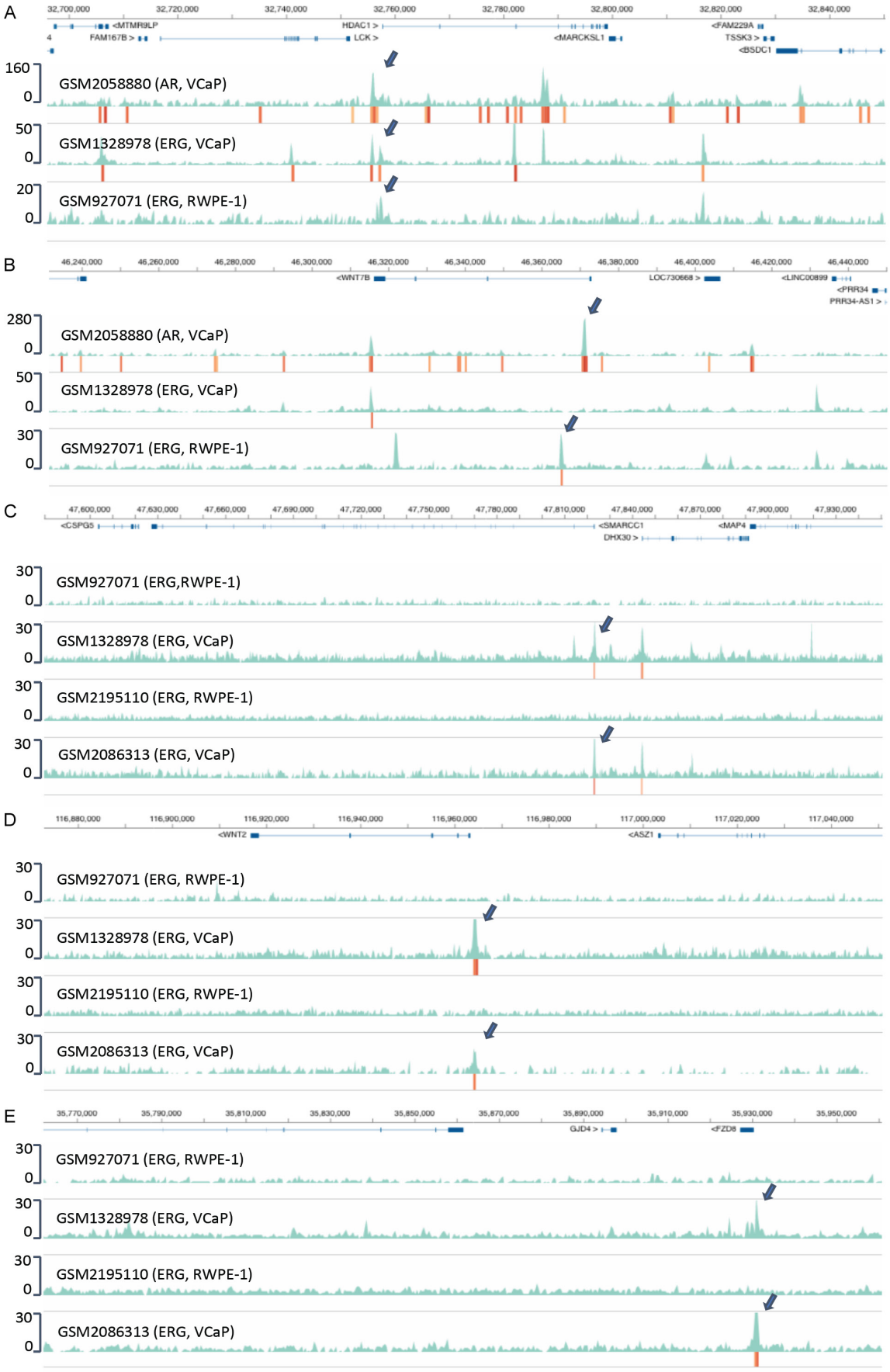

E
F
P5
P7
DEG Up
DEG Down
Mutation
Mutation & DEG Up/Down
Mutation & DEG Down
Mutation & DEG Up
Nodes having potential to be a drug target
G
H
P8
P11



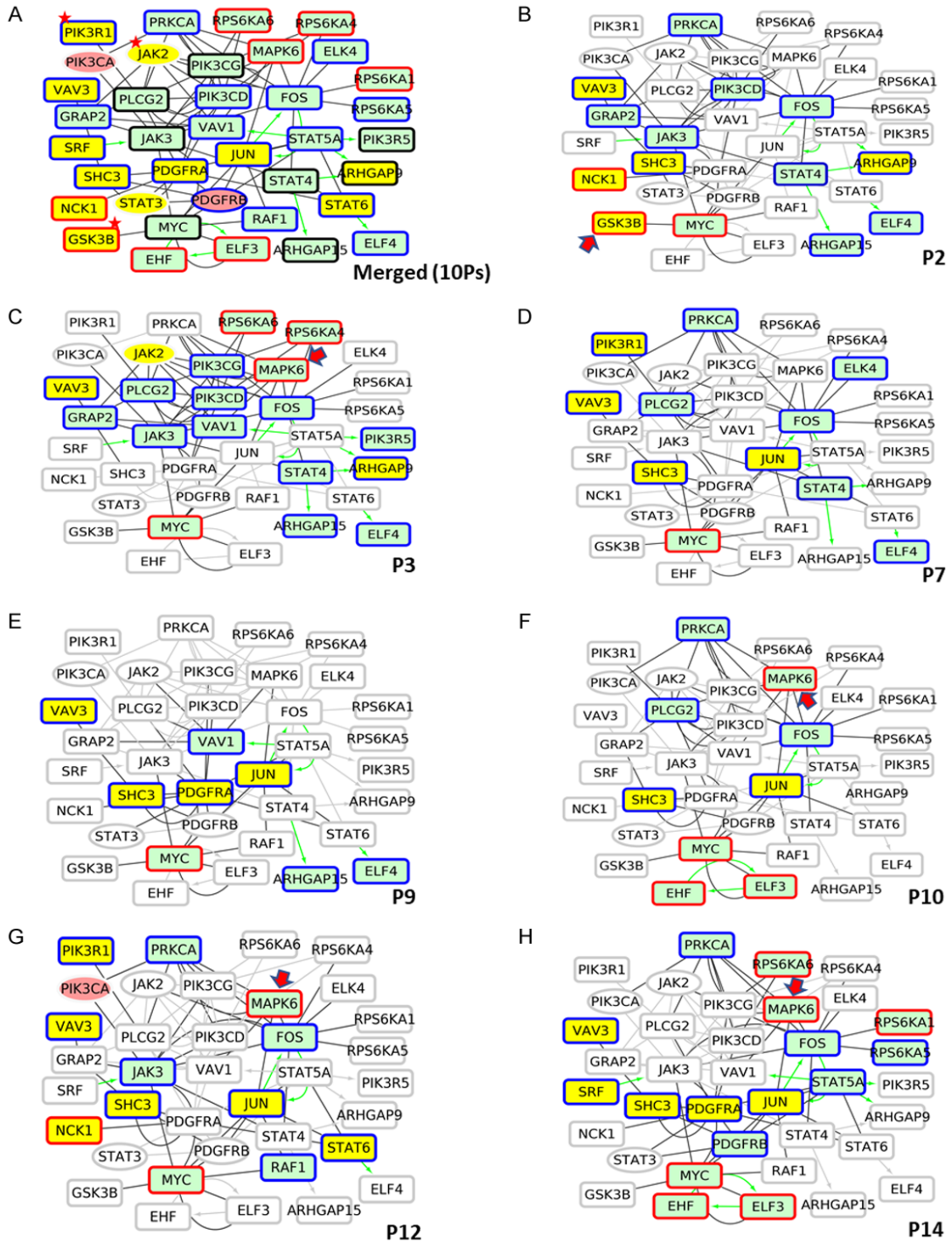
Supplementary Figure 4. Patient-specific deregulations of the Wnt pathway components for the indicated patients (P). (A) Merged network of all 15 patients used a background in (B-L). Frame and color codes are shown below P5. (B-L) Deregulated and/or mutated components of the Wnt pathway in each patient. Note that for the Wnt pathways components both inhibitory and activating drugs are in clinical trails such as Foxy-5 (WNT5A-mimicking peptide that activating Fzd2 and Fzd5) and Ipafricept (Fzd8 antibody blocking). (Clara et al (2020) Nat Rev Clin Onc 17,204).

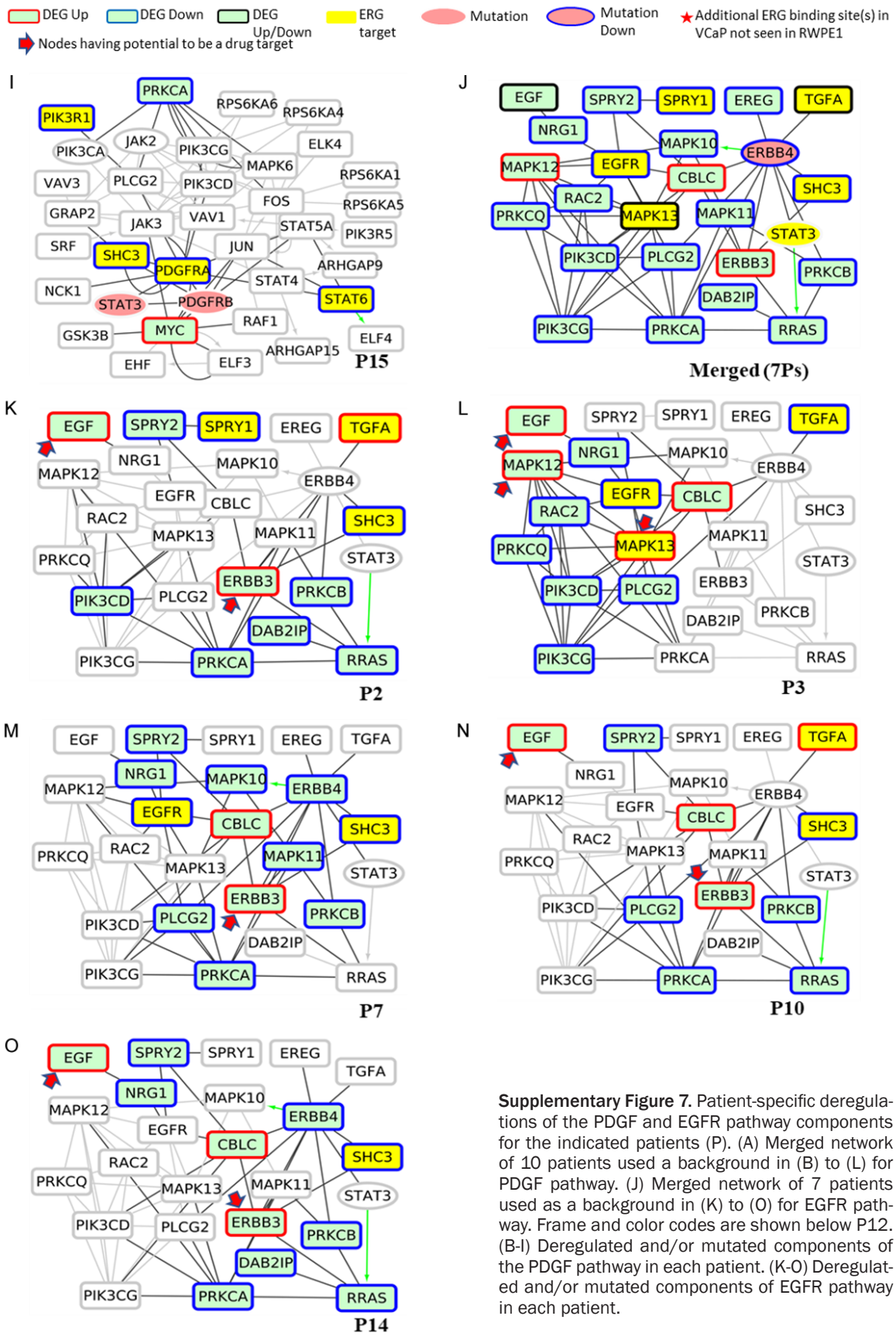


Supplementary Figure 5. Heat map showing differentially expressed genes of the WNT pathway. Color codes represent log2 fold changes. P1 to P15: Patients 1 to 15.

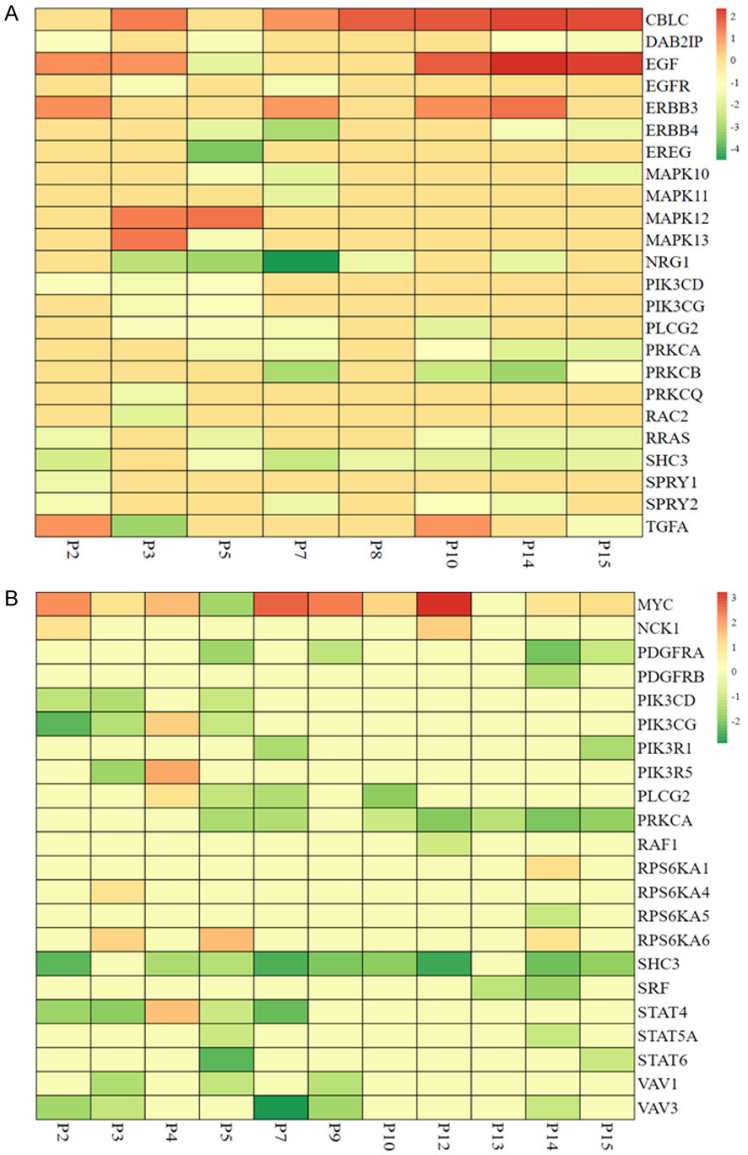

A
32,700,000
32,720,000
32,740,000
32,760,000
32,780,000
32,800,000
32,820,000
32,840,000
<MTMR9LP
HDAC1 >
<FAM229A
FAM167B >
LCK >
<MARCKSL1
TSSK3 >
<BSDC1
160
0
GSM2058880 (AR, VCaP)
50
0
GSM1328978 (ERG, VCaP)
20
0
GSM927071 (ERG, RWPE-1)
B
46,240,000
46,260,000
46,280,000
46,300,000
46,320,000
46,340,000
46,360,000
46,380,000
46,400,000
46,420,000
46,440,000
<WNT7B
LOC730668 >
<LINC00899
<PRR34
PRR34-AS1 >
280
0
GSM2058880 (AR, VCaP)
50
0
GSM1328978 (ERG, VCaP)
30
0
GSM927071 (ERG, RWPE-1)
C
47,600,000
47,630,000
47,660,000
47,690,000
47,720,000
47,750,000
47,780,000
47,810,000
47,840,000
47,870,000
47,900,000
47,930,000
<CSPG5
<SMARCC1
<MAP4
DHX30 >
30
0
GSM927071 (ERG,RWPE-1)
30
0
GSM1328978 (ERG, VCaP)
30
0
GSM2195110 (ERG, RWPE-1)
30
0
GSM2086313 (ERG, VCaP)
D
116,880,000
116,900,000
116,920,000
116,940,000
116,960,000
116,980,000
117,000,000
117,020,000
117,040,000
<WNT2
<ASZ1
30
0
GSM927071 (ERG, RWPE-1)
30
0
GSM1328978 (ERG, VCaP)
30
0
GSM2195110 (ERG, RWPE-1)
30
0
GSM2086313 (ERG, VCaP)
E
35,770,000
35,790,000
35,810,000
35,830,000
35,850,000
35,870,000
35,890,000
35,910,000
35,930,000
35,950,000
GJD4 >
<FZD8
30
0
GSM927071 (ERG, RWPE-1)
30
0
GSM1328978 (ERG, VCaP)
30
0
GSM2195110 (ERG, RWPE-1)
30
0
GSM2086313 (ERG, VCaP)

Supplementary Figure 6. AR binding and acquisition of ERG binding sites in VCaP prostate cancer cells relative to RWPE-1 prostate epithelial cells. Screenshots of qcGenomics browser NAVi displaying genes that show AR and/or ERG binding in their promoter regions. (A), *HDAC1*; (B), *WNT7B; (*C), *SMARCCl*; (D), *WNT2*; (E), *FDZ8*. ChIP-seq data sets in (A and B) are from GEO accession numbers (from top to bottom) GSM2058880 (AR, VCaP), GSM1328978 (ERG, VCaP) and GSM927071 (ERG, RWPE-1), as specified. Note that in (A) ERG binding at the *HDAC1* promoter is seen in VCaP and RWPE-1 cells, while in (B) for *WNT7B* a promoter-proximal ERG binding is seen in 'normal' RWPE-1 but not in VCaP cells; this ERG binding site is distant from the AR binding site. The ERG ChIP-seq data sets in (C-E) are from GEO accession numbers (from top to bottom) GSM927071 for RWPE-1, GSM1328978 for VCaP (both use anti-ERG antibody Epitomics 2805-1), GSM2195110 for RWPE-1 and GSM2086313 for VCaP. GSM2195110 was done by using Anti-ERG Clone 9FY Biocare #CM421 C, GSM2195110 used an anti-ERG antibody but did not provide the source. Note the consistency between corresponding experiments with different antibodies.

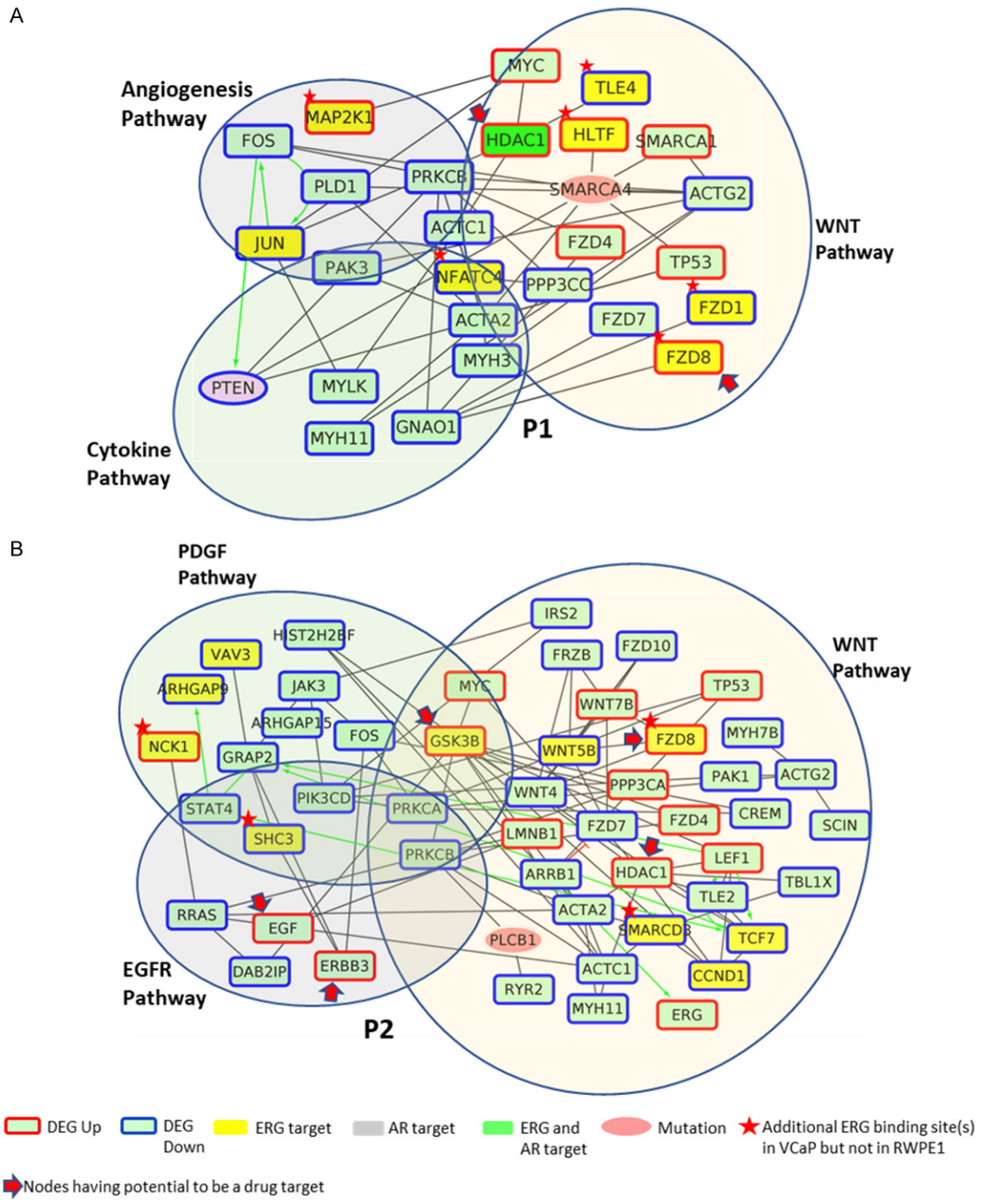




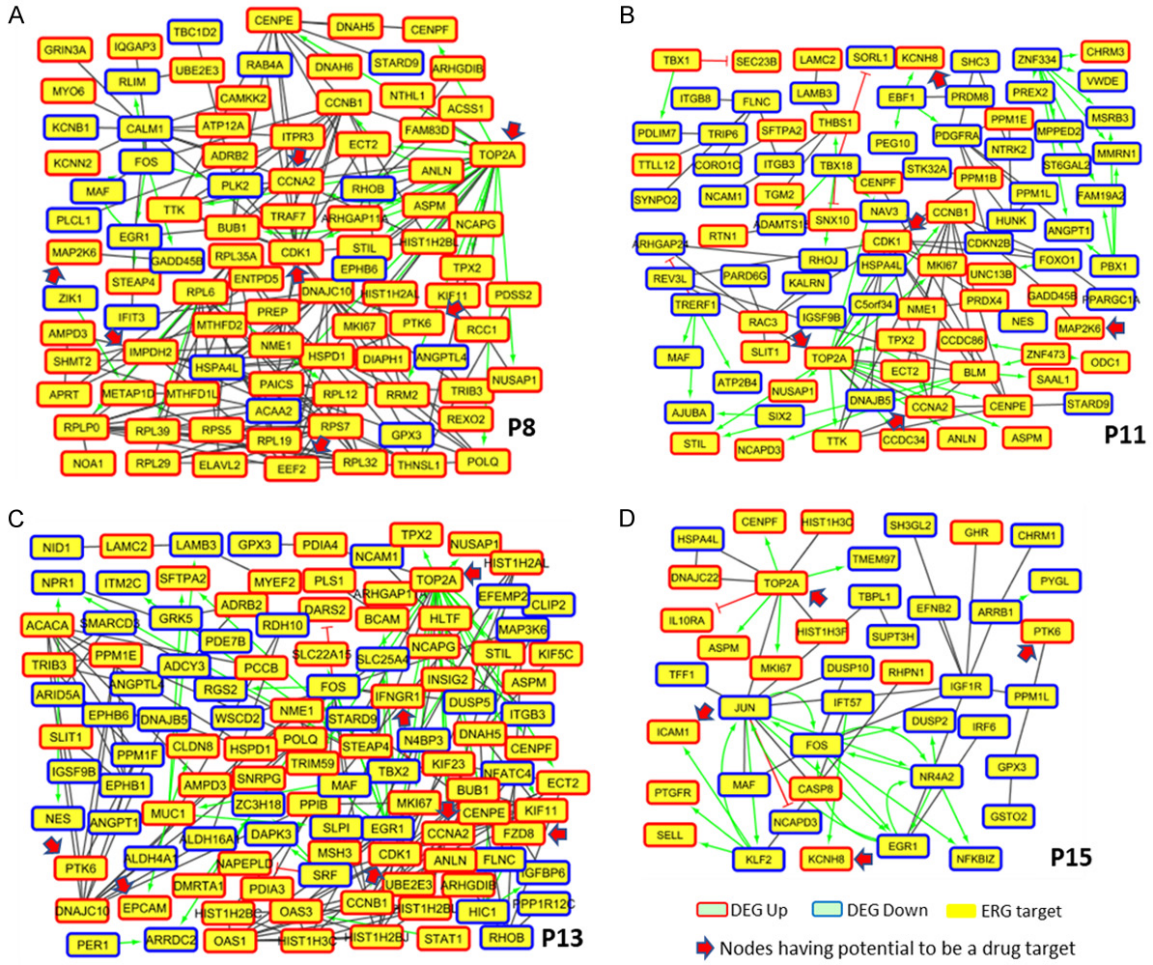
Supplementary Figure 7. Patient-specific deregulations of the PDGF and EGFR pathway components for the indicated patients (P). (A) Merged network of 10 patients used a background in (B) to (L) for PDGF pathway. (J) Merged network of 7 patients used as a background in (K) to (O) for EGFR pathway. Frame and color codes are shown below P12. (B-I) Deregulated and/or mutated components of the PDGF pathway in each patient. (K-O) Deregulated and/or mutated components of EGFR pathway in each patient.



Supplementary Figure 8. Heat maps showing differentially expressed genes of the EGFR (A) and PDGF (B) pathway. Color codes represent log2 fold changes. Px, patient number.



Supplementary Figure 9. Patient-specific deregulations of the PDGF and EGFR pathway components for the indicated patients (P). A. Network display of deregulated and mutated factors of patient P1 in the WNT, Angiogenesis and Cytokine Pathways to reveal connectivities between the different pathways. B. Similar representation of the de-regulated and mutated factors in P2 for the WNT, PDGF and EGFR pathways.



Supplementary Figure 10. Patient-specific networks comprising ERG transcription factor regulated DEGs. Majority of ERG-regulated DEGs are not common among (A) patient 8, (B) patient 11, (C) patient 13 and (D) patient 15 suggesting the need for a patient-centric approach.

Supplementary Figure 11. Patient-specific networks comprising MAZ transcription factor regulated DEGs. Majority of MAZ-regulated DEGs are not common among (A) patient 7 and (B) patient 12 suggesting the need for a patient-centric approach.

PUBLICATION N°2

Complexity against current cancer research: Are we on the wrong track?

Yasenya Kasikci, Hinrich Gronemeyer

International Journal of Cancer, 2021

Received: 27 October 2021 | Revised: 8 December 2021 | Accepted: 9 December 2021

DOI: 10.1002/ijc.33912

INVITED REVIEW

IJC INTERNATIONAL JOURNAL of CANCER

Complexity against current cancer research: Are we on the wrong track?

Yasenya Kasikci[1,2,3,4] | Hinrich Gronemeyer[3,4]

[1]Department of Functional Genomics and Cancer, Institut de Génétique et de Biologie Moléculaire et Cellulaire (IGBMC), Illkirch, France

[2]Centre National de la Recherche Scientifique, UMR7104, Illkirch, France

[3]Institut National de la Santé et de la Recherche Médicale, U1258, Illkirch, France

[4]Université de Strasbourg, Illkirch, France

Correspondence
Hinrich Gronemeyer, Department of Functional Genomics and Cancer, Institut de Génétique et de Biologie Moléculaire et Cellulaire (IGBMC), Illkirch, France.
Email: hg@igbmc.fr

Funding information
Agence Nationale de la Recherche; Institut National Du Cancer

Abstract

Cancer genetics has led to major discoveries, including protooncogene and tumor-suppressor concepts, and cancer genomics generated concepts like driver and passenger genes, revealed tumor heterogeneity and clonal evolution. Reconstructing trajectories of tumorigenesis using spatial and single-cell genomics is possible. Patient stratification and prognostic parameters have been improved. Yet, despite these advances, successful translation into targeted therapies has been scarce and mostly limited to kinase inhibitors. Here, we argue that current cancer research may be on the wrong track, by considering cancer more as a "monogenic" disease, trying to extract common information from thousands of patients, while not properly considering complexity and individual diversity. We suggest to empower a systems cancer approach which reconstructs the information network that has been altered by the tumorigenic events, to analyze hierarchies and predict (druggable) key nodes that could interfere with/block the aberrant information transfer. We also argue that the interindividual variability between patients of similar cohorts is too high to extract common polygenic network information from large numbers of patients and argue in favor of an individualized approach. The analysis we propose would require a structured multinational and multidisciplinary effort, in which clinicians, and cancer, developmental, cell and computational biologists together with mathematicians and informaticians develop dynamic regulatory networks which integrate the entire information transfer in and between cells and organs in (patho)physiological conditions, revealing hierarchies and available drugs to interfere with key regulators. Based on this blueprint, the altered information transfer in individual cancers could be modeled and possible targeted (combo)therapies proposed.

KEYWORDS
complexity challenge, conceptual problems in cancer genomics, information transfer, integrated network analysis

1 | INTRODUCTION

In order to develop our criticism to current concepts of cancer genomics, we will first elute to the enormous complexity of the human body which relies on dynamic regulated information transfer reminiscent of social networks.

Abbreviations: APL, acute promyelocytic leukemia; AR, androgen receptor; CDK, cyclin-dependent kinase; CML, chronic myeloid leukemia; DEG, differentially expressed gene; ENCODE, Encyclopedia of DNA Elements (encodeproject.org); ERG, member of the erythroblast transformation-specific (ETS) family of transcriptions factors; HCA, Human Cell Atlas (humancellatlas.org); IHEC, International Human Epigenome Consortium (ihec-epigenomes.org); RB1, retinoblastoma transcriptional co-repressor 1; SPC(+) alveolar type II cells, surfactant-associated protein C (SPC)-expressing alveolar type II cells.

The genome of a fertilized human zygote contains already all the information that is necessary to develop and sustain an extremely complex system like the human body. A fully developed human comprises 37 trillion cells[1] which are functionally and spatially organized according to a highly sophisticated body plan, able to respond to diverse sets of internal and environmental cues. None of these cells lives in isolation or without instructive functional information. Development, death, survival and homeostasis of the individual components and higher-order communities in such multicellular organisms depend on elaborate intracellular and intercellular communication. This communication is ensured by networks of information transfer within and between organelles, cells, organs and other body compartments to coordinate and direct growth, survival, differentiation/functionalization and metabolism. These compartments are dynamically regulated, frequently redundant and highly plastic to maintain functionality. This system of information transfer occurs at multiple organizational levels by employing a multiplicity of molecular actors.

The actors (nodes) in this network are highly diverse, both from their chemical constitution and mode of communication. Small chemical substances or peptides, like steroid hormones may act in endocrine, paracrine, autocrine or intracrine fashion to initiate a network of responses in cells distant to the origin of the signal, or they may regulate intracellular events corresponding to differentiation state and acquired functionalization of a given cell/organelle in a given body compartment. Steroid hormones have evolved to fit into pockets of their cognate nuclear receptors which can act as inducible transcription factors and trigger a plethora of temporally organized gene-regulatory responses of sets of genes, which then regulate multiorgan phenomena like pregnancy. Protein ligands can bind to cell membrane receptors and upon forming complexes with associated/regulatory factors, initiate for example kinase cascades. Proteins may homodimerize or heterodimerize or form multicomponent complexes which together with other specified cell components can generate specialized machineries, like ribosomes, or transcription and epigenetic regulatory complexes.

Communication means are as diverse as the nature of the actors. Enzymatically catalyzed reactions are at the basis of metabolic pathways, regulatory cascades and cell fate phenomena and the involved enzymes are often organized as directional (oxidative phosphorylation, kinase cascade) or cyclic (Krebs cycle, cell cycle) information transfer, regulated by feed-back or feed-forward processes or external cues. While these systems have a variable degree of autonomy, they are limited by their cell fate destiny and overarching control systems if certain checkpoints are surpassed (differentiation, apoptosis and autophagy).

2 | THE FUNCTIONS OF THE BODY AND OF HUMAN SOCIETIES RELY SIMILARLY ON NETWORK-BASED INFORMATION TRANSFER

The level of organization and complexity of the human body resembles that of human societies. Indeed, both systems are similarly hierarchically organized and depend on dynamic, responsive, flexible, redundant and resilient networks of information transfer.[2,3] In the case of human communities, this communication has various means of information exchange, such as written or visual information transfer (media), individual contacts and relationships, regulations (laws), control (police) and defense (military) systems and hierarchies (government). All these have approximations of information transfer in the human body. There is one written law, the genetic information encoded in the genome and its architecture. Access to this information can be restricted (epigenetics) and it can be edited to target invaders (immune defense). Repair and regeneration systems could be regarded as "healthcare system" equivalent of cells and organs. Fail-safe systems like tumor surveillance, autophagy, apoptosis or senescence are reminiscent of police and judicial systems. Information transfer by social interactions corresponds to (hetero)oligomeric interaction of molecules and there is a close approximation of industry with molecular machines in the cell. Governance and administration find their counterparts in the gene regulatory systems, such as transcription regulation, epigenetics, chromatin architecture and a large variety of effects exerted by regulators like the huge and still underexplored class of noncoding RNAs. Thus, the human body functions like a society (family, city and country), with spatially organized compartments (organelles, cells and organs) that rely on communication networks established between the system components.

3 | MODELING AND REPRESENTATION OF BIOLOGICAL NETWORKS AND THEIR PERTURBATIONS

Communication can be formally represented as a network[4] of elements ("nodes"—molecules) and their connections by lines ("edges"—protein-protein, DNA-protein, transcription factor-target gene, enzyme-substrate or any other type of interaction). A biological network corresponds to a theoretical model for the functional flux of information through any type of biological structure (eg, cells or subcellular structure, groups of related cells and organs).[5] Various aspects (eg, importance of a note, type and strengths of connections) can be integrated and mathematical graph theory-based analyses[6] can provide crucial information about the nature and function of a network. The network theory has been extensively applied to understand the topological information transfer in brain,[4] but its use in developmental or cancer biology[7] is rather limited.

Molecular changes, such as mutations and deregulation of gene expression, are perturbations of the corresponding biological network. The functional consequences of perturbations can propagate along the connected edges and affect a subset of nodes. These systemic spreading effects have been inferred recently in a patient-matched analysis of cancer and normal prostate epithelial tissues.[8] This analysis has two important messages. First, the deregulated and mutated genes are not isolated but form a subnetwork within larger communities, exemplified here by the Wnt-network[9] (Figure 1A). Note that a similar network reveals the global interactions established by the

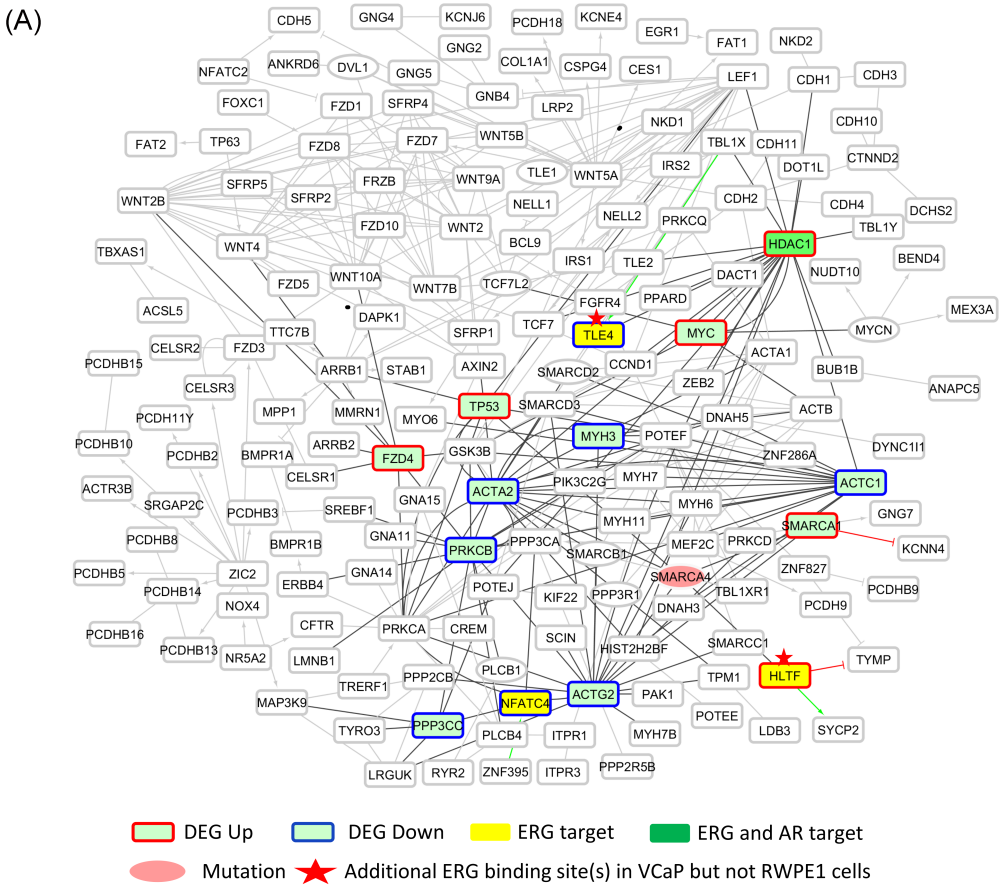


FIGURE 1 Legend on next page.

International Committee on Intellectual Cooperation (ICIC),[10] which included leading scientists like Albert Einstein and Marie Curie (Figure 1B). The second key message from the patient-matched analysis was that even for the very similar group of ERG (member of the erythroblast transformation-specific (ETS) family of transcriptions factors)-positive prostate cancers, subnetworks of differentially expressed genes (DEGs) within the Wnt network varied from one patient to the other.[8] Thus, the Wnt network was commonly affected in the large majority of patients, but the deregulation affected very different subsets of interconnected actors.

Dramatic systemic effects can be exerted by perturbation of higher-order regulatory actors. If for example epigenetic regulation is impaired, the normal spatio-temporal flux of genetic information is altered and can lead to pathologies, including cancer. It is indeed known that epigenetic mediators are deregulated or mutated in cancer.[11,12] Determining how disease-associated variations affect (the hierarchical organization of) cellular networks and revealing the consequences (propagation) of perturbations are therefore important diagnostic steps to guide treatment. Bioinformatics tools for the hierarchical analysis of networks have been developed (Ref. 13 for a recent example) and the predictions can be validated by genome engineering in model systems.

4 | PERTURBATIONS CAN ADAPT TO CELL TYPE-SPECIFIC COMMUNICATION NETWORKS TO BECOME ONCOGENIC

The term "oncogene" describes the capacity of a mutated or overexpressed gene, to turn a cell into a cancer cell. However, this does not mean that any type of cell becomes a tumor cell when an oncogene is activated, or a tumor suppressor gene is inactivated. Even though cancer can develop in all cell types of the body, the genetic causes of tumorigenesis are generally cell-type selective. The most convincing experiment demonstrating oncogene susceptibility of a single tissue—and resistance to the oncogenic insult of the others—was the inducible ubiquitous expression of *K-Ras*^{G12V}. These mice developed specifically lung adenomas and adenocarcinomas despite harboring the oncogene in all cells.[14] Furthermore, in a subsequent study, the Barbacid team showed that only one type of lung cell integrates the *K-Ras*^{G12V} perturbation into its signaling network to allow tumorigenesis to occur and identified the cancer-initiating cells by demonstrating that only SPC+ (surfactant-associated protein C) alveolar type II cells were able to form hyperplastic lesions that progressed to adenocarcinomas.[15] These data show that an oncogene does not function universally and that a particular perturbation can lead to cancer only in the context of a defined cell type which will develop along a defined developmental trajectory. Mutational landscaping across major cancer types confirmed the existence of tumor type-specific mutations.[16] It is therefore the specific makeup of the regulatory network at a certain differentiation state and its programmed evolution which cooperates with a (or multiple) perturbation(s) to drive a cell into a cancer trajectory.

However, that single perturbations like K-ras G12V drive tumorigenesis on their own like in the mouse model is rare. Over the approximate 1 to 50 years of tumor evolution, multiple mutations accumulate, facilitated upon perturbation of control mechanisms and at diagnosis a patient's tumor frequently has a large number of genetic aberrations, including not only point mutations but also specific chromosomal translocations and other genome lesions which provide a growth advantage. Mutations are generally separated into one or a limited number of "driver" and a large number of "passenger" mutations[16,17] which cooperate in the context of a given cell fate trajectory, albeit the precise impact of low-frequency mutations on a given cancer trajectory still has to be elucidated.[18] From the perspective of an individual tumor, also high impact passenger mutations should be considered for potential effects on the entire perturbed cellular communication network.

In addition, tumors evolve clonally and present at diagnosis frequently with intratumoral heterogeneity.[19] Even the tumor microenvironment of a single tumor can diverge spatially.[20] Thus, as it has been pointed out by Mel Greaves, each cancer is unique, has its specific history and most of its complexity can be explained by classical evolutionary principles, including the appearance of clones during tumor evolution and therapy.[19,21-23]

5 | THE "HALLMARKS OF CANCER" DESCRIBE PERTURBATION OF COMMUNICATION NETWORKS

The recent increase in "hallmark" publications[24-27] appears indicative of the desire to understand (patho)physiological phenomena in a systemic manner by trying to connect molecular data with complex

FIGURE 1 Networks of information transfer in the human body and human societies are similar. (A) Wnt-network (gray) in which differentially regulated (red lines upregulated, blue downregulated) and mutated (pink sphere) genes are displayed in color. From exome-seq and RNA-seq analysis of in a patient-matched prostate cancer and normal prostate epithelium. Yellow spheres, ERG target genes, asterisk indicates gain-of-regulation by ERG due to the TMPRSS2-ERG fusion (from Ref. 8). Pink oval, mutated gene; rectangles represent upregulated (red line), downregulated (blue line), ERG (yellow filling) or ERG and AR (green filling) target genes; red asterisks reveal genes which acquired ERG-binding in prostate cancer cells. (B) Graph representing the social network established by the International Committee on Intellectual Cooperation (ICIC).[10] Founded in 1922 by the League of Nations, its goal was to coordinate the restructuration of knowledge circulation. The ICIC involved leading scientists, including Albert Einstein and Marie Curie, and generated a complex network between transnational institutions and societies, congresses and individuals. Node sizes illustrate the number of connections. This file is licensed under the Creative Commons Attribution 4.0 International license [Color figure can be viewed at wileyonlinelibrary.com]

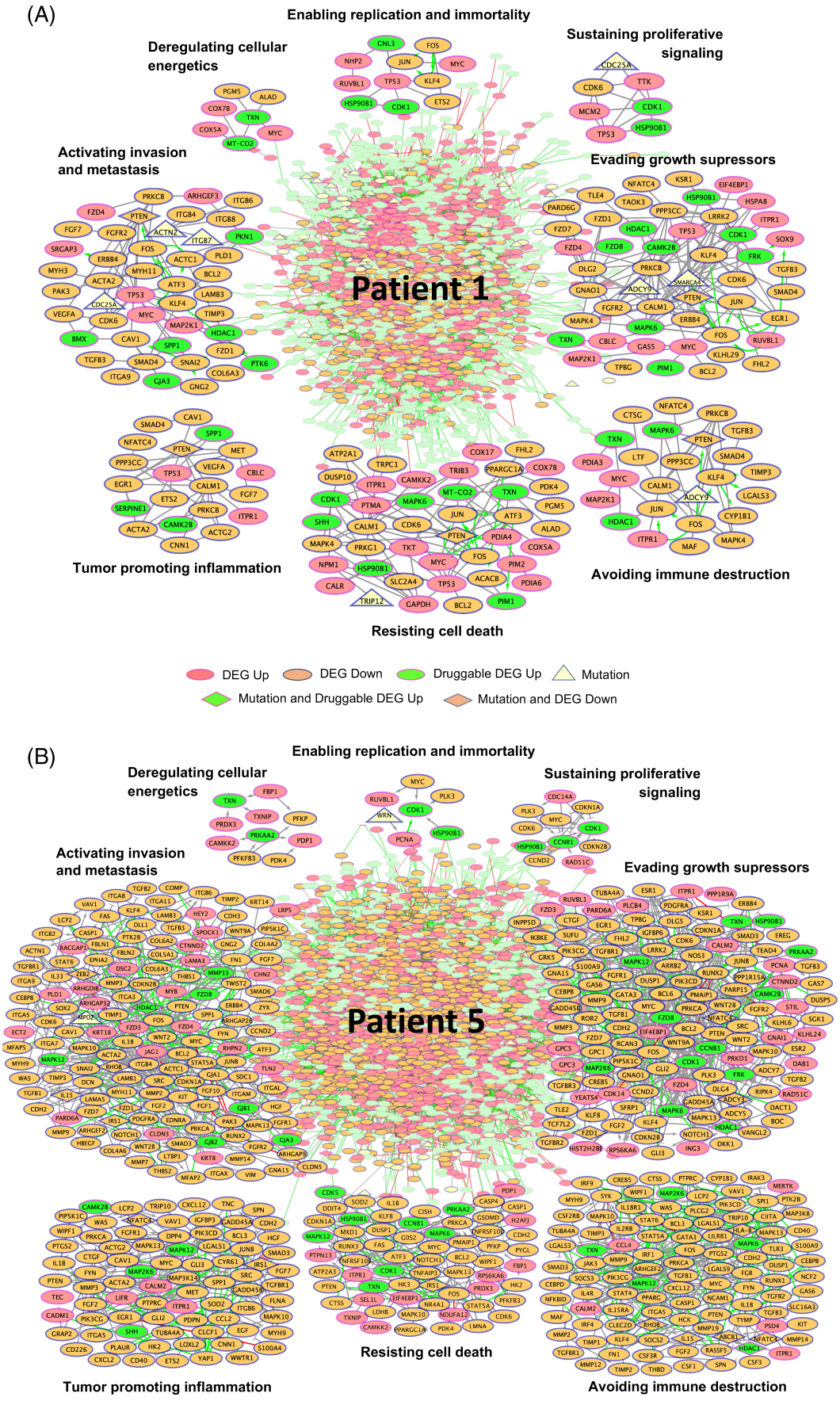


FIGURE 2 Legend on next page.

(aberrant) physiological functions. Indeed, the huge amount of "big data" information continuously accumulating from functional genomics, proteomics and metabolomics asks for being integrated in a conceptual framework for exploitation.

This framework has been provided by the conceptualization of hallmarks of cancer[28,29] comprising cell-autonomous features like unrestricted proliferation, replicative immortality, adapting metabolism, opposing genome integrity and escaping suicidal death programs, as well as cell-interactive phenomena like angiogenesis induction, immune evasion and promotion of inflammation, invasion and metastasis. All these hallmarks correspond actually to regulated communication networks adapted to specific cell biological phenomena. They can be formalized as simplified directional networks with generally known key players[28] or as more sophisticated signaling networks powered by Google Maps.[30] Figure 2 illustrates the deregulated subnetworks corresponding to the eight hallmarks of cancer, extracted from a complex network of DEGs that was established by comparing mutations and gene expressions in two ERG-positive prostate cancers with the patient-matched normal prostate epithelium.

Thus, in principle, with the large number of functional genomics technologies, already existing data from international efforts like International Human Epigenome Consortium (IHEC), Encyclopedia of DNA Elements (ENCODE), ongoing efforts like the Human Cell Atlas (HCA)[31,32] and the complex disease-focused LifeTime Initiative[33] together with an enormous, continuously increasing sequencing power, we should be able to construct cell/organ-specific networks from healthy individuals as a blueprint and superimpose the deregulated networks from individual cancers.

6 | GENE-CENTRIC ANALYSES DISREGARD THE SYSTEMIC CHANGES IN CANCERS

Traditionally, molecular (cancer) biologists are used to gene-centric views and analyses. While this has led to important discoveries, like (proto)oncogenes and tumor suppressor genes and therapeutics like Herceptin (Trastuzumab), there has been rather limited conceptual progress in cancer therapy. Indeed, cancers that are now considered curable, such as acute promyelocytic leukemia (APL) or chronic myeloid leukemia (CML) are caused by specific genetic events, including the generation of fusion genes due to chromosomal translocations. In such cases, targeting the cancer cause is promising and has been a major success in APL by destabilizing the PML-RARα fusion protein[34] and in CML by targeting the kinase in the BCR-Abl1 protein.[35,36] It has been argued that CML is not a typical cancer but corresponds rather to a premalignant condition[19] and it is likely that in these cases the limited number of persistent cancer-causing aberrations accounts for the therapeutic success. Certain childhood leukemias, like B-cell precursor acute lymphoblastic leukemia, may be even linked to contemporary lifestyle and be preventable.[37] However, in contrast to pediatric leukemia and cases like CML and APL, leukemias of the adult can be highly complex.

In solid cancers, where a plethora of genetic aberrations accumulate before the disease becomes diagnosed, progress in the sense of providing novel therapeutic paradigms beyond the mere announcement of having identified a novel target, be it a protein-coding or one of the ever-increasing numbers of regulatory RNAs, has been scarce. This is somewhat surprising, given the huge amount of studies on cancer genomes, transcriptomes, epigenomes, metabolomes and other global profiling of cancer cells. Therefore, the question appears justified, if there could be a conceptual issue with today's cancer genomics.

One of the salient features of current cancer genomics studies is that tumor DNAs and transcriptomes of hundreds to thousands of patients are sequenced. Indeed, the cancer genome atlas program stores over 2.5 petabytes of genomic, epigenomic, transcriptomic and proteomic data (www.cancer.gov/tcga). In the majority of studies, the aim is to find common alterations in a defined set of patients, generally according to histopathologic and clinical parameters. This includes studies analyzing 2000 breast cancer tumors[38] or an integrative genomics analysis of advanced prostate cancer.[39] Examples of prototypic conclusions derived from the latter study are "… retinoblastoma transcriptional co-repressor 1 (RB1) loss in 21% of cases …," "…71.3% of cases harbored androgen receptor (*AR*) pathway aberrations …" or "non-AR related clinically actionable alterations include aberrations in the PI3K pathway (49%), DNA repair pathway (19%), RAF kinases (3%), CDK (Cyclin-dependent kinase) inhibitors (7%) and the WNT pathway (5%)." Remarkably, although mutations in coding regions play a fundamental role in the development of cancer, a recent reanalysis of the frequencies of cancer gene mutations in the United States

FIGURE 2 Deregulated hallmarks of cancer correspond to subnetworks of deregulated genes. (A) Illustration of the DEG networks in an ERG-positive prostate cancer (patient 1), derived from comparing tumor and adjacent normal epithelium. Due to its complexity, the global DEG network cannot be displayed such that nodes and edges are discernible (central illustration). From this network, subnetworks have been extracted which contain DEGs that are part of the ensemble of genes constituting a particular cancer hallmark.[30] (B) Similar network established for another ERG-positive prostate cancer (patient 5). Note that despite their classification as ERG-positive prostate cancers, the deregulated networks differ dramatically. Networks depicted in A and B were established from exome-seq (to identify mutations) and RNA-seq (to identify differentially expressed genes) and integration of validated transcription factor-target gene and protein-protein interaction information. Note the dramatic differences for each of the "hallmarks of cancer" between the prostate cancers of patient 1 and 5, both of which are ERG-positive. For details, see Ref. 8. Triangles reveal mutated, diamonds mutated and deregulated genes; oval shapes indicate up-regulated (red lining, pink color) upregulated and drug-targetable (red lining, green color) or downregulated (blue lining, orange color) genes. When known, connectivities are displayed as green (activation) or red (inhibition) lines; unknown connectivities and protein-protein interactions are shown as black lines [Color figure can be viewed at wileyonlinelibrary.com]

TABLE 1 Existing drugs that target genes in the cancer hallmark DEG networks of patients 1 and 5

Gene name	Drug name	Drug action
CDK1	SELICICLIB (Clinical Trial)	Pan-CDK inhibitor
	ALVOCIDIB (Clinical Trial)	Pan-CDK inhibitor
FRK	REGORAFENIB (Approved)	Small molecule multikinase inhibitor
	NINTEDANIB (Approved)	Small molecule multikinase inhibitor
	VANDETANIB (Approved)	Small molecule multikinase inhibitor
HDAC1	MOCETINOSTAT (Clinical Trial)	Isotype-selective inhibitor of class 1 HDACs
	PANOBINOSTAT (Approved)	Pan-HDAC inhibitor
	BELINOSTAT (Approved)	Pan-HDAC inhibitor
	VORINOSTAT (Approved)	Pan-HDAC inhibitor
	ROMIDEPSIN (Approved)	Pan-HDAC inhibitor
	ABEXINOSTAT (Clinical Trial)	Pan-HDAC inhibitor
BMX	FOSTAMATINIB (Approved)	Small molecule multikinase inhibitor
	ZANUBRUTINIB (Approved)	Small molecule inhibitor of Bruton's tyrosine kinase
CAMK2B	NIFEDIPINE (Approved)	Calcium channel blocker
CCNB1	SELICICLIB (Approved)	Pan-CDK inhibitor
CDK5	OLOMOUCINE (Clinical Trial)	Pan-CDK inhibitor
	ALSTERPAULLONE (Clinical Trial)	Pan-CDK inhibitor
FZD8	CARBAMAZEPINE (Approved)	Inhibits Wnt/β-catenin signaling by binding to an allosteric site of the Wnt ligand receptor FZD8
GJA3	CARBENOXOLONE (Approved)	Gap junction blocker
GJB1	CARBENOXOLONE (Approved)	Gap junction blocker
GJB2	CARBENOXOLONE (Approved)	Gap junction blocker
HSP90B1	GEDANAMYCIN (Clinical Trial)	Hsp90 inhibitor
	RADICICOL (Clinical Trial)	Hsp90 inhibitor
MAP2K6	TRAMETINIB (Approved)	Reversible allosteric MEK inhibitor
	SELUMETINIB (Clinical Trial)	Non-ATP-competitive MEK inhibitor
	BINIMETINIB (Approved)	Non-ATP-competitive MEK inhibitor
	COBIMETINIB (Approved)	Small molecule MEK inhibitor
	FOSTAMATINIB (Approved, Clinical Trial, orphan drug)	Small molecule multikinase inhibitor
MAPK12	SORAFENIB (Approved)	Targets the RAF/MEK/ERK pathway
MAPK6	SORAFENIB (Approved)	Targets the RAF/MEK/ERK pathway
MMP15	MARIMASTAT (Clinical Trial)	Broad-spectrum inhibitor of all major MMPs
MT-CO2	CELECOXIB (Approved)	Cyclooxygenase-2 (COX-2) inhibitor
PIM1	MITOXANTRONE (Approved)	Nanomolar inhibitor of PIM1 kinase
PRKAA2	METFORMIN (Approved)	Metformin activates AMPK in certain cells. PRKAA2 encodes AMPK subunit a2
PTK6	VANDETANIB (Approved)	Small molecule multikinase inhibitor
	DASATINIB (Approved)	Small molecule multikinase inhibitor
SHH	VISMODEGIB (Approved)	Inhibits the transmembrane protein Smoothened homolog (SMO) to inhibit the Hedgehog signaling pathway
TXN	PX-12 (Clinical Trail)	Small-molecule inhibitor of TXN

Note: This list may not be exhaustive and does not imply that patients should be treated with these compounds which may lack target and/or tissue selectivity. What this table is supposed to show is that drugs targeting nodes within the DEG networks, which are even part of the genes constituting cancer hallmarks, can be developed and there is no reason to exclude that more active or selective drugs can be generated for combinatorial therapy approaches.

concluded that cancer genetics is less dominated by high-frequency, high-profile cancer driver genes than studies limited to a subset of cancer types have suggested in the past.[40] Such frequencies of genome aberrations deliver important information on the possible alterations in a specific cancer type and are useful for patient stratification. However, they do not provide comprehensive insight into the altered molecular communication networks of a given patient and do not consider mutations that affect regulatory regions of genes.

In the case of prostate cancer genome-wide association studies have predicted more than 160 disease-susceptibility loci, but for the large majority the molecular basis of this susceptibility and their clinical implications have remained elusive.[41] Numerous large-scale genomic or transcriptomic studies have been performed,[39,42-46] some of which predicted therapeutic targets, but in a more recent study *RB1* was the only gene associated with clinical outcome.[47] Essentially, despite all the efforts, an improved 37-gene signature for stratification[48] was the most tangible result of these studies, which are all based on common high-frequency alterations. Similarly, for other solid cancers, many gene "markers" and "signatures" have been described which have value for prognosis and patient stratification, but did not lead to a therapeutic breakthrough.

7 | THE NEED FOR A SYSTEMIC INDIVIDUALIZED ANALYSIS OF CANCER WITH INTEGRATED DRUG TARGET PREDICTION FOR COMBINATORIAL THERAPIES

We have previously argued in favor of a personalized analysis of deregulated networks in cancer.[8] This premise was based on the observation of large intertumoral/interindividual variation of signaling pathways in a very well-defined subset of prostate patients, including functionally divergent paralogues like the many *Wnt* genes and their receptors, and leading to different cross-talks between pathways, not even considering the potential impact of tumor-specifically altered expression of noncoding regulatory RNAs.[8] Patient-specific networks have also been derived from an integrative genomic, transcriptomic and phosphoproteomic study which compared treatment-naïve and metastatic prostate cancer.[49]

There is no doubt that cancer at the time of diagnosis, is an advanced disease with complex temporally deteriorating genetic lesions. While lesions may have a different impact on the various communication modules within the overall deregulated network (eg, deregulated *Myc* is present in several "hallmarks" in Figure 2), many other genes are directly or indirectly connected, forming a deregulated network, thereby affecting the overall communication that ultimately drives the hallmarks of cancer. We propose to comprehend the deregulated communication network in cancer analogously to the "omnigenic" model for complex genetic traits formulated by Boyle, Li and Pritchard.[50] This model considers that most traits are first directly affected by a modest number of "core genes" with specific roles in disease etiology. In the cancer case this could correspond to key nodes, such as oncogenes, tumor suppressors, but also other (cell/tissue-specific) master regulators. Secondly, upon realizing that the "core genes" contribute only a small part to the total heritability, the "omnigenic" model proposes that a large number of genes could make nonzero contributions in the context of cell regulatory networks. In the cancer case, these genes would be those that are interconnected within the cancer vs normal deregulated communication networks.

Such a deregulated communication network is a formalistic tool that can be mathematically described and extracted. It is unlimited to input data. While it uses genes as nodes, the direct and indirect effects of other molecules and systems that communicate with genes and their products can be integrated, including signaling molecules, metabolomes, epigenomes, chromatin landscapes and architecture. "Multiomics" approaches are frequently used and even the information encoded in chromatin architecture can be integrated into gene regulatory networks.[51] In addition, spatial and single-cell genomics can be added to provide information about tumor architecture and clonality.

This network analysis reveals also the large differences between individual cancers, even if they belong to the same type. In Figure 2 we show the deregulated network of cancer hallmark genes for patient 1 (Figure 2A) and patient 5 (Figure 2B), both of which are classified as castration-resistant ERG-positive prostate cancers. It is immediately obvious, that both the number and the genes vary dramatically in the two cases. Indeed, we have observed that even for the same signaling pathway, the affected genes (mutated, deregulated) are often distinct, and that common deregulation is rather an exception.[8] As a consequence, also the drugs that have been developed—albeit not necessarily for prostate cancer—and are available, vary for the two cancers (Table 1).

Indeed, a salient feature of these networks is the integration of information on targets for existing drugs, especially approved ones and those at clinical trial stage. We have incorporated this information in two prostate cancer patient-specific deregulated networks in Figure 2 (green colored nodes). Such information could assist the testing/development of combinatorial therapies, ideally in patient-derived cell systems. The efficacy of therapies can be monitored by liquid biopsy-based genomics analyses, which together with single-cell/spatial genomics of the primary tumor (clones) will provide information on drug response/escape and clonal tumor evolution.

8 | CONCLUSIONS

Our vision is an international human gene network consortium that assembles, generates and integrates the various levels of information to establish networks governing the development and homeostasis of normal cells and tissues and compares them to diseased cells and tissues to define the deregulated networks. This would be a multidisciplinary effort of clinicians, and cancer, developmental, cell and computational biologists, together with mathematicians and informaticians. The goal would be to develop a huge dynamic database of communication networks at scales from cells to organs to the entire organism. Existing

high-quality information in databases and from various consortia, including the HCA and LifeTime Initiative, will already provide an enormous resource to start with. Various omics, including metabolomics, should be integrated. Individual (patient-matched) cancer data would be compared in this resource with normal cell/organ information and differential networks established. Such an effort, even with the inclusion of millions of additional personalized omics analyses would be negligible in cost, compared to the 1 trillion bipartisan infrastructure bill that has been signed by the US president in November 2021. This effort should include strong informatics components for machine learning/artificial intelligence-assisted modeling,[52] the development of cell fate trajectories in healthy and diseased states from single-cell and spatial genomics and comprise the computational prediction, design and testing of combinatorial therapies.

ACKNOWLEDGMENTS

We thank Geneviève Almouzni, Marco Mendoza-Parra, Susan Chan and Philippe Kastner for encouragement and suggestions.

CONFLICT OF INTEREST

The authors declare no conflicts of interest.

ORCID

Yasenya Kasikci https://orcid.org/0000-0002-4714-4867
Hinrich Gronemeyer https://orcid.org/0000-0001-9454-2449

REFERENCES

1. Bianconi E, Piovesan A, Facchin F, et al. An estimation of the number of cells in the human body. *Ann Hum Biol*. 2013;40:463-471.
2. Borgatti SP, Mehra A, Brass DJ, Labianca G. Network analysis in the social sciences. *Science*. 2009;323:892-895.
3. Kadushin C. *Understanding Social Networks: Theories, Concepts and Findings*. New York: Oxford University Press; 2012:251.
4. Bullmore E, Sporns O. Complex brain networks: graph theoretical analysis of structural and functional systems. *Nat Rev Neurosci*. 2009; 10:186-198. doi:10.1093/oso/9780198805090.001.0001
5. Newman M. *Networks*. Second Ed., Oxford: Oxford University Press; 2018. doi:10.1093/oso/9780198805090.001.0001
6. Bondy JA, Murty USR. *Graph Theory with Applications*. London: Macmillan; 1976.
7. Zhang W, Chien J, Yong J, Kuang R. Network-based machine learning and graph theory algorithms for precision oncology. *npj Precis Oncol*. 2017;1:25.
8. Kumar A, Kasikci Y, Badredine A, et al. Patient-matched analysis identifies deregulated networks in prostate cancer to guide personalized therapeutic intervention. *Am J Cancer Res*. 2021;11:5299-5318.
9. Logan CY, Nusse R. The Wnt signaling pathway in development and disease. *Annu Rev Cell Dev Biol*. 2004;20:781-810.
10. La GM. Connaissance est un réseau. *Cah Numér*. 2014;10:37-54.
11. Nebbioso A, Tambaro FP, Dell'Aversana C, Altucci L. Cancer epigenetics: moving forward. *PLoS Genet*. 2018;14:e1007362.
12. Dawson MA, Kouzarides T. Cancer epigenetics: from mechanism to therapy. *Cell*. 2012;150:12-27.
13. Cholley PE, Moehlin J, Rohmer A, et al. Modeling gene-regulatory networks to describe cell fate transitions and predict master regulators. *npj Syst Biol Appl*. 2018;4:29.
14. Guerra C, Mijimolle N, Dhawahir A, et al. Tumor induction by an endogenous K-ras oncogene is highly dependent on cellular context. *Cancer Cell*. 2003;4:111-120.
15. Mainardi S, Mijimolle N, Francoz S, Vicente-Duenas C, Sanchez-Garcia I, Barbacid M. Identification of cancer initiating cells in K-Ras driven lung adenocarcinoma. *Proc Natl Acad Sci U S A*. 2014;111:255-260.
16. Kandoth C, McLellan MD, Vandin F, et al. Mutational landscape and significance across 12 major cancer types. *Nature*. 2013;502: 333-339.
17. Bailey MH, Tokheim C, Porta-Pardo E, et al. Comprehensive characterization of cancer driver genes and mutations. *Cell*. 2018;173: 371-85.e18.
18. Armenia J, Wankowicz SAM, Liu D, et al. The long tail of oncogenic drivers in prostate cancer. *Nat Genet*. 2018;50:645-651.
19. Greaves M, Maley CC. Clonal evolution in cancer. *Nature*. 2012;481: 306-313.
20. Grunwald BT, Devisme A, Andrieux G, et al. Spatially confined sub-tumor microenvironments in pancreatic cancer. *Cell*. 2021;184:5577–5592.
21. Greaves M. Evolutionary determinants of cancer. *Cancer Discov*. 2015;5:806-820.
22. McGranahan N, Swanton C. Clonal heterogeneity and tumor evolution: past, present, and the future. *Cell*. 2017;168:613-628.
23. Birkbak NJ, McGranahan N. Cancer genome evolutionary trajectories in metastasis. *Cancer Cell*. 2020;37:8-19.
24. Pavlova NN, Thompson CB. The emerging hallmarks of cancer metabolism. *Cell Metab*. 2016;23:27-47.
25. Lopez-Otin C, Blasco MA, Partridge L, Serrano M, Kroemer G. The hallmarks of aging. *Cell*. 2013;153:1194-1217.
26. Lopez-Otin C, Kroemer G. Hallmarks of health. *Cell*. 2021;184: 1929-1939.
27. Mittelbrunn M, Kroemer G. Hallmarks of T cell aging. *Nat Immunol*. 2021;22:687-698.
28. Hanahan D, Weinberg RA. Hallmarks of cancer: the next generation. *Cell*. 2011;144:646-674.
29. Hanahan D, Weinberg RA. The hallmarks of cancer. *Cell*. 2000;100: 57-70.
30. Kuperstein I, Bonnet E, Nguyen HA, et al. Atlas of cancer signalling network: a systems biology resource for integrative analysis of cancer data with Google maps. *Oncogenesis*. 2015;4:e160.
31. Regev A, Teichmann SA, Lander ES, et al. The human cell atlas. *Elife*. 2017;6:e67886. doi:10.7554/eLife.67886. PMID: 34612202
32. Snyder MP, Lin S, Posgai A, et al. The human body at cellular resolution: the NIH human biomolecular atlas program. *Nature*. 2019;574:187-192.
33. Rajewsky N, Almouzni G, Gorski SA, et al. LifeTime and improving European healthcare through cell-based interceptive medicine. *Nature*. 2020;587:377-386.
34. de Thé H, Pandolfi PP, Chen Z. Acute promyelocytic leukemia: a paradigm for oncoprotein-targeted cure. *Cancer Cell*. 2017;32:552-560. doi:10.1016/j.ccell.2017.10.002
35. Apperley JF. Chronic myeloid leukaemia. *Lancet*. 2015;385:1447-1459.
36. Deininger MW, Goldman JM, Melo JV. The molecular biology of chronic myeloid leukemia. *Blood*. 2000;96:3343-3356.
37. Greaves M. A causal mechanism for childhood acute lymphoblastic leukaemia. *Nat Rev Cancer*. 2018;18:471-484.
38. Curtis C, Shah SP, Chin SF, et al. The genomic and transcriptomic architecture of 2000 breast tumours reveals novel subgroups. *Nature*. 2012;486:346-352. doi:10.1038/nature10983
39. Robinson D, van Allen EM, Wu YM, et al. Integrative clinical genomics of advanced prostate cancer. *Cell*. 2015;161:1215-1228. doi:10.1016/j.cell.2015.06.053
40. Mendiratta G, Ke E, Aziz M, Liarakos D, Tong M, Stites EC. Cancer gene mutation frequencies for the US population. *Nat Commun*. 2021;12:5961.
41. Farashi S, Kryza T, Clements J, Batra J. Post-GWAS in prostate cancer: from genetic association to biological contribution. *Nat Rev Cancer*. 2019;19:46-59.

42. Baca SC, Prandi D, Lawrence MS, et al. Punctuated evolution of prostate cancer genomes. *Cell*. 2013;153:666-677.
43. Barbieri CE, Baca SC, Lawrence MS, et al. Exome sequencing identifies recurrent SPOP, FOXA1 and MED12 mutations in prostate cancer. *Nat Genet*. 2012;44:685-689.
44. Grasso CS, Wu YM, Robinson DR, et al. The mutational landscape of lethal castration-resistant prostate cancer. *Nature*. 2012;487:239-243. doi:10.1038/nature11125
45. Haffner MC, Mosbruger T, Esopi DM, et al. Tracking the clonal origin of lethal prostate cancer. *J Clin Invest*. 2013;123:4918-4922.
46. Mancuso N, Gayther S, Gusev A, et al. Large-scale transcriptome-wide association study identifies new prostate cancer risk regions. *Nat Commun*. 2018;9:4079.
47. Abida W, Cyrta J, Heller G, et al. Genomic correlates of clinical outcome in advanced prostate cancer. *Proc Natl Acad Sci U S A*. 2019;116:11428-11436.
48. You S, Knudsen BS, Erho N, et al. Integrated classification of prostate cancer reveals a novel luminal subtype with poor outcome. *Cancer Res*. 2016;76:4948-4958.
49. Drake JM, Paull EO, Graham NA, et al. Phosphoproteome integration reveals patient-specific networks in prostate cancer. *Cell*. 2016;166:1041-1054.
50. Boyle EA, Li YI, Pritchard JK. An expanded view of complex traits: from polygenic to omnigenic. *Cell*. 2017;169:1177-1186.
51. Malysheva V, Mendoza-Parra MA, Blum M, Spivakov M, Gronemeyer H. Gene regulatory network reconstruction incorporating 3D chromosomal architecture reveals key transcription factors and DNA elements driving neural lineage commitment. *bioRxiv*. 2019;303842. doi:10.1101/303842
52. Eraslan G, Avsec Z, Gagneur J, Theis FJ. Deep learning: new computational modelling techniques for genomics. *Nat Rev Genet*. 2019;20:389-403.

How to cite this article: Kasikci Y, Gronemeyer H. Complexity against current cancer research: Are we on the wrong track? *Int. J. Cancer*. 2022;150(10):1569-1578. doi:10.1002/ijc.33912

PUBLICATION N°3

Role of miRNAs in the differential inter-individual gene-regulatory networks in prostate cancer

Yasenya Kasikci, et al.

In preparation

ROLE OF miRNAs IN THE DIFFERENTIAL INTER-INDIVIDUAL GENE-REGULATORY NETWORKS IN PROSTATE CANCER

Yasenya Kasikci[1,2,3,4], et al.

[1]Institut de Génétique et de Biologie Moléculaire et Cellulaire (IGBMC), Department of Functional Genomics and Cancer, Illkirch, France

[2]Centre National de la Recherche Scientifique, UMR7104, Illkirch, France

[3]Institut National de la Santé et de la Recherche Médicale, U1258, Illkirch, France

[4]Université de Strasbourg, Illkirch, France

Yasenya Kasikci

E-mail: kasikciy@igbmc.fr

Phone: +(33) 7 82 97 72 94

Running Title: miRNA-Mediated Differential Gene Regulation in Prostate Cancer

Key words: DEmiRs, DEGs, GRN, prostate cancer, personalized medicine

Abstract: Prostate cancer is a common malignancy affecting a significant portion of the male population worldwide. The complexity and heterogeneity of this disease require a comprehensive understanding of the underlying molecular mechanisms driving its progression and treatment resistance. Cancer deregulation can occur not only through direct mutations in genes that regulate the genome, such as structural mutations or mutations in promoters or enhancers but also through deregulations in long non-coding RNAs (lncRNAs) and microRNAs (miRNAs), which can affect the expression of mRNA and proteins indirectly. In this study, we investigated the role of deregulated ncRNAs in the context of differential inter-individual gene regulatory networks (GRNs) established for each patient. Through the integration of high-throughput omics data, including gene and non-coding RNA expression profiles from a cohort of primary ERG-positive prostate cancer patients, a systematic analysis was performed to identify dysregulated ncRNAs among individuals. Functional enrichment analyses were conducted to gain insight into the involved biological pathways and GRNs were constructed by integrating deregulated miRNA-target interactions. Validation experiments were performed with dual-luciferase assays to confirm direct targeting between miRNAs and their putative target genes. Additionally, quantitative real-time PCR was performed to validate expression levels of selected miRNA target genes after exposing the LnCaP cell line to agomirs and antagomirs. Validation experiments provided evidence for the functional relevance of the identified mRNA-miRNA interaction network, indicating the potential utility of these targets in personalized miRNA network-based diagnosis and treatment of prostate cancer. By enhancing our understanding of the molecular mechanisms of deregulation in prostate cancer, these findings offer valuable insights for the development of personalized therapeutic approaches, reinforcing the correlation with our hypothesis.

INTRODUCTION

Cancer is a complex and heterogeneous disease with numerous genetic and environmental causes. Prostate cancer is the most frequently diagnosed malignancy and the fifth leading cause of cancer-related deaths among men worldwide[1]. The heterogeneity of prostate cancer tumors characterized by individual differences highlights the importance of molecular markers in facilitating diagnosis and predicting the prognosis. The occurrence of disease often results from complex systems involving multiple levels, rather than being attributed to the alteration of a single molecule. Therefore, the detection of a single molecular biomarker from groups of patients with the same disease is often not reliable in a set of heterogeneous patients because of the complexity of the disease[2]. To address this challenge, one potential solution lies in the identification of network biomarkers, which include the altered molecules (nodes) as well as their links to others (edges)[3]. Network analysis provides a comprehensive perspective beyond individual disease genes and offers insight into key regulatory pathways that contribute to disease pathogenesis. Although the impact of individual genes may appear minimal, their collective influence can significantly impact regulatory pathways.

MicroRNAs (miRNAs) are small, noncoding regulatory RNA molecules of 19 to 25 nucleotides that play a crucial role in the regulation of gene expression[4]. By binding to complementary sites in the 3' untranslated regions (3'UTRs) of specific target mRNAs, these regulatory RNAs post-transcriptionally control gene expression, resulting in mRNA degradation. Over 2000 human miRNAs have been identified, regulating up to 60% of gene transcripts. A significant proportion of these miRNAs are located inside or near regions associated with cancer (such as fragile sites, minimal regions of loss of heterozygosity, minimal regions of amplification, and common breakpoints[5]), making them highly relevant to tumor growth, invasion, and metastasis by either increasing the activation of oncogenic pathways or restricting the expression of tumor suppressors[5]. Advances in deep sequencing, high-throughput screening, and chip technologies have provided valuable insights into the dysregulation of miRNAs in cancer cells. The characterization of miRNAs in human tumors has not only enhanced our understanding of carcinogenesis but also paved the way for the identification of novel tumor markers and the development of innovative therapies[6].

It is noteworthy that miRNAs can act as either oncogenes or tumor suppressors, depending on the specific tissue in which they are expressed[7], as demonstrated by several well-studied examples such as miR-17[8], miR-21[9], miR-221/222[10-13], let-7[14-16], miR-15/16[17,18], miR-200[19-21], and miR-34[22-25]. Dysregulation of miRNAs has been linked to the hallmark features of cancer[26], and restoration of expression of certain critical downregulated miRNAs has been investigated as a potential treatment for various cancers[27-29]. Increasing the understanding of the molecular dysregulation underlying carcinogenesis has created opportunities to use miRNAs as diagnostic and prognostic indicators. Signature of miRNAs such as miR-10b in glioma[30], miR-29 in head and neck squamous cell carcinoma[31], and miR-34 in lung[32-34] and prostate cancer[35] have been identified and studied in clinical trials. Research on miRNA-based cancer treatments is attracting increasing attention due to miRNAs ability to target multiple effectors involved in cell differentiation, proliferation, and survival[36].

Dysregulation of miRNA expression may lead to disruption of tumor-related gene expression, thereby affecting tumor development and progression[37]. In prostate cancer, understanding the miRNA-mRNA interactions and their involvement in various signaling pathways is critical. In fact, prostate cancer is known to share numerous signaling pathways with other cancers such as Wnt, TNF and TGF-beta pathways, and also has unique pathways such as androgen receptor (AR) pathway that play a crucial role in its development[38]. In these complicated processes, miRNAs participate in many key cellular pathways and their interactions with many mRNAs can be leveraged when designing therapies to effectively disrupt cancer cell signaling networks[39].Advancements in bioinformatics and gene expression analysis have facilitated the construction of gene networks specific to cancer samples, allowing for the identification of potential oncogenic and tumor-suppressive miRNAs[40]. By analyzing miRNA targets, a large number of possible miRNA-mRNA interactions can be identified and ranked through network analysis, providing valuable insights into the disease mechanisms.

In this study, we used a network-based approach that leverages multiple types of ‘omics data' to identify functionally relevant disease targets. We performed extensive duplicate strand-specific total (Ribo-minus) RNA-seq and analyzed both coding and non-coding RNAs. Through bioinformatic analysis, we identified deregulated miRNAs and mRNAs by comparing tumor (T) vs

normal (N) samples and we generated personalized differential miRNA-mRNA networks for each patient to show inter-individual heterogeneity. Through experimental validation, we confirmed the direct interaction between miRNAs and target genes, supporting the development of personalized miRNA-based therapies.

MATERIALS AND METHODS

1. Patient Sample Collection

Samples were collected within 15 minutes after radical prostatectomy to minimize delays between de-vascularization and freezing and preserve labile molecules. Punch biopsies ("carrots") of 8 mm diameter were taken from tumor and adjacent normal tissue, snap-frozen in liquid nitrogen, and stored at -80°C. Carrots were cut into sequential sections for genomic and transcriptomic studies, and tumor cellularity was monitored by histological staining to ensure homogeneity of tumor and normal sections stored in LoBind tubes at -80°C.

2. Tissue Microarrays

Paraffin-embedded prostate cancer and adjacent normal tissue cores were used to create tissues microarrays (TMA). EZ-TMA™ Manual Tissue Microarray Kit was used in a semi-automated way to perform the TMA on the histopathology platform of the Biological Resource Center (CRB) of the Toulouse University Hospital. Two cores of 2 mm in diameter from two representative tumor areas and two adjacent normal tissues of each selected prostatic sample were included in paraffin recipient blocs. Immunohistochemistry was done using an automated Dako Autostainer with ERG, EZH2, and Androgen Receptor antibodies. The results were analyzed under an optical microscope by two pathologists (CM and M-LQ), blinded to the clinical data after digitalization of the slides using a Hamamatsu NanoZoomer slide scanner (Japan) at 20× magnification with a resolution of 0.46 microns per pixel.

3. Whole Exome Sequencing (WES)

WES, we utilized the QIAamp DNA micro kit according to the manufacturer's instructions to isolate DNA from both frozen tumor and matched normal tissue. GATC Biotech processed the DNA for exome capture, library preparation, and sequencing, using the SureSelectXT Human all

exon V6 kit to capture exons, preparing libraries with the TruSeq DNA library preparation kit, and performing Paired-End 125-base sequencing on an Illumina HiSeq 2000. We used the Genome Analysis Toolkit (GATK, 3.7) with default parameters to process the FastQ files provided by GATC Biotech for variant discovery. To assist with WES analysis, we created a WES Analysis Pipeline written in Python3 and managed with the Snakemake (3.13.3) tool and followed the authors' instructions to use the Genome Analysis Toolkit (GATK, 3.7). For a detailed description of the analysis pipeline, please refer to the comprehensive documentation available in the American Journal of Cancer Research[41].

4. RNA-sequencing

RNA was extracted from frozen tumor and matched normal tissue using Trizol reagent (Invitrogen), and two independent sets of adjacent 10 μm sections (N=10) were used to generate biological duplicate RNA-seq data. Consistent data between the biological duplicates were retained for subsequent analysis. The RNA was then purified using the RNeasy MinElute Cleanup Kit and sent to GATC Biotech (Konstanz, Germany) for strand-specific, paired-end, and Ribo-minus total RNA-seq. To prepare the RNA-seq libraries, ribosomal RNA depletion was performed with the Ribo Zero gold kit (Illumina Inc.), and libraries were created using the TruSeq stranded total RNA library prep kit (Illumina Inc.). The samples were sequenced using Illumina HiSeq 2000 with either paired end 125 base or 150 base reads, and the resulting FastQ files were used for further analysis. For a detailed description of the analysis pipeline, please refer to the comprehensive documentation available in the American Journal of Cancer Research[41].

5. Differential miRNA and mRNA Analysis

The patient-specific differential gene expression analysis was done using DESeq2 (version 1.20.0)[42], following a general set of steps with specific parameters. The analysis involved matching raw read counts from normal/tumor duplicates as input. The following steps were employed: First, rows with a sum of counts equal to 0 were removed. CooksCutoff was disabled, and an alpha value of 0.01 was used. Genes were then selected based on an adjusted p-value of ≤ 0.01. Additionally, genes with a log2 fold change (Log2FC) ≤ -1 or ≥ +1 were subsetted. Subsequently, the resulting list of Differentially Expressed Genes (DEGs), lncRNAs (DElncRs) and

miRNAs (DEmiRs) for each patient was obtained (Supplementary File 1). Subsequently, these were utilized for further analysis and investigation.

6. Identification of miRNA Targets

The predicted target genes of microRNAs were obtained from the miRDB online database[43]. This particular miRNA target prediction program relies on support vector machines (SVMs) and high-throughput training datasets. It considers a predicted target with a prediction score greater than 80 to be highly likely to be accurate. Therefore, only genes with a target score above 80 were chosen for each miRNA and a list of predicted targets of patient-specific deregulated miRNAs was generated (Supplementary File 2).

7. Construction of miRNA-mRNA Regulatory Networks

To generate gene networks for individual patients, we extracted a list of mutated genes from whole exome sequencing (WES) data and a list of differentially expressed genes (DEGs) from RNA-seq data, comparing tumor samples to normal samples for each patient. To explore the interactions among these genes, we utilized STRING, a Protein-Protein Interaction (PPI) database, and CellNet, a gene regulatory network (GRN) database. For the STRING database, we merged the list of genes (including DEGs and mutated genes) while retaining the attribute information indicating whether a gene was differentially expressed or mutated. Duplicated genes were removed, and the merged list was queried against the STRING database using an in-house script. We only considered edge interactions that had been experimentally validated, as denoted by a non-zero exp_score. Regarding the CellNet database, we queried only the differentially expressed genes on the target genes and retrieved along the cognate transcription factors. We selected interactions that had a z-score ≥5. After obtaining the networks from both databases, we incorporated additional information from the WES and RNA-seq data, specifying whether the genes were mutated, differentially expressed or both. This information was integrated with the existing database information. The individual master networks generated by using CellNet and STRING for each patient were visualized using Cytoscape. Next, we proceeded to query the target genes of all patient-specific deregulated miRNAs within Cytoscape, using the patient-specific networks we had generated earlier. This step involved identifying and extracting the target genes

of the deregulated miRNAs from the patient-specific networks, allowing us to focus specifically on the interactions between these miRNAs and their respective target genes. By incorporating this information into Cytoscape, we were able to visualize and analyze the interactions between the deregulated miRNAs and their target genes within the context of the patient-specific networks.

8. Gene Function Enrichment Analysis

For each patient, pathway enrichment analysis was performed on the target genes of deregulated miRNAs. This analysis was conducted using the ShinyGO web server (version 0.77), a comprehensive annotation database that combines information from Ensembl and STRING-db. During the analysis, a significance threshold of FDR < 0.05 was applied to determine which pathways were considered significant. The fold enrichment scores were used to assess the significance of the resulting pathway enrichments. Only categories with a minimum of ten overlapping genes were selected.

9. Cell Lines and Cell Culture

Human prostate cancer cells (PC3) were cultured with F-12K Medium (Kaighn's Modification of Ham's F-12 Medium) containing 10% fetal bovine serum, 40 µg/mL gentamicin. Androgen-dependent human prostate cancer (LNCaP) cells were cultured in RPMI 1640 w/10mM HEPES supplemented with 10% fetal calf serum, 1mM PyrNa and 40 µg/mL gentamicin. All cells were maintained in a 5% CO2-humidified atmosphere at 37°C.

10. RNA Oligonucleotides and Cell Transfection

The miR-27a-3p inhibitor, miR-27a-3p mimic, miR-23a-3p inhibitor, miR-23a-3p mimic, the inhibitor scrambled control and the mimic scrambled control were purchased from GenePharma Company (Shanghai, China). Cell transfection was performed by using HiPerfect Transfection Reagent (Qiagen) according to the manufacturer's protocols. LNCaP cells were cultured in 6-well plates and transiently transfected with miR-27a-3p inhibitor at a final concentration of 50 nM; miR-27a-3p mimic at a final concentration of 5 nM; miR-23a-3p inhibitor at a final concentration of 50 nM; miR-23a-3p mimic at a final concentration of 5 nM. At 24h

post transfection, the transfected cells were collected for quantitative real-time PCR analysis. Primer sequences are provided in Supplementary Table 1.

11. RNA Isolation and Quantitative Real-Time PCR Analysis

Total RNA was isolated by using TRIzol reagent (Invitrogen) according to the manufacturer's instructions. All samples were then treated with DNase I (Thermo Fisher) to remove contaminated genomic DNA from the purified RNA samples, followed by RNA cleanup using the RNeasy MinElute Cleanup Kit (Qiagen). RNA samples (1 μg each) were then reverse-transcribed into cDNA using the SuperScript™ IV First-Strand Synthesis Kit (Invitrogen). Quantitative real-time PCR reaction was performed by using the LightCycler® 480 SYBR Green I Master kit and device (Roche). Relative quantities of mRNA levels were determined using the 2-ΔΔCT method after normalization with GAPDH as a standard reference. All samples were performed in triplicate and the average CT was normalized by subtracting the average CT of GAPDH (ΔCT). All samples were compared to the nontreated LNCaP cell line by subtracting the experimental ΔCT from the ΔCT value of the treated LNCaP cell line as the calibrator. Data are presented as the mean ± standard deviation and statistical analyses were conducted in the Microsoft Excel software platform, using an independent Student's t-test, assuming equal variance between the two groups. Statistical significance was defined as $p \leqq 0.05$.

12. Dual Luciferase Reporter Activity

pmirGLO-GATA6 3'-UTR and pmirGLO-TOX3 3'-UTR vectors were constructed by inserting the amplified 3'-UTR of human GATA6 (513-1091 bp) and TOX3 (153-528 bp) into pmirGLO luciferase reporter plasmid, separately. The cloning process involved the use of SacI (NEB) and XbaI (NEB) restriction enzymes to cut both the vector and the genes, followed by ligation using T4 DNA ligase (Thermo Fisher) enzyme. The cloned vectors were then verified by DNA sequencing. For the luciferase assay, PC3 cells were seeded into 12-well plates and pmirGLO empty vector and pmirGLO-GATA6 3'-UTR vector co-transfected together with miR-27a mimic or negative control (NC) while pmirGLO empty vector and pmirGLO-TOX3 3'-UTR vector co-transfected together with miR-23a mimic and NC, by using Lipofectamine™ 2000 Transfection Reagent (Invitrogen). 24h after transfection, cells were lysed for luciferase activity measurement.

Luciferase assay was performed using the Dual-Luciferase reporter assay system (Promega). Reporter activity is expressed as the ratio of firefly to renilla activity. Each assay was repeated three times and the standard deviation within each sample is noted.

RESULTS and DISCUSSION

1. Profiling and Significance of Dysregulated miRNAs in ERG-positive Prostate Cancer Patients

We have previously analyzed prostate cancer patient samples with the aim of integrating patient-specific differentially expressed genes (DEGs) and mutated genes[41]. This integration involved utilizing data from protein-protein interaction and gene-regulatory network databases to construct patient-specific GRNs. To obtain highly characterized samples, we worked with 15 primary ERG-positive prostate tumors and matched normal tissue samples that were carefully selected by expert pathologists. The selection criteria ensured that consecutive sections of the same biopsy exhibited minimal differences in tumor cellularity, while the sections of normal biopsies from the same prostate showed no presence of tumor cells (0% tumor cells). To identify somatic variations, we performed whole-exome sequencing, which allowed us to examine the genetic differences between the cancer samples and their corresponding normal prostate tissue. By using the MuTect2 tool, we called variants and discovered a range of 49 to 114 mutations in each cancer sample compared to the corresponding normal tissue. Additionally, to identify deregulated gene expression levels, we conducted high-throughput strand-specific paired-end total RNA sequencing on the matched tumor and normal biopsy sections, ensuring biological duplicates. Prior to sequencing, we performed ribosomal RNA depletion to eliminate unwanted RNA. The analysis of the RNA-seq datasets comparing T and N samples revealed tumor-specific differentially expressed genes with diverse functions, including oncogenes, tumor suppressors, and transcription factors. Here, in addition to protein-coding genes, we focused on changes in the expression of specific regulatory RNAs, such as miRNAs and lncRNAs (Supplementary File 1). The log2 fold change values were calculated based on the average expression levels obtained from duplicate RNA sequencing (RNA-seq) of tumor and normal samples. Figure 1 illustrates the heat map analysis showing that different sets of DEmiRs and DElncRs deregulated in each patient. The heat map uses a color scale to represent the expression levels of DEmiRs and DElncRs. Genes

showing increased expression in tumor samples are represented in red, indicating upregulation, while genes with decreased expression in tumor samples are shown in blue, indicating downregulation. The varying color intensities and patterns provide insights into the expression and heterogeneity of these non-coding RNA molecules across the patient.

Several of these ncRNAs are known to play roles in the progression and treatment of prostate cancer. For instance, the lncRNA LINC00844 acts as a tumor suppressor and is down-regulated in metastatic prostate cancer cells, indicating that LINC00844 mechanistically facilitates AR binding to chromatin and its expression is crucial for promoting NDRG1 gene expression, which suppresses migration and invasion of prostate cancer[44]. It is also downregulated in 13 of the 15 PCa samples. Another lncRNA, HOTTIP, promotes prostate tumor proliferation and triggers cisplatin resistance. Knocking down HOTTIP inhibits the Wnt pathway, leading to cell death, cell cycle arrest, and increased sensitivity to cisplatin in prostate cancer cells[45]. HOTTIP is also downregulated in 9 of the 15 PCa samples. Furthermore, overexpression of LINC00665 is associated with reduced overall survival in prostate cancer patients, indicating its tumor-promoting role and potential as a prognostic and diagnostic tool[46]. It is also overexpressed in 12 of the 15 PCa samples. The lncRNA Nuclear enriched abundant transcript 1 (NEAT1) is involved in tumorigenesis, functioning as a p53 target gene that supports oncogene-expressing cells[47]. High NEAT1 expression correlates with poor prognosis in various cancers, including prostate cancer. In prostate cancer, elevated NEAT1 expression confers resistance to androgen receptor antagonists, leading to a worse disease outcome[48]. It is also overexpressed in 7 of the 15 PCa samples. The Prostate Cancer Associated Transcript 1 (PCAT-1) was identified as a cell proliferation activator in prostate cancer and has been detected in other types of cancer as well[49]. It is also implicated in DNA damage repair by reducing the RNA levels of the tumor suppressor BRCA2[50]. Additionally, PCAT-1 overexpressing cells showed increased sensitivity to Poly ADP Ribosyl Phosphatase (PARP) inhibitors, commonly used to treat tumors with BRCA1 or BRCA2 mutations. Thus, PCAT-1 expression may serve as a prognostic marker and an indicator of PARP sensitivity[50]. It is upregulated in 12 of the 15 PCa samples. Furthermore, PCA3 was approved by the FDA as a diagnostic test for prostate cancer, making it the first FDA-approved test based on lncRNA. PCA3 has become a useful noninvasive test for prostate cancer, with

improved specificity, positive predictive value, and negative predictive value compared to serum PSA testing[51]. It has also upregulated in 12 of the 15 samples.

A similar patient-specific divergence was seen in putative tumor suppressor and oncogenic miRNAs that are considered active for clinical development. For instance, miR-205 consistently shows downregulation in prostate cancer, in line with its reported tumor-suppressive role. Ectopic expression of miR-205 has been found to diminish the Androgen Receptor (AR) and Mitogen-Activated Protein Kinase (MAPK) signaling pathways, leading to repress tumor development in prostate cancer[52,53]. It is downregulated in 13 of the 15 PCa samples. Surprisingly, miR-222 has been identified as an oncogene in neural crest cancers, including glioblastoma multiforme[54] and thyroid papillary carcinoma[13] while in prostate cancer it shows a tumor-suppressing role[11] and is downregulated in 9 out of 15 PCa samples. In prostate and bladder cancer, miR-27a functions as a putative tumor suppressor by modulating proliferation and epithelial-mesenchymal transition[55]. It is downregulated in 4 of the 15 PCa samples. In a study on let-7b, its upregulation in tumor-associated macrophages (TAMs) was found to contribute to their tumor-promoting role by affecting the expression of inflammatory cytokines. Inhibiting let-7b reduces angiogenesis and mobility of prostate cancer cells, suggesting its potential as a modulator for macrophage polarization[56]. It is upregulated in 3 out of 15 PCa samples Additionally, miR-34a acts as a suppressor in various tumors by targeting genes involved in cell cycle control and apoptosis, similar to p53-mediated phenotypes[23,25]. While miR-34a was downregulated in 3 out of 15 PCa samples, it was unexpectedly upregulated in one patient. While certain miRNAs and lncRNAs exhibited deregulation across most patients, there were specific miRNAs and lncRNAs that showed deregulation only in individual patients. For instance, miR1251-Patient11(P11), miR155-(P9), miR17-(P12), miR200c-(P12), and miR561-(P6) were found to be deregulated in specific patients. Furthermore, different sets of DEmiRs and DElncRs were observed in each patient, emphasizing the unique molecular profiles exhibited by each individual. Additionally, some miRNAs, such as miR34a and miR646, were downregulated in certain patients but upregulated in others, further underscoring the heterogeneity of miRNA expression patterns within the patient population. The significantly different patterns of altered gene expression emphasize the importance of understanding the complex deregulated systems

in each patient. This understanding is crucial for identifying key targets in critical signaling pathways and essential nodes in networks.

2. Unveiling Regulatory Landscapes: Functional Insights from Dysregulated miRNAs

After identifying the sets of deregulated miRNAs in each patient, we analyzed their corresponding targets using a computational prediction algorithm such as miRDB. miRDB is an online database that uses a bioinformatics tool called MirTarget to predict miRNA targets. This tool was developed by analyzing a vast number of miRNA-target interactions derived from high-throughput sequencing experiments. Within this database, common features associated with miRNA binding and target downregulation have been identified and utilized in predicting miRNA targets through machine-learning methods like support vector machines (SVMs). All the predicted targets in miRDB are assigned target prediction scores ranging between 50 and 100, as determined by the new computational target prediction algorithm. A higher score indicates greater confidence in the prediction. Notably, a predicted target with a prediction score exceeding 80 is highly likely to represent a real target. Consequently, in our experiment, we exclusively considered target genes with prediction scores exceeding 80. By examining the targets of dysregulated miRNAs in each patient, we sought to identify the specific genes and pathways influenced by these dysregulated miRNAs in individual cases.

After generating a list of predicted target genes for each deregulated miRNA in the patients, we conducted a cross-referencing analysis with our dataset of DEGs to identify and classify DEmiRs-target genes for each patient. Subsequently, pathway enrichment analysis was performed on the target genes of deregulated miRNAs. For this analysis, we utilized the Panther pathway database on ShinyGO web server version 0.77, which is a comprehensive annotation database integrating information from Ensembl and STRING-db. A significance threshold of FDR < 0.05 was applied during the analysis to determine the pathways considered significant and only categories with a minimum of ten overlapping genes were selected for further analysis and consideration.

While analyzing the deregulated pathways in each patient, we noted that certain pathways, such as Cadherin, Integrin, and Wnt pathways, were consistently deregulated across

a majority of the patients[41]. However, it is important to note that each patient also displayed deregulation in different sets of pathways, indicating the personalized nature of these alterations. To gain a better understanding of the similarities and differences in these deregulated genes and pathways among the patients, we conducted a more detailed examination of selected pathways. Figure 2 illustrates the list of deregulated pathways for four patients (P1, P6, P9, P12) and focuses on the representation of deregulated genes within the Wnt pathway since the Wnt pathway emerged as a common deregulated pathway in all four patients. We observed distinct patterns of dysregulated miRNAs and their target genes within the miRNA-mRNA interaction Wnt networks. For instance, in P9, miR2052 and let7b were found to be upregulated, while their target genes were downregulated. The deregulated target genes of miR2052 were identified as FZD7 and FZD3, whereas the deregulated target genes of let7b included FZD3, MEF2C, ZNF827, and ACTA1. Similarly, in P12, miR25 was found to be upregulated, and its target genes included FZD10 and ACTC1. The deregulation of these specific genes and miRs play important roles in regulating Wnt signaling pathway, which is known to be involved in cancer development and progression. FZD3, FZD7 and FZD10 are cell surface receptors that play a crucial role in transducing Wnt signaling. Activation of Frizzled receptors by Wnt ligands initiates a signaling cascade that regulates various cellular processes, including cell proliferation, differentiation, and migration. ACTA1 and ACTC1 participate in the organization and dynamics of signaling complexes, influencing the transmission of Wnt signals. Dysregulation of ACTA1 and ACTC1 can disrupt the cytoskeletal architecture and affect the proper functioning of the Wnt pathway[57]. Furthermore, other important deregulated genes, such as ZNF827 and MEF2C, which function as transcription factors, have also been identified. These genes play crucial roles in regulating the expression of various target genes and contribute significantly to the observed molecular changes[58].

However, the significance of miRNA therapy lies in the possibility that a single miRNA may be responsible for the deregulation of multiple genes. Instead of focusing on each gene individually, we can potentially target multiple genes by focusing on miRNAs like let7b. This highlights one of the most important features of miRNA-targeted therapy, where a single miRNA

can regulate multiple genes and pathways, offering a potential avenue for therapeutic intervention.

3. Unraveling Complexity: Construction and Analysis of miRNA-mRNA Regulatory Networks

To gain a comprehensive understanding of the miRNA-mRNA interactions, it is important to expand our focus beyond specific pathways, such as the Wnt pathway. While studying pathway-specific dysregulation may provide valuable insights into these specific mechanisms, it might restrict the exploration of the broader miRNA-mRNA interactions. By integrating the differentially expressed targets of all deregulated miRNAs, we can construct a more extensive network, enabling the identification of shared regulatory mechanisms that extend beyond individual pathways. This comprehensive approach reveals common regulatory elements and potential master regulators that govern multiple pathways, leading to a more holistic understanding of the underlying molecular processes. Therefore, in our study, we constructed patient-specific networks that incorporate the differentially expressed predicted targets of all deregulated miRNAs. First, we identified the potential target genes of the miRNAs by using the miRNA target prediction database[43] and then these target genes were subsequently integrated into the patient-specific network, allowing us to assess their differential expression patterns within the tumors.

Figure 3 illustrates the miRNA-mRNA interaction networks for two distinct patients, providing a visual representation of the intricate regulatory relationships between miRNAs and their target genes within each patient's specific context. The miRNA-mRNA interaction networks illustrate the nodes and edges that represent miRNAs and their corresponding target genes, respectively. Each node represents a specific miRNA or target gene, while the edges denote the regulatory interactions between them. The connections between nodes indicate the miRNA-mediated regulation of target genes, highlighting the complexity and interconnectedness of these regulatory relationships. By visualizing the miRNA-mRNA interaction networks for different patients, we can observe patient-specific variations in the regulatory landscape. These variations may arise from differences in miRNA expression profiles, alterations in target gene expression, or distinct regulatory mechanisms operating within each patient's tumor.

When we performed a more detailed analysis on patients within the network, we observed instances where miRNAs target genes that have diverse functional roles. For example, some of the target genes identified in P7 includes transcription factors such as GATA6 (by miR-27a), TOX3 (by miR-23a) or ONECUT2 (by miR22, miR-27a). These transcription factors play critical roles in gene expression regulation and can have wide-ranging effects on cellular processes. When targeted by miRNAs, the expression levels and activity of these transcription factors may be modulated, leading to downstream effects on the activation or repression of other genes within the regulatory network. Similarly, in P12, our analysis revealed the deregulation of several transcription factors, including SOX4 (by miR-4664), ONECUT2 (by miR-4697), FOXI1 (by miR-4664), and SALL3 (by miR-646). By modulating the activity of these transcription factors, miRNAs can indirectly impact the expression of downstream genes, leading to cascading effects on cellular processes and potentially contributing to prostate cancer progression. Indeed, within the miRNA-mRNA interaction network, we also came across crucial genes such as tumor suppressors and oncogenes that exhibit intriguing expression patterns. Taking P15 as an example, we observed the downregulation of tumor suppressor genes ITIH5 and DCC, along with the unexpected downregulation of the MET oncogene. Tumor suppressor genes play a vital role in maintaining genomic stability and inhibiting tumor development. The downregulation of ITIH5 and DCC suggests a potential disruption of their tumor-suppressive functions in this particular patient.

The observed variations in miRNA expression patterns among patients reflect the heterogeneity of prostate cancer at the molecular level. Different patients exhibit unique miRNA expression profiles, which in turn lead to distinct miRNA-mRNA target interactions within their respective networks. These patient-specific miRNA-mRNA interactions contribute to the intricate regulatory complexity underlying prostate cancer progression. Moreover, the diverse sets of deregulated miRNAs and their target genes among patients suggest the involvement of multiple pathways and biological processes in prostate cancer. Each patient's network exhibits unique regulatory modules and functional components that contribute to disease pathogenesis. Understanding these patient-specific regulatory elements can provide insights into the

underlying molecular processes driving prostate cancer and identify potential biomarkers and therapeutic targets tailored to each patient's unique profile.

However, as the network analyses conducted in this study involved predicted miRNA targets, it is essential to verify whether these targets are indeed the direct regulatory targets of the corresponding miRNAs and ascertain their regulation by miRNAs. In order to gain a deeper understanding of the functional implications of the identified miRNA-mRNA interactions, experimental validation through in-vitro tests becomes necessary. To address this objective, we aimed to perform a series of in-vitro experiments to provide experimental evidence supporting the predicted miRNA-mRNA interactions and to validate their regulatory relationship. By undertaking these experimental assays, we aimed to enhance the reliability and credibility of the network analyses by validating the predicted interactions in a biological context.

4. Validating miRNA-mRNA Interactions: Confirming the Building Blocks of Regulatory Networks

One commonly employed experimental approach for validating miRNA-mRNA interactions is the luciferase reporter assay. In our study, we investigated the regulatory relationship between miR-27a and the gene GATA6, as well as miR-23a and the gene TOX3. To determine whether these genes are indeed the target genes of their respective miRNAs, we employed the Dual-Luciferase reporter assay. The Dual-Luciferase assay involves the construction of a pmiRGLO reporter vector containing the putative target sequences (3’ UTR region of GATA6 or TOX3) of the miRNAs of interest downstream of a luciferase gene and Renilla control gene. Simultaneously, we introduced either miR-27a mimic or miR-23a mimic to mimic the overexpression of the respective miRNAs. As a control in our study, we included a negative control mimic to account for any non-specific effects of the transfection process. Additionally, we transfected separately the miR-27a and miR-23a mimics into an empty pmirGlo vector as unregulated vector control. This control allowed us to assess the specific effects of miR-27a and miR-23a overexpression on our target genes, independent of any potential regulatory elements present in the vector. By including this unregulated vector control, we were able to differentiate between the effects of the miRNA mimic itself and any potential regulatory elements present in the vector backbone. This control is essential for ensuring the specificity of the observed changes

in gene expression that can be attributed solely to the modulation of miR-27a or miR-23a levels. Following the transfection, we measured the luciferase activity in the transfected PC3 cells. By quantifying the luciferase activity, we were able to determine whether miR-27a and miR-23a directly target GATA6 and TOX3, respectively. A significant reduction in luciferase activity in PC3 cells transfected with the miRNA mimics, compared to the negative controls, indicated a direct interaction between the miRNA and the target gene, validating their regulatory relationship (Figure 4).

To further validate the regulatory effects of miRNAs on their target genes, modulation of miRNA expression levels through miRNA overexpression or knockdown techniques was performed. Although we aimed to explore the personalized aspect of miRNA regulation, our study was limited by the availability of only frozen tissue samples. Unfortunately, we could not utilize patient-derived xenografting (PDX) with primary tumor tissue as the patients no longer exhibited prostate cancer. Despite this substitution introducing certain limitations, we proceeded with the analysis to gain insights into the interplay between miRNAs and their targets. For this investigation, we focused on the LNCaP cell line as a surrogate, treating it as a representation of a new patient. By introducing miRNA mimics or inhibitors, we manipulated the expression levels of miR-27a and miR-23 in LNCaP cells. Then we compared the expression levels of the 8 predicted target genes (MAP7, SATB2, HNF4G, ESRP1, GATA6, ITGA8, SMOC2, SCG2) in cells with manipulated miRNA expression by RT-PCR (Figure 5).

During our detailed examination of the patients, we focused on functionally important transcription factors and discovered distinct patterns of deregulation for instance in the GATA6 and TOX3 genes. We observed that these deregulation patterns varied depending on DEmiRs. For example, while miR-27b upregulation led to the downregulation of GATA6 and its interaction with different genes, miR-27a downregulation resulted in GATA6 upregulation and its interaction with other target genes. A similar pattern was also observed for the TOX3 gene (Figure 5A).

Specifically, when we separately manipulated LNCaP cells using miR-27a and miR-23a mimics and inhibitors, we observed changes in the expression levels of some predicted target genes. The observed deregulation of these target genes provided further evidence of the

regulatory effects of miR-27a and miR-23a on gene expression. The observed deregulation of the target genes substantiated the regulatory effects of miR-27a and miR-23a on gene expression. However, it is noteworthy that certain discrepancies were observed between the RT-PCR analyses of patient samples and the LNCaP cell line. Specifically, SCG2 and ESRP1 exhibited opposite regulation. While SCG2 was downregulated in P3 and P7, it was upregulated in the LNCaP cell line. Conversely, ESRP1 showed upregulation in P2 and P7 but was downregulated in the LNCaP cell line.

This finding suggested that different miRNAs can deregulate target genes, leading to varying expression levels. This variation arises from the context-specific functions of miRNAs as tumor suppressors or oncogenes within specific tissues. As a result, the deregulated genes interact with different gene sets, contributing to the complexity of miRNA-mediated gene regulation. The observed context-dependent effects of miRNAs and their deregulated target genes provide insights into the molecular basis of cancer and suggest potential opportunities for personalized therapeutic interventions.

CONCLUSION

In conclusion, this study focused on the role of miRNAs in the differential inter-individual gene regulatory networks in prostate cancer. The heterogeneity of prostate cancer tumors emphasizes the importance of molecular markers for accurate diagnosis and prognosis. By utilizing a network-based approach and analyzing both coding and non-coding RNAs, the study identified deregulated miRNAs and mRNAs in prostate cancer patients. Through experimental validation, the direct interaction between miRNAs and target genes was confirmed, supporting the development of personalized miRNA-based therapies. Further research in this field is needed to explore the potential of miRNA-based therapies in prostate cancer and to uncover additional miRNA-mRNA interactions. By continuing to investigate the complex interplay between miRNAs and gene regulatory networks, we can advance our understanding of prostate cancer and pave the way for more effective personalized treatment strategies.

Acknowledgments

We thank to Marina Roussel to provide us pmirglo empty vector.

REFERENCES

1 Sung, H. *et al.* Global Cancer Statistics 2020: GLOBOCAN Estimates of Incidence and Mortality Worldwide for 36 Cancers in 185 Countries. *CA: a cancer journal for clinicians* **71**, 209-249, doi:10.3322/caac.21660 (2021).

2 Cava, C., Bertoli, G. & Castiglioni, I. In silico identification of drug target pathways in breast cancer subtypes using pathway cross-talk inhibition. *Journal of translational medicine* **16**, 154, doi:10.1186/s12967-018-1535-2 (2018).

3 Cava, C., Bertoli, G. & Castiglioni, I. Portrait of Tissue-Specific Coexpression Networks of Noncoding RNAs (miRNA and lncRNA) and mRNAs in Normal Tissues. *Computational and mathematical methods in medicine* **2019**, 9029351, doi:10.1155/2019/9029351 (2019).

4 Lin, S. & Gregory, R. I. MicroRNA biogenesis pathways in cancer. *Nature reviews. Cancer* **15**, 321-333, doi:10.1038/nrc3932 (2015).

5 Calin, G. A. *et al.* Human microRNA genes are frequently located at fragile sites and genomic regions involved in cancers. *Proceedings of the National Academy of Sciences* **101**, 2999-3004, doi:doi:10.1073/pnas.0307323101 (2004).

6 Rupaimoole, R. & Slack, F. J. MicroRNA therapeutics: towards a new era for the management of cancer and other diseases. *Nature reviews. Drug discovery* **16**, 203-222, doi:10.1038/nrd.2016.246 (2017).

7 Calin, G. A. & Croce, C. M. MicroRNA-cancer connection: the beginning of a new tale. *Cancer research* **66**, 7390-7394, doi:10.1158/0008-5472.can-06-0800 (2006).

8 O'Donnell, K. A., Wentzel, E. A., Zeller, K. I., Dang, C. V. & Mendell, J. T. c-Myc-regulated microRNAs modulate E2F1 expression. *Nature* **435**, 839-843, doi:10.1038/nature03677 (2005).

9 Fabbri, M. *et al.* MicroRNAs bind to Toll-like receptors to induce prometastatic inflammatory response. *Proceedings of the National Academy of Sciences* **109**, E2110-E2116, doi:doi:10.1073/pnas.1209414109 (2012).

10 Gillies, J. K. & Lorimer, I. A. Regulation of p27Kip1 by miRNA 221/222 in glioblastoma. *Cell cycle (Georgetown, Tex.)* **6**, 2005-2009, doi:10.4161/cc.6.16.4526 (2007).

11 Galardi, S. *et al.* miR-221 and miR-222 expression affects the proliferation potential of human prostate carcinoma cell lines by targeting p27Kip1. *The Journal of biological chemistry* **282**, 23716-23724, doi:10.1074/jbc.M701805200 (2007).

12 le Sage, C. *et al.* Regulation of the p27(Kip1) tumor suppressor by miR-221 and miR-222 promotes cancer cell proliferation. *The EMBO journal* **26**, 3699-3708, doi:10.1038/sj.emboj.7601790 (2007).

13 Visone, R. *et al.* MicroRNAs (miR)-221 and miR-222, both overexpressed in human thyroid papillary carcinomas, regulate p27Kip1 protein levels and cell cycle. *Endocrine-related cancer* **14**, 791-798, doi:10.1677/erc-07-0129 (2007).

14 Yanaihara, N. *et al.* Unique microRNA molecular profiles in lung cancer diagnosis and prognosis. *Cancer cell* **9**, 189-198, doi:10.1016/j.ccr.2006.01.025 (2006).

15 Takamizawa, J. *et al.* Reduced expression of the let-7 microRNAs in human lung cancers in association with shortened postoperative survival. *Cancer research* **64**, 3753-3756, doi:10.1158/0008-5472.can-04-0637 (2004).

16 Johnson, S. M. *et al.* RAS is regulated by the let-7 microRNA family. *Cell* **120**, 635-647, doi:10.1016/j.cell.2005.01.014 (2005).

17 Cimmino, A. *et al.* miR-15 and miR-16 induce apoptosis by targeting BCL2. *Proceedings of the National Academy of Sciences* **102**, 13944-13949, doi:doi:10.1073/pnas.0506654102 (2005).

18 Calin, G. A. *et al.* Frequent deletions and down-regulation of micro- RNA genes miR15 and miR16 at 13q14 in chronic lymphocytic leukemia. *Proceedings of the National Academy of Sciences of the United States of America* **99**, 15524-15529, doi:10.1073/pnas.242606799 (2002).
19 Gregory, P. A. *et al.* The miR-200 family and miR-205 regulate epithelial to mesenchymal transition by targeting ZEB1 and SIP1. *Nature cell biology* **10**, 593-601, doi:10.1038/ncb1722 (2008).
20 Bracken, C. P. *et al.* A double-negative feedback loop between ZEB1-SIP1 and the microRNA-200 family regulates epithelial-mesenchymal transition. *Cancer research* **68**, 7846-7854, doi:10.1158/0008-5472.can-08-1942 (2008).
21 Hao, J. *et al.* MicroRNA control of epithelial-mesenchymal transition in cancer stem cells. *International journal of cancer* **135**, 1019-1027, doi:10.1002/ijc.28761 (2014).
22 Hermeking, H. The miR-34 family in cancer and apoptosis. *Cell death and differentiation* **17**, 193-199, doi:10.1038/cdd.2009.56 (2010).
23 Raver-Shapira, N. *et al.* Transcriptional activation of miR-34a contributes to p53-mediated apoptosis. *Molecular cell* **26**, 731-743, doi:10.1016/j.molcel.2007.05.017 (2007).
24 Yamakuchi, M. & Lowenstein, C. J. MiR-34, SIRT1 and p53: the feedback loop. *Cell cycle (Georgetown, Tex.)* **8**, 712-715, doi:10.4161/cc.8.5.7753 (2009).
25 Chang, T. C. *et al.* Transactivation of miR-34a by p53 broadly influences gene expression and promotes apoptosis. *Molecular cell* **26**, 745-752, doi:10.1016/j.molcel.2007.05.010 (2007).
26 Hanahan, D. & Weinberg, R. A. The hallmarks of cancer. *Cell* **100**, 57-70, doi:10.1016/s0092-8674(00)81683-9 (2000).
27 Dhawan, A., Scott, J. G., Harris, A. L. & Buffa, F. M. Pan-cancer characterisation of microRNA across cancer hallmarks reveals microRNA-mediated downregulation of tumour suppressors. *Nature Communications* **9**, 5228, doi:10.1038/s41467-018-07657-1 (2018).
28 Lin, K. *et al.* Loss of MIR15A and MIR16-1 at 13q14 is associated with increased TP53 mRNA, de-repression of BCL2 and adverse outcome in chronic lymphocytic leukaemia. *British journal of haematology* **167**, 346-355, doi:10.1111/bjh.13043 (2014).
29 Lin, X., Yang, Z., Zhang, P. & Shao, G. miR-154 suppresses non-small cell lung cancer growth in vitro and in vivo. *Oncology reports* **33**, 3053-3060, doi:10.3892/or.2015.3895 (2015).
30 Sun, L. *et al.* MicroRNA-10b induces glioma cell invasion by modulating MMP-14 and uPAR expression via HOXD10. *Brain research* **1389**, 9-18, doi:10.1016/j.brainres.2011.03.013 (2011).
31 Yamamoto, N. *et al.* Tumor-suppressive microRNA-29a inhibits cancer cell migration and invasion via targeting HSP47 in cervical squamous cell carcinoma. *Int J Oncol* **43**, 1855-1863, doi:10.3892/ijo.2013.2145 (2013).
32 Kasinski, A. L. & Slack, F. J. miRNA-34 prevents cancer initiation and progression in a therapeutically resistant K-ras and p53-induced mouse model of lung adenocarcinoma. *Cancer research* **72**, 5576-5587, doi:10.1158/0008-5472.can-12-2001 (2012).
33 Wiggins, J. F. *et al.* Development of a lung cancer therapeutic based on the tumor suppressor microRNA-34. *Cancer research* **70**, 5923-5930, doi:10.1158/0008-5472.can-10-0655 (2010).
34 Stahlhut, C. & Slack, F. J. Combinatorial Action of MicroRNAs let-7 and miR-34 Effectively Synergizes with Erlotinib to Suppress Non-small Cell Lung Cancer Cell Proliferation. *Cell cycle (Georgetown, Tex.)* **14**, 2171-2180, doi:10.1080/15384101.2014.1003008 (2015).
35 Liu, C. *et al.* The microRNA miR-34a inhibits prostate cancer stem cells and metastasis by directly repressing CD44. *Nature medicine* **17**, 211-215, doi:10.1038/nm.2284 (2011).
36 Garzon, R., Marcucci, G. & Croce, C. M. Targeting microRNAs in cancer: rationale, strategies and challenges. *Nature reviews. Drug discovery* **9**, 775-789, doi:10.1038/nrd3179 (2010).
37 Jackson, B. L., Grabowska, A. & Ratan, H. L. MicroRNA in prostate cancer: functional importance and potential as circulating biomarkers. *BMC cancer* **14**, 930, doi:10.1186/1471-2407-14-930 (2014).

38 Suzuki, H., Ueda, T., Ichikawa, T. & Ito, H. Androgen receptor involvement in the progression of prostate cancer. *Endocrine-related cancer* **10**, 209-216, doi:10.1677/erc.0.0100209 (2003).

39 Li, D., Hao, X. & Song, Y. Identification of the Key MicroRNAs and the miRNA-mRNA Regulatory Pathways in Prostate Cancer by Bioinformatics Methods. *BioMed research international* **2018**, 6204128, doi:10.1155/2018/6204128 (2018).

40 Kumar, B. *et al.* Cell-type specific expression of oncogenic and tumor suppressive microRNAs in the human prostate and prostate cancer. *Scientific Reports* **8**, 7189, doi:10.1038/s41598-018-25320-z (2018).

41 Kumar, A. *et al.* Patient-matched analysis identifies deregulated networks in prostate cancer to guide personalized therapeutic intervention. *American journal of cancer research* **11**, 5299-5318 (2021).

42 Love, M. I., Huber, W. & Anders, S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biology* **15**, 550, doi:10.1186/s13059-014-0550-8 (2014).

43 Chen, Y. & Wang, X. miRDB: an online database for prediction of functional microRNA targets. *Nucleic acids research* **48**, D127-d131, doi:10.1093/nar/gkz757 (2020).

44 Lingadahalli, S. *et al.* Novel lncRNA LINC00844 Regulates Prostate Cancer Cell Migration and Invasion through AR Signaling. *Molecular cancer research : MCR* **16**, 1865-1878, doi:10.1158/1541-7786.mcr-18-0087 (2018).

45 Jiang, H. *et al.* Knockdown of the long noncoding RNA HOTTIP inhibits cell proliferation and enhances cell sensitivity to cisplatin by suppressing the Wnt/β-catenin pathway in prostate cancer. *Journal of cellular biochemistry* **120**, 8965-8974, doi:10.1002/jcb.27851 (2019).

46 Eke, I. *et al.* The lncRNAs LINC00261 and LINC00665 are upregulated in long-term prostate cancer adaptation after radiotherapy. *Molecular therapy. Nucleic acids* **24**, 175-187, doi:10.1016/j.omtn.2021.02.024 (2021).

47 Adriaens, C. *et al.* p53 induces formation of NEAT1 lncRNA-containing paraspeckles that modulate replication stress response and chemosensitivity. *Nature medicine* **22**, 861-868, doi:10.1038/nm.4135 (2016).

48 Chakravarty, D. *et al.* The oestrogen receptor alpha-regulated lncRNA NEAT1 is a critical modulator of prostate cancer. *Nat Commun* **5**, 5383, doi:10.1038/ncomms6383 (2014).

49 Prensner, J. R. *et al.* Transcriptome sequencing across a prostate cancer cohort identifies PCAT-1, an unannotated lincRNA implicated in disease progression. *Nature biotechnology* **29**, 742-749, doi:10.1038/nbt.1914 (2011).

50 Prensner, J. R. *et al.* PCAT-1, a long noncoding RNA, regulates BRCA2 and controls homologous recombination in cancer. *Cancer research* **74**, 1651-1660, doi:10.1158/0008-5472.can-13-3159 (2014).

51 Bussemakers, M. J. *et al.* DD3: a new prostate-specific gene, highly overexpressed in prostate cancer. *Cancer research* **59**, 5975-5979 (1999).

52 Majid, S. *et al.* MicroRNA-205-directed transcriptional activation of tumor suppressor genes in prostate cancer. *Cancer* **116**, 5637-5649, doi:10.1002/cncr.25488 (2010).

53 Hagman, Z. *et al.* miR-205 negatively regulates the androgen receptor and is associated with adverse outcome of prostate cancer patients. *British journal of cancer* **108**, 1668-1676, doi:10.1038/bjc.2013.131 (2013).

54 Zhang, C.-Z. *et al.* MiR-221 and miR-222 target PUMA to induce cell survival in glioblastoma. *Molecular Cancer* **9**, 229, doi:10.1186/1476-4598-9-229 (2010).

55 Wan, X. *et al.* Androgen-induced miR-27A acted as a tumor suppressor by targeting MAP2K4 and mediated prostate cancer progression. *The International Journal of Biochemistry & Cell Biology* **79**, 249-260, doi:10.1016/j.biocel.2016.08.043 (2016).

56 Wang, Z. *et al.* miRNA let-7b modulates macrophage polarization and enhances tumor-associated macrophages to promote angiogenesis and mobility in prostate cancer. *Sci Rep* **6**, 25602, doi:10.1038/srep25602 (2016).
57 Galli, C. *et al.* Actin cytoskeleton controls activation of Wnt/β-catenin signaling in mesenchymal cells on implant surfaces with different topographies. *Acta biomaterialia* **8**, 2963-2968, doi:10.1016/j.actbio.2012.04.043 (2012).
58 Pon, J. R. & Marra, M. A. MEF2 transcription factors: developmental regulators and emerging cancer genes. *Oncotarget* **7** (2015).

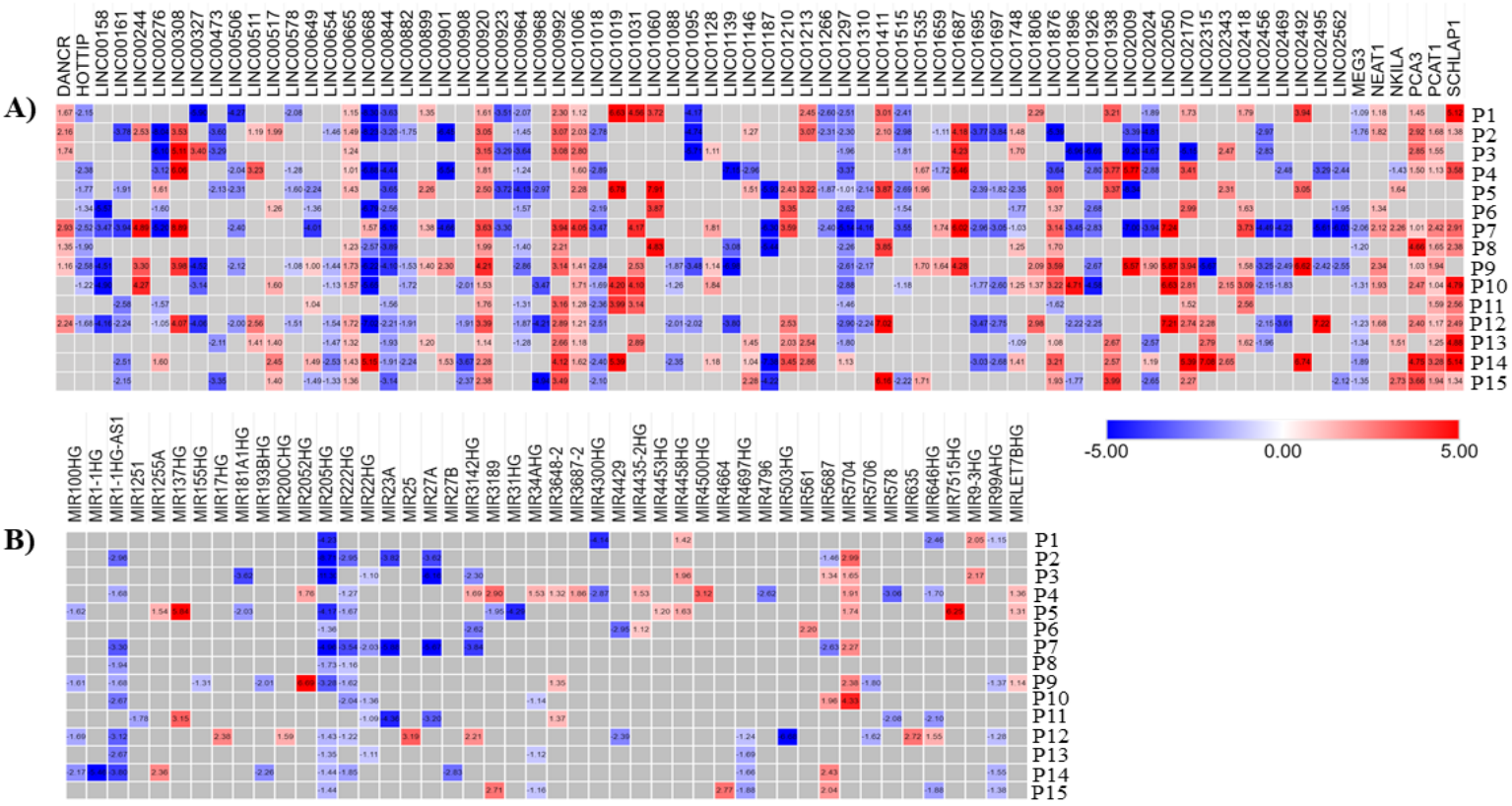
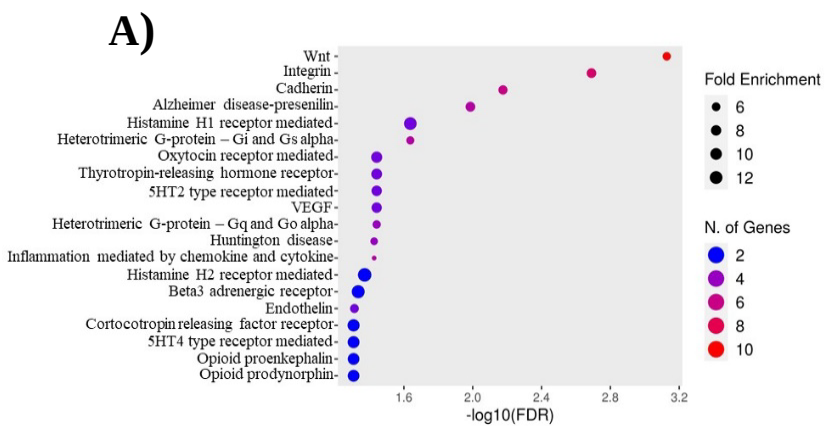
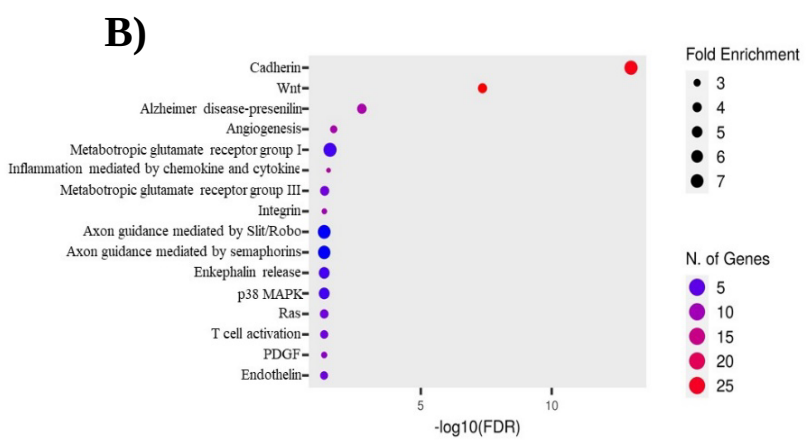
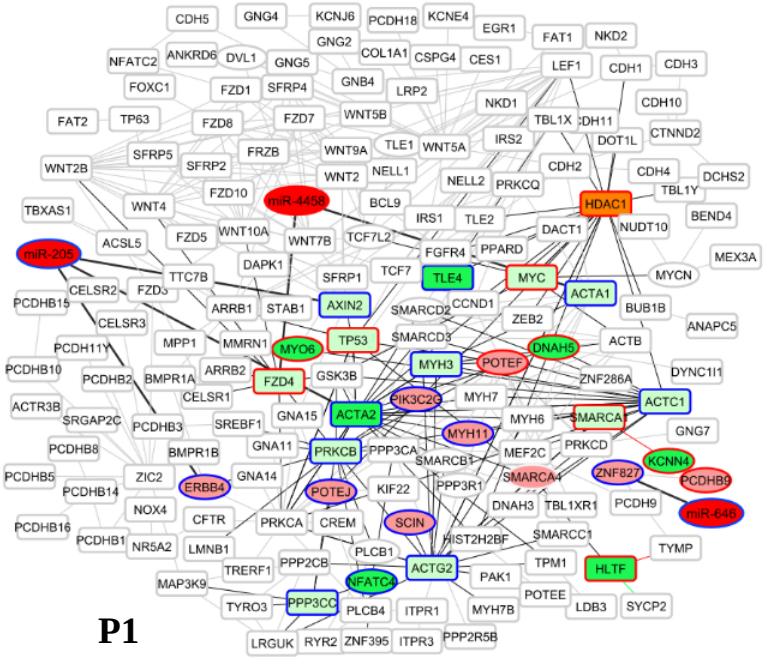
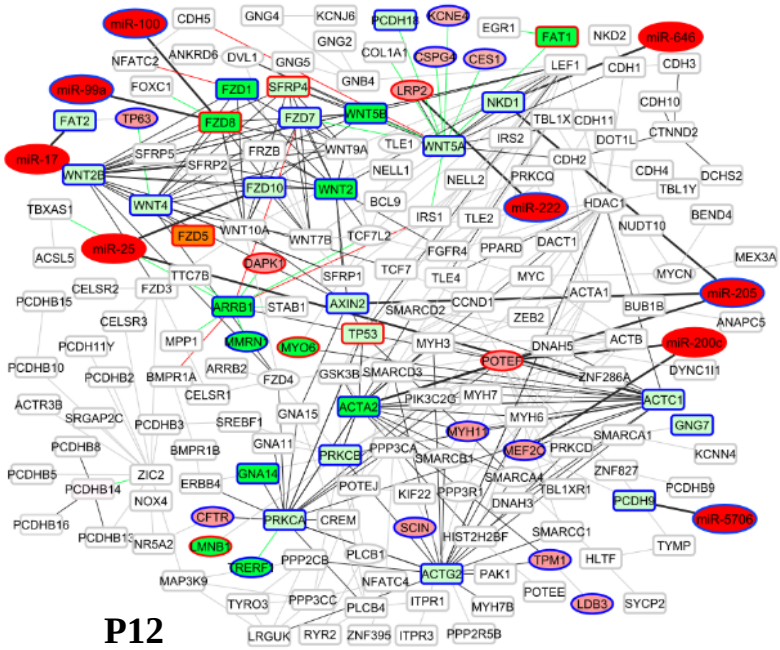


Figure 1 The log2 fold change heat map showing **A)** DElncRs, **B)** DEmiRs based on the average expression levels obtained from duplicate RNA sequencing (RNA-seq) of tumor and normal samples. Heat map analysis generated using the Morpheus software (https://software.broadinstitute.org/morpheus). In the color scheme, the red color represents up-regulation of expression, while the blue color represents down-regulation.









The figure legend is on the next page.

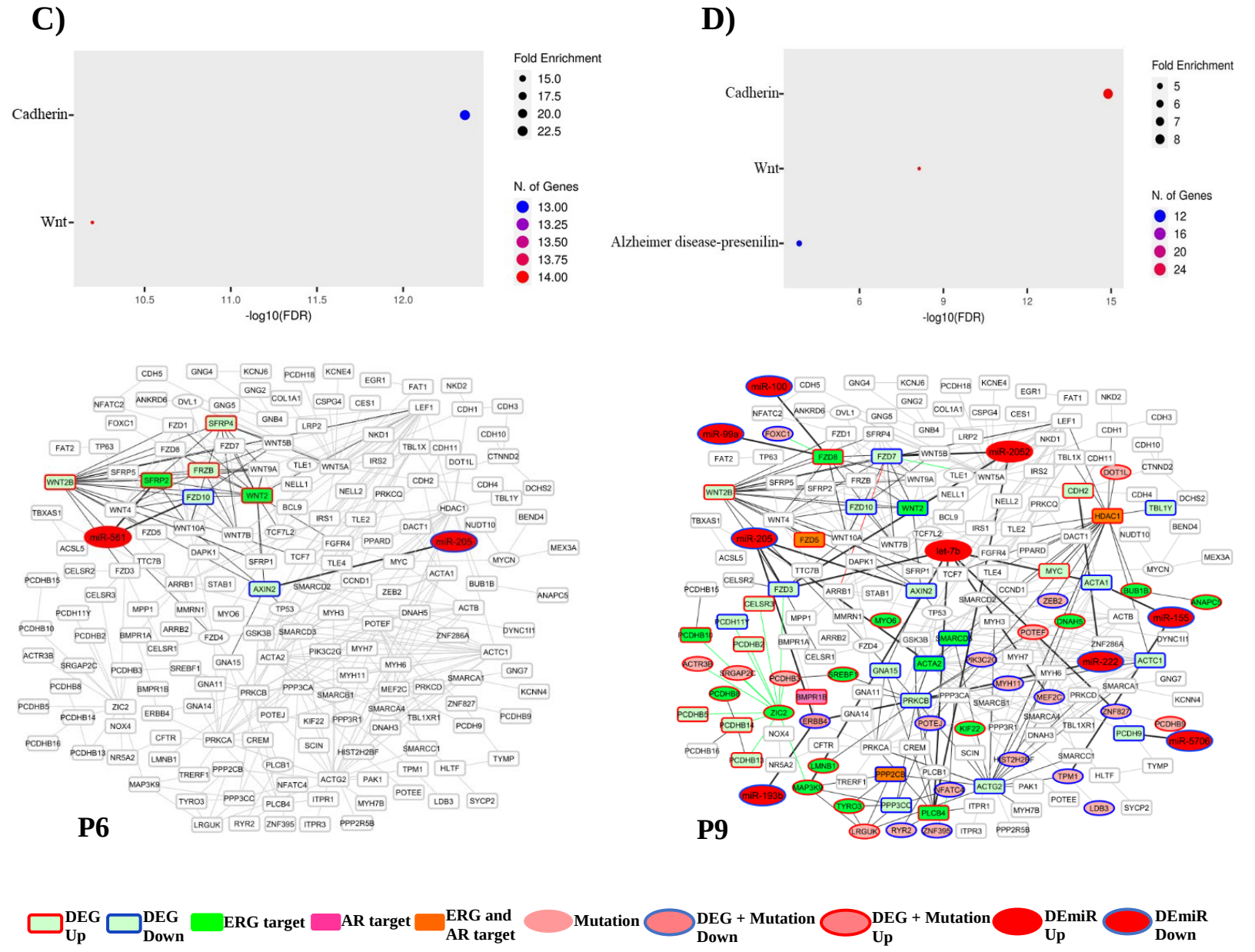
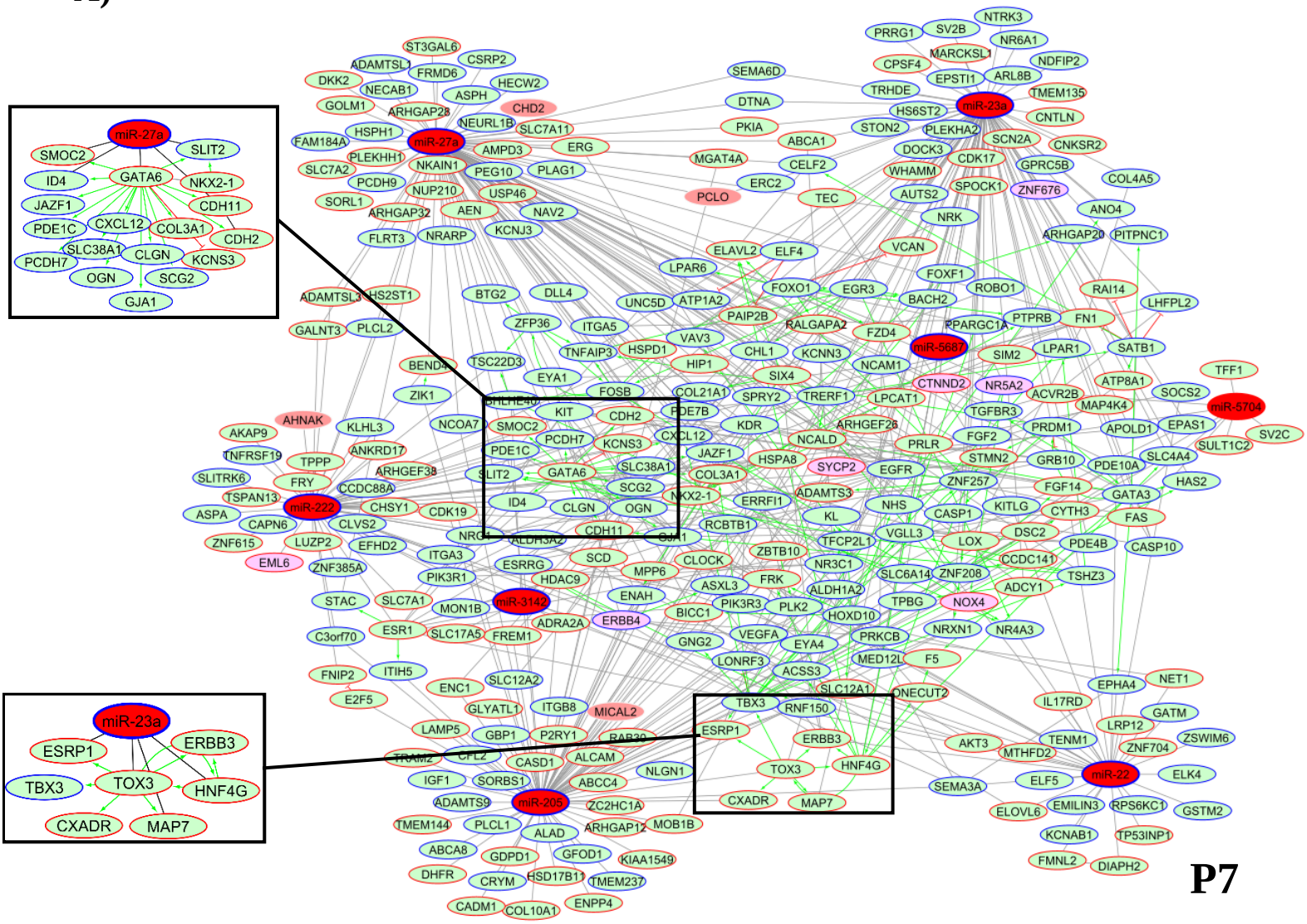


Figure 2 The Panther pathway analysis by ShinyGO (top) and the visualization of the patient-specific Wnt pathway networks (bottom) for **A)** Patient 1 **B)** Patient 12 **C)** Patient 6 **D)** Patient 9. When known, connectivities are displayed as green (activation) or red (inhibition) lines; unknown connectivities and protein-protein interactions are displayed as grey lines. DEG specifics, DEmiRs and mutations are color-coded as described below the figure.



The figure legend is on the next page.

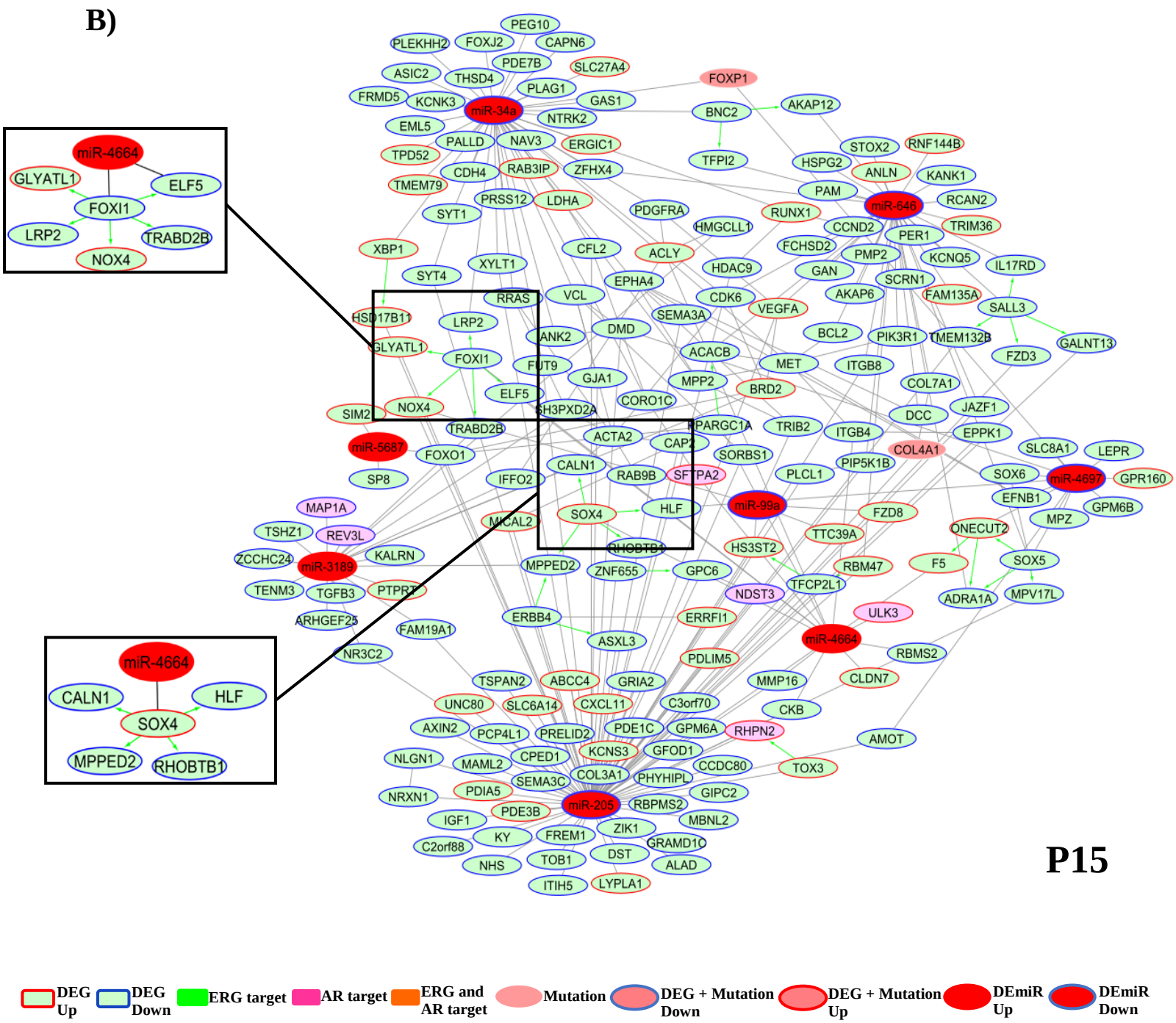


Figure 3 The figure represents a miRNA-mRNA interactions network for two different patients, **A)** Patient 7 and **B)** Patient 15. When known, connectivities are displayed as green (activation) or red (inhibition) lines; unknown connectivities and protein-protein interactions are displayed as grey lines. DEG specifics, DEmiRs and mutations are color-coded as described below the figure.

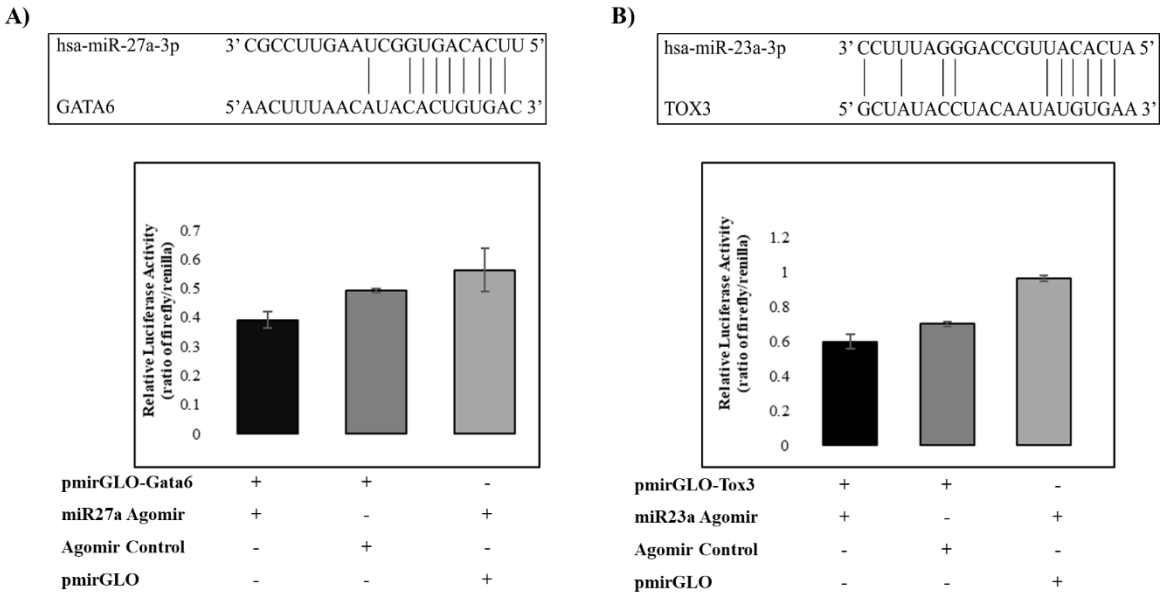


Figure 4 GATA6 is a direct target of miR-27a and TOX3 is a direct target of miR-23a in PC-3 prostate cancer cells. **A)** The figure shows the miR-27a binding site in the 3'-UTR of the human GATA6 gene. The top panel presents a schematic representation of this binding site. In the bottom panel, the GATA6-3'UTR-WT vector was co-transfected with either miR-27a mimic (pmirGLO-Gata6 + miR27a Agomir) or a non-targeting control miRNA (pmirGLO-Gata6 + Agomir Control) into PC-3 cells. Additionally, an unregulated vector control (pmirGlo without GATA6 3'UTR region) was co-transfected with miR-27a mimic (pmirGLO + miR27a Agomir) into PC3 cells. After 24 hours, firefly and renilla activities were measured (n=3). **B)** The same experimental procedure described in A) was applied for miR-23a and its target TOX3. The figure provides information about the miR-23a binding site in the TOX3 gene and presents the results of co-transfection of TOX3-3'UTR-WT vector with miR-23a mimic (pmirGLO-Tox3 + miR23a Agomir) and non-targeting control miRNA (pmirGLO-Tox3 + Agomir Control) and an unregulated vector control with miR-23a mimic (pmirGLO + miR23a Agomir) into PC-3 cells. Firefly and renilla activities were measured after 24 hours (n=3).

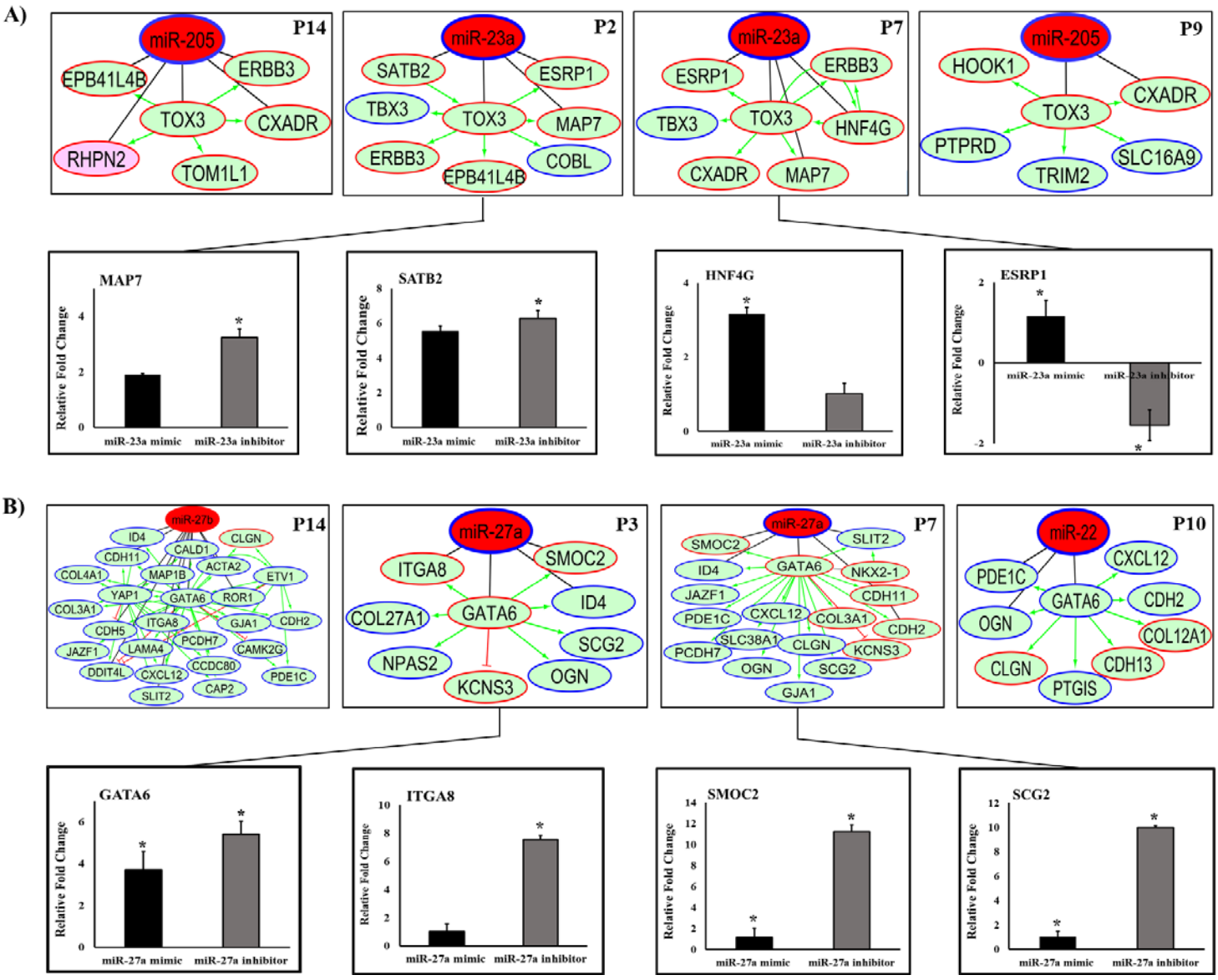


Figure 5 A) The top section illustrates the deregulation of the TOX3 transcription factor when targeted by different miRNAs in various patients, along with the affected genes. This analysis enables us to observe the diverse regulatory patterns of TOX3 under the influence of different miRNAs in individual patients. In the bottom section, we present the expression levels of the deregulated target genes using RT-PCR analysis after 24 hours of transfection with either the miR-23a mimic or the miR-23a inhibitor in Lncap cells. The mRNA expression values were normalized to the internal control GAPDH mRNA levels. Following the treatments, the fold changes in target mRNA levels in LNCaP cells were calculated by comparing them to the mRNA level in untreated LNCaP cells. The results are presented as means ± S.E.M. (n = 3 experiments). *P < 0.05.**B)** The same procedure was applied to the GATA6 gene, where LNCaP cells were manipulated either with the miR-27a mimic or the miR-27a inhibitor.

Name	**Sequence (5'->3')**
Primers used in quantitative RT-PCR	
GATA6-F	CCCAGACCACTTGCTATGAA
GATA6-R	GGAATTATTGCTATTACCAGAGC
SCG2-F	GGTGAGAGACAGCAAAGAGAA
SCG2-R	GACATCATCTGAGAGTTGGTC
ITGA8-F	GTTTCTACTGGCAAGGACAAGT
ITGA8-R	CAACTGAGTATCCAAGGTAACTG
SMOC2-F	GTGTCACAGCTACACGGGAT
SMOC2-R	GGAGCTGCGGCATCATCTG
MAP7-F	CACTAAACTGTACTCACCCGAC
MAP7-R	GTCTGGTTTATTTCCTGAGTGGT
SATB2-F	CAAGTCAGAGATGAGCTGAAG
SATB2-R	TCTCAGACAACAATCCCTGTG
ESRP1-F	CTTCCAAGGTTACCAGTGTTTG
ESRP1-R	AGCCTGAGTGTAGACAAACTTC
HNF4G-F	CGTTTATTCTTGCAGGTTCAGTC
HNF4G-R	CTTCTTCTGGTGCTTATTCTGTC
Primers used in plasmid construction	
GATA6-3'-UTR-F	AGCCGGATCAGCTTGCATGCAGCTCTAGAGCTACCTGC
GATA6-3'-UTR-R	AGCCGGATCAGCTTGCATGCAGCTCTAGAGCTACCTGC
TOX3-3'-UTR-F	TAGCGAGCTCGAGGAACTGCAGTGTAGCTGAG
TOX3-3'-UTR-R	TGCTCTAGAGCAGAGCTTTGGCAAAAGTCTGT

Supplementary Table 1 Primers used in quantitative RT-PCR and plasmid construction (Red highlights indicates restriction enzyme sites).

PERSPECTIVES AND CONCLUSIONS

General Conclusion

In general conclusion, our study sheds light on the immense complexity of cancer and raises important questions about the effectiveness of current approaches. The human body, with its intricate network of cells and information transfer, resembles a highly organized and regulated system. Understanding the dynamics of this network is crucial for comprehending the development and progression of cancer. So, in this study, we presented a comprehensive analysis of patient-matched data to identify DEGs and DEncRNAs in prostate cancer. By integrating genomic and transcriptomic datasets, we were able to generate patient-specific networks and identify key nodes involved in different deregulated pathways and also in the cross-talk between these pathways. Afterward, we validated the network analysis conducted using both experimental and in-silico methods. To enhance the complexity of the network, we incorporated miRNAs and target genes, expanding our understanding of the interconnected pathways. Furthermore, through experimental validation, we confirmed the direct interaction between miRNAs and target genes, providing evidence for the functional relevance of the identified miRNA-mRNA interaction network. Exploring the functional implications of these interactions and their involvement in key regulatory pathways provided deeper insights into the underlying molecular mechanism of prostate cancer.

In summary, the complexity of cancer research calls for a paradigm shift in general cancer therapy approach. By embracing the interconnectedness of networks and leveraging advanced computational methods, we can unlock new insights into cancer biology and pave the way for more effective diagnostic, therapeutic, and preventive strategies. Furthermore, the concept of targeting networks proposes that genes, which may not be deregulated but serve as crucial nodes within the deregulated network, could offer substantial value as potential targets for testing.

Perspectives

One important future perspective is the integration of multi-omics data to gain a more comprehensive understanding of the regulatory networks involved in prostate cancer. By combining gene expression profiles with other omics data such as DNA methylation, histone modifications, and protein expression, researchers can uncover additional layers of regulation and identify key molecular players in the disease. This integrative approach will provide a more holistic view of the molecular mechanisms underlying prostate cancer and may reveal novel therapeutic targets.

Second, the use of patient-derived organoid cultures could provide a valuable platform for testing novel molecules against putative therapeutic targets. This approach allows for more accurate and personalized screening of potential treatments, leading to more effective interventions.

Another future direction is the investigation of the functional implications of miRNA-mRNA interactions in prostate cancer. While this study validated the regulatory relationships between miRNAs and target genes, further experiments like Crispr-Cas9 gene knockdown or knockin are needed to elucidate the downstream effects of these interactions on cellular processes and signaling pathways. Understanding the functional consequences of miRNA dysregulation will not only enhance our knowledge of prostate cancer pathogenesis but also inform the development of targeted therapies.

Moreover, the integration of spatial genomics and single-cell analysis can provide a deeper understanding of tumor heterogeneity and aid in the design of personalized treatment strategies. Spatial genomics techniques, such as spatial transcriptomics and spatial proteomics, allow to study the spatial distribution of gene expression and protein levels within the tumor microenvironment. Similarly, single-cell analysis enables the examination of individual tumor cells, unveiling cellular diversity and heterogeneity within the tumor. Identifying rare cell populations, such as cancer stem cells or therapy-resistant clones, can be instrumental in designing targeted therapies that address specific cellular subpopulations. By combining spatial genomics and single-cell analysis, we can gain a more nuanced understanding of the tumor's spatial organization, cellular interactions, and the dynamic nature of the tumor microenvironment.

Furthermore, the complexity of cancer research is becoming enormous, as data from various (single cell) (epi)genomics, (epi/spatial)transcriptomics, (phospho)proteomics, metabolomics, lipidomics, and nuclear architecture alteration will far exceed what we are used to today. Such Big Data, however, is becoming more interpretable and manageable given the dramatic development of artificial intelligence, machine learning, and neuronal network algorithms, provided that the data are properly normalized and integrable, and the algorithms are efficiently trained.

In conclusion, these future perspectives offer the potential for advancing our comprehension of prostate cancer biology and formulating enhanced, personalized, and targeted treatment strategies. The ongoing exploration of these innovative avenues will

undoubtedly shape the landscape of cancer research and revolutionize clinical management for the benefit of patients worldwide.

BIBLIOGRAPHY

Bibliography

1 Sung, H. *et al.* Global Cancer Statistics 2020: GLOBOCAN Estimates of Incidence and Mortality Worldwide for 36 Cancers in 185 Countries. *CA: a cancer journal for clinicians* **71**, 209-249, doi:10.3322/caac.21660 (2021).

2 Euler, L. Solutio problematis ad geometrian situs pertinentis. *Coment. Acad. Sci. Petropolitanae* **8**, 128–140 (1736).

3 Winter, J., Jung, S., Keller, S., Gregory, R. I. & Diederichs, S. Many roads to maturity: microRNA biogenesis pathways and their regulation. *Nature cell biology* **11**, 228-234, doi:10.1038/ncb0309-228 (2009).

4 McNeal, J. E. The zonal anatomy of the prostate. *Prostate* **2**, 35-49, doi:10.1002/pros.2990020105 (1981).

5 Ittmann, M. Anatomy and Histology of the Human and Murine Prostate. *Cold Spring Harb Perspect Med* **8**, a030346, doi:10.1101/cshperspect.a030346 (2018).

6 Hanahan, D. & Weinberg, R. A. Hallmarks of cancer: the next generation. *Cell* **144**, 646-674, doi:10.1016/j.cell.2011.02.013 (2011).

7 Barron, D. A. & Rowley, D. R. The reactive stroma microenvironment and prostate cancer progression. *Endocrine-related cancer* **19**, R187-204, doi:10.1530/erc-12-0085 (2012).

8 Kong, Y. W., Ferland-McCollough, D., Jackson, T. J. & Bushell, M. microRNAs in cancer management. *Lancet Oncol* **13**, e249-258, doi:10.1016/s1470-2045(12)70073-6 (2012).

9 Siegel, D. A., O'Neil, M. E., Richards, T. B., Dowling, N. F. & Weir, H. K. Prostate Cancer Incidence and Survival, by Stage and Race/Ethnicity - United States, 2001-2017. *MMWR. Morbidity and mortality weekly report* **69**, 1473-1480, doi:10.15585/mmwr.mm6941a1 (2020).

10 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* **124**, 2801-2814, doi:10.1002/cncr.31549 (2018).

11 LeBlanc, A. G., Demers, A. & Shaw, A. Recent trends in prostate cancer in Canada. *Health Rep* **30**, 12-17, doi:10.25318/82-003-x201900400002-eng (2019).

12 Collin, S. M. *et al.* Prostate-cancer mortality in the USA and UK in 1975-2004: an ecological study. *Lancet Oncol* **9**, 445-452, doi:10.1016/s1470-2045(08)70104-9 (2008).

13 Pakzad, R., Mohammadian-Hafshejani, A., Ghoncheh, M., Pakzad, I. & Salehiniya, H. The incidence and mortality of prostate cancer and its relationship with development in Asia. *Prostate Int* **3**, 135-140, doi:10.1016/j.prnil.2015.09.001 (2015).

14 Loeb, S. *et al.* Overdiagnosis and overtreatment of prostate cancer. *Eur Urol* **65**, 1046-1055, doi:10.1016/j.eururo.2013.12.062 (2014).

15 Hsing, A. W. & Devesa, S. S. Trends and patterns of prostate cancer: what do they suggest? *Epidemiol Rev* **23**, 3-13, doi:10.1093/oxfordjournals.epirev.a000792 (2001).

16 McNeal, J. E. Normal histology of the prostate. *Am J Surg Pathol* **12**, 619-633, doi:10.1097/00000478-198808000-00003 (1988).

17 Isaacs, J. T. Prostatic structure and function in relation to the etiology of prostatic cancer. *Prostate* **4**, 351-366, doi:10.1002/pros.2990040405 (1983).

18 Roehrborn, C. G. Pathology of benign prostatic hyperplasia. *International journal of impotence research* **20** S11-18, doi:10.1038/ijir.2008.55 (2008).

19 Arrighi, H. M., Metter, E. J., Guess, H. A. & Fozzard, J. L. Natural history of benign prostatichyperplasia and risk of prostatectomy: The Baltimore Longitudinal Study of Aging. *Urology* **38**, 4-8, doi:10.1016/0090-4295(91)80191-9 (1991).

20 Young, J. M., Muscatello, D. J. & Ward, J. E. Are men with lower urinary tract symptoms at increased risk of prostate cancer? A systematic review and critique of the available evidence. *BJU international* **85**, 1037-1048, doi:10.1046/j.1464-410x.2000.00659.x (2000).

21 Plochocki, A. & King, B. Medical Treatment of Benign Prostatic Hyperplasia. *The Urologic clinics of North America* **49**, 231-238, doi:10.1016/j.ucl.2021.12.003 (2022).
22 Miernik, A. & Gratzke, C. Current Treatment for Benign Prostatic Hyperplasia. *Deutsches Arzteblatt international* **117**, 843-854, doi:10.3238/arztebl.2020.0843 (2020).
23 McNeal, J. E. & Yemoto, C. E. Spread of adenocarcinoma within prostatic ducts and acini. Morphologic and clinical correlations. *Am J Surg Pathol* **20**, 802-814, doi:10.1097/00000478-199607000-00003 (1996).
24 McNeal, J. E. & Bostwick, D. G. Intraductal dysplasia: a premalignant lesion of the prostate. *Human pathology* **17**, 64-71, doi:10.1016/s0046-8177(86)80156-3 (1986).
25 Bostwick, D. G. & Brawer, M. K. Prostatic Intra-Epithelial Neoplasia and Early Invasion in Prostate Cancer. *Cancer* **59**, 788-794, doi:10.1002/1097-0142(19870215)59:4<788::AID-CNCR2820590421>3.0.CO;2-I (1987).
26 Häggman, M., Macoska, J. A., Wojno, K. J. & Oesterling, J. E. The relationship between prostatic intraepithelial neoplasia and prostate cancer: critical issues. *The Journal of urology* **158 1**, 12-22, doi:10.1097/00005392-199707000-00004 (1998).
27 Alcaraz, A. *et al.* High-grade prostate intraepithelial neoplasia shares cytogenetic alterations with invasive prostate cancer. *Prostate* **47**, 29-35, doi:10.1002/pros.1044 (2001).
28 Alexander, E. E., Qian, J., Wollan, P. C., Myers, R. P. & Bostwick, D. G. Prostatic intraepithelial neoplasia does not appear to raise serum prostate-specific antigen concentration. *Urology* **47**, 693-698, doi:10.1016/S0090-4295(96)00004-0 (1996).
29 Epstein, J. I. & Herawi, M. Prostate Needle Biopsies Containing Prostatic Intraepithelial Neoplasia or Atypical Foci Suspicious for Carcinoma: Implications for Patient Care. *The Journal of Urology* **175**, 820-834, doi:10.1016/S0022-5347(05)00337-X (2006).
30 Herawi, M., Kahane, H., Cavallo, C. & Epstein, J. I. Risk of Prostate Cancer on First Re-Biopsy Within 1 Year Following a Diagnosis of High Grade Prostatic Intraepithelial Neoplasia is Related to the Number of Cores Sampled. *The Journal of Urology* **175**, 121-124, doi:10.1016/S0022-5347(05)00064-9 (2006).
31 Montironi, R. *et al.* Atypical foci suspicious but not diagnostic of malignancy in prostate needle biopsies. *European urology* **50** 666-674, doi:10.1016/j.eururo.2006.07.048 (2006).
32 Borboroglu, P. G., Comer, S. W., Riffenburgh, R. H. & Amling, C. L. Extensive repeat transrectal ultrasound guided prostate biopsy in patients with previous benign sextant biopsies. *J Urol* **163**, 158-162 (2000).
33 Moore, C. K. *et al.* Prognostic significance of high grade prostatic intraepithelial neoplasia and atypical small acinar proliferation in the contemporary era. *The Journal of urology* **173**, 70-72, doi:10.1097/01.ju.0000148260.69779.c5 (2005).
34 Humphrey, P. A. Histological variants of prostatic carcinoma and their significance. *Histopathology* **60**, 59-74, doi:10.1111/j.1365-2559.2011.04039.x (2012).
35 Calle, E. E., Rodriguez, C., Walker-Thurmond, K. & Thun, M. J. Overweight, obesity, and mortality from cancer in a prospectively studied cohort of U.S. adults. *The New England journal of medicine* **348**, 1625-1638, doi:10.1056/NEJMoa021423 (2003).
36 Siegel, R. L., Miller, K. D., Fuchs, H. E. & Jemal, A. Cancer Statistics, 2021. *CA: a cancer journal for clinicians* **71**, 7-33, doi:10.3322/caac.21654 (2021).
37 Ward, E. *et al.* Cancer disparities by race/ethnicity and socioeconomic status. *CA: a cancer journal for clinicians* **54**, 78-93, doi:10.3322/canjclin.54.2.78 (2004).
38 Lloyd, T. *et al.* Lifetime risk of being diagnosed with, or dying from, prostate cancer by major ethnic group in England 2008-2010. *BMC medicine* **13**, 171, doi:10.1186/s12916-015-0405-5 (2015).
39 Steinberg, G. D., Carter, B. S., Beaty, T. H., Childs, B. & Walsh, P. C. Family history and the risk of prostate cancer. *The Prostate* **17**, 337-347, doi:10.1002/pros.2990170409 (1990).

40 Hjelmborg, J. B. *et al.* The Heritability of Prostate Cancer in the Nordic Twin Study of Cancer. *Cancer Epidemiology, Biomarkers & Prevention* **23**, 2303-2310, doi:10.1158/1055-9965.epi-13-0568 (2014).
41 Schumacher, F. R. *et al.* Association analyses of more than 140,000 men identify 63 new prostate cancer susceptibility loci. *Nature genetics* **50**, 928-936, doi:10.1038/s41588-018-0142-8 (2018).
42 Eeles, R. A. *et al.* Identification of 23 new prostate cancer susceptibility loci using the iCOGS custom genotyping array. *Nature genetics* **45**, 385-391, 391e381-382, doi:10.1038/ng.2560 (2013).
43 Amundadottir, L. T. *et al.* A common variant associated with prostate cancer in European and African populations. *Nature genetics* **38**, 652-658, doi:10.1038/ng1808 (2006).
44 Dupont, W. D. *et al.* 8q24 genetic variation and comprehensive haplotypes altering familial risk of prostate cancer. *Nature Communications* **11**, 1523, doi:10.1038/s41467-020-15122-1 (2020).
45 Panagopoulos, I., Möller, E., Collin, A. & Mertens, F. The POU5F1P1 pseudogene encodes a putative protein similar to POU5F1 isoform 1. *Oncology reports* **20**, 1029-1033 (2008).
46 Sotelo, J. *et al.* Long-range enhancers on 8q24 regulate c-Myc. *Proceedings of the National Academy of Sciences of the United States of America* **107**, 3001-3005, doi:10.1073/pnas.0906067107 (2010).
47 Wasserman, N. F., Aneas, I. & Nóbrega, M. A. An 8q24 gene desert variant associated with prostate cancer risk confers differential in vivo activity to a MYC enhancer. *Genome research* **20 9**, 1191-1197 (2010).
48 Du, M. *et al.* Prostate cancer risk locus at 8q24 as a regulatory hub by physical interactions with multiple genomic loci across the genome. *Human molecular genetics* **24**, 154-166, doi:10.1093/hmg/ddu426 (2015).
49 Eeles, R. A. *et al.* Multiple newly identified loci associated with prostate cancer susceptibility. *Nature genetics* **40**, 316-321, doi:10.1038/ng.90 (2008).
50 Takata, R. *et al.* Genome-wide association study identifies five new susceptibility loci for prostate cancer in the Japanese population. *Nature genetics* **42**, 751-754, doi:10.1038/ng.635 (2010).
51 Thomas, G. *et al.* Multiple loci identified in a genome-wide association study of prostate cancer. *Nature genetics* **40**, 310-315, doi:10.1038/ng.91 (2008).
52 Gudmundsson, J. *et al.* Two variants on chromosome 17 confer prostate cancer risk, and the one in TCF2 protects against type 2 diabetes. *Nature genetics* **39**, 977-983, doi:10.1038/ng2062 (2007).
53 Edwards, S. M. *et al.* Two percent of men with early-onset prostate cancer harbor germline mutations in the BRCA2 gene. *American journal of human genetics* **72**, 1-12, doi:10.1086/345310 (2003).
54 Cuzick, J. *et al.* Prevention and early detection of prostate cancer. *Lancet Oncol* **15**, e484-492, doi:10.1016/s1470-2045(14)70211-6 (2014).
55 Leongamornlert, D. *et al.* Germline BRCA1 mutations increase prostate cancer risk. *British journal of cancer* **106**, 1697-1701, doi:10.1038/bjc.2012.146 (2012).
56 Kote-Jarai, Z. *et al.* BRCA2 is a moderate penetrance gene contributing to young-onset prostate cancer: implications for genetic testing in prostate cancer patients. *British journal of cancer* **105**, 1230-1234, doi:10.1038/bjc.2011.383 (2011).
57 Bancroft, E. K. *et al.* Targeted prostate cancer screening in BRCA1 and BRCA2 mutation carriers: results from the initial screening round of the IMPACT study. *Eur Urol* **66**, 489-499, doi:10.1016/j.eururo.2014.01.003 (2014).
58 Cagney, D. N. *et al.* The FDA NIH Biomarkers, EndpointS, and other Tools (BEST) resource in neuro-oncology. *Neuro-oncology* **20**, 1162-1172, doi:10.1093/neuonc/nox242 (2018).

59 Velonas, V. M., Woo, H. H., dos Remedios, C. G. & Assinder, S. J. Current status of biomarkers for prostate cancer. *International journal of molecular sciences* **14**, 11034-11060, doi:10.3390/ijms140611034 (2013).
60 Tomlins, S. A. *et al.* Recurrent fusion of TMPRSS2 and ETS transcription factor genes in prostate cancer. *Science (New York, N.Y.)* **310**, 644-648, doi:10.1126/science.1117679 (2005).
61 Hossain, D. & Bostwick, D. G. Significance of the TMPRSS2:ERG gene fusion in prostate cancer. *BJU international* **111**, 834-835, doi:10.1111/bju.12120 (2013).
62 Yu, J. *et al.* An integrated network of androgen receptor, polycomb, and TMPRSS2-ERG gene fusions in prostate cancer progression. *Cancer cell* **17**, 443-454, doi:10.1016/j.ccr.2010.03.018 (2010).
63 Gasi Tandefelt, D., Boormans, J., Hermans, K. & Trapman, J. ETS fusion genes in prostate cancer. *Endocrine-related cancer* **21**, R143-152, doi:10.1530/erc-13-0390 (2014).
64 Tomlins, S. A., Palanisamy, N., Siddiqui, J., Chinnaiyan, A. M. & Kunju, L. P. Antibody-based detection of ERG rearrangements in prostate core biopsies, including diagnostically challenging cases: ERG staining in prostate core biopsies. *Archives of pathology & laboratory medicine* **136**, 935-946, doi:10.5858/arpa.2011-0424-OA (2012).
65 Pettersson, A. *et al.* The TMPRSS2:ERG rearrangement, ERG expression, and prostate cancer outcomes: a cohort study and meta-analysis. *Cancer epidemiology, biomarkers & prevention : a publication of the American Association for Cancer Research, cosponsored by the American Society of Preventive Oncology* **21**, 1497-1509, doi:10.1158/1055-9965.epi-12-0042 (2012).
66 Demichelis, F. *et al.* TMPRSS2:ERG gene fusion associated with lethal prostate cancer in a watchful waiting cohort. *Oncogene* **26**, 4596-4599, doi:10.1038/sj.onc.1210237 (2007).
67 St John, J., Powell, K., Conley-Lacomb, M. K. & Chinni, S. R. TMPRSS2-ERG Fusion Gene Expression in Prostate Tumor Cells and Its Clinical and Biological Significance in Prostate Cancer Progression. *Journal of cancer science & therapy* **4**, 94-101, doi:10.4172/1948-5956.1000119 (2012).
68 Gerhauser, C. *et al.* Molecular Evolution of Early-Onset Prostate Cancer Identifies Molecular Risk Markers and Clinical Trajectories. *Cancer cell* **34**, 996-1011.e1018, doi:10.1016/j.ccell.2018.10.016 (2018).
69 Sun, C. *et al.* TMPRSS2-ERG fusion, a common genomic alteration in prostate cancer activates C-MYC and abrogates prostate epithelial differentiation. *Oncogene* **27**, 5348-5353, doi:10.1038/onc.2008.183 (2008).
70 RCo, P. Dataset for histopathology reports for prostatic carcinoma (2016).
71 Panzone, J., Byler, T., Bratslavsky, G. & Goldberg, H. Transrectal Ultrasound in Prostate Cancer: Current Utilization, Integration with mpMRI, HIFU and Other Emerging Applications. *Cancer management and research* **14**, 1209-1228, doi:10.2147/cmar.s265058 (2022).
72 Drost, F. H. *et al.* Prostate MRI, with or without MRI-targeted biopsy, and systematic biopsy for detecting prostate cancer. *The Cochrane database of systematic reviews* **4**, Cd012663, doi:10.1002/14651858.CD012663.pub2 (2019).
73 Padhani, A. R. *et al.* Prostate Imaging-Reporting and Data System Steering Committee: PI-RADS v2 Status Update and Future Directions. *Eur Urol* **75**, 385-396, doi:10.1016/j.eururo.2018.05.035 (2019).
74 Pokorny, M. R. *et al.* Prospective study of diagnostic accuracy comparing prostate cancer detection by transrectal ultrasound-guided biopsy versus magnetic resonance (MR) imaging with subsequent MR-guided biopsy in men without previous prostate biopsies. *Eur Urol* **66**, 22-29, doi:10.1016/j.eururo.2014.03.002 (2014).
75 Ahmed, H. U. *et al.* Diagnostic accuracy of multi-parametric MRI and TRUS biopsy in prostate cancer (PROMIS): a paired validating confirmatory study. *Lancet (London, England)* **389**, 815-822, doi:10.1016/s0140-6736(16)32401-1 (2017).

76 Benndorf, M. *et al.* Diagnostic performance and reproducibility of T2w based and diffusion weighted imaging (DWI) based PI-RADSv2 lexicon descriptors for prostate MRI. *European journal of radiology* **93**, 9-15, doi:10.1016/j.ejrad.2017.05.015 (2017).

77 Chen, F., Cen, S. & Palmer, S. Application of Prostate Imaging Reporting and Data System Version 2 (PI-RADS v2): Interobserver Agreement and Positive Predictive Value for Localization of Intermediate- and High-Grade Prostate Cancers on Multiparametric Magnetic Resonance Imaging. *Academic radiology* **24**, 1101-1106, doi:10.1016/j.acra.2017.03.019 (2017).

78 Wang, X. *et al.* The diagnostic value of PI-RADS V1 and V2 using multiparametric MRI in transition zone prostate clinical cancer. *Oncol Lett* **16**, 3201-3206, doi:10.3892/ol.2018.9038 (2018).

79 NICE. Prostate Cancer: Diagnosis and Treatment. Clinical Guideline. (2014).

80 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* **68**, 464-470, doi:10.1016/j.eururo.2014.10.021 (2015).

81 Loeb, S. & Catalona, W. J. The Prostate Health Index: a new test for the detection of prostate cancer. *Therapeutic advances in urology* **6**, 74-77, doi:10.1177/1756287213513488 (2014).

82 Van Neste, L. *et al.* Detection of High-grade Prostate Cancer Using a Urinary Molecular Biomarker-Based Risk Score. *Eur Urol* **70**, 740-748, doi:10.1016/j.eururo.2016.04.012 (2016).

83 McKiernan, J. *et al.* A Novel Urine Exosome Gene Expression Assay to Predict High-grade Prostate Cancer at Initial Biopsy. *JAMA oncology* **2**, 882-889, doi:10.1001/jamaoncol.2016.0097 (2016).

84 Jiao, B. *et al.* Economic Evaluation of Urine-Based or Magnetic Resonance Imaging Reflex Tests in Men With Intermediate Prostate-Specific Antigen Levels in the United States. *Value in health : the journal of the International Society for Pharmacoeconomics and Outcomes Research* **24**, 1111-1117, doi:10.1016/j.jval.2021.02.009 (2021).

85 Hessels, D. *et al.* DD3(PCA3)-based molecular urine analysis for the diagnosis of prostate cancer. *Eur Urol* **44**, 8-15; discussion 15-16, doi:10.1016/s0302-2838(03)00201-x (2003).

86 Kamien, M. Can first-year medical students contribute to better care for patients with a chronic disease? *Medical Education* **24**, 23-26, doi:10.1111/j.1365-2923.1990.tb02432.x (1990).

87 Erho, N. *et al.* Discovery and validation of a prostate cancer genomic classifier that predicts early metastasis following radical prostatectomy. *PloS one* **8**, e66855, doi:10.1371/journal.pone.0066855 (2013).

88 Albala, D. *et al.* Health Economic Impact and Prospective Clinical Utility of Oncotype DX® Genomic Prostate Score. *Reviews in urology* **18**, 123-132, doi:10.3909/riu0725 (2016).

89 Cuzick, J. *et al.* Prognostic value of an RNA expression signature derived from cell cycle proliferation genes in patients with prostate cancer: a retrospective study. *Lancet Oncol* **12**, 245-255, doi:10.1016/s1470-2045(10)70295-3 (2011).

90 Humphrey, P. A. Gleason grading and prognostic factors in carcinoma of the prostate. *Modern pathology : an official journal of the United States and Canadian Academy of Pathology, Inc* **17**, 292-306, doi:10.1038/modpathol.3800054 (2004).

91 Wu, H. *et al.* Watchful Waiting and Factors Predictive of Secondary Treatment of Localized Prostate Cancer. *Journal of Urology* **171**, 1111-1116, doi:10.1097/01.ju.0000113300.74132.8b (2004).

92 Roberts, C. B. *et al.* Patterns and Correlates of Prostate Cancer Treatment in Older Men. *The American Journal of Medicine* **124**, 235-243, doi:10.1016/j.amjmed.2010.10.016 (2011).

93 Bill-Axelson, A. *et al.* Radical prostatectomy versus watchful waiting in early prostate cancer. *The New England journal of medicine* **352**, 1977-1984, doi:10.1056/NEJMoa043739 (2005).

94 Chen, C. *et al.* Comparisons of health-related quality of life among surgery and radiotherapy for localized prostate cancer: a systematic review and meta-analysis. *Oncotarget* **8**, 99057-99065, doi:10.18632/oncotarget.21519 (2017).

95 Nag, S., Beyer, D., Friedland, J., Grimm, P. & Nath, R. American Brachytherapy Society (ABS) recommendations for transperineal permanent brachytherapy of prostate cancer. *International journal of radiation oncology, biology, physics* **44**, 789-799, doi:10.1016/s0360-3016(99)00069-3 (1999).

96 Laing, R. *et al.* Low-dose-rate brachytherapy for the treatment of localised prostate cancer in men with a high risk of disease relapse. *BJU international* **122**, 610-617, doi:10.1111/bju.14223 (2018).

97 Gardner, T. A. & Koch, M. O. Prostate cancer therapy with high-intensity focused ultrasound. *Clinical genitourinary cancer* **4**, 187-192, doi:10.3816/CGC.2005.n.031 (2005).

98 Golan, R. *et al.* Partial Gland Treatment of Prostate Cancer Using High-Intensity Focused Ultrasound in the Primary and Salvage Settings: A Systematic Review. *J Urol* **198**, 1000-1009, doi:10.1016/j.juro.2017.03.137 (2017).

99 Valerio, M. *et al.* New and Established Technology in Focal Ablation of the Prostate: A Systematic Review. *Eur Urol* **71**, 17-34, doi:10.1016/j.eururo.2016.08.044 (2017).

100 van der Poel, H. G. *et al.* Focal Therapy in Primary Localised Prostate Cancer: The European Association of Urology Position in 2018. *Eur Urol* **74**, 84-91, doi:10.1016/j.eururo.2018.01.001 (2018).

101 Huggins, C. & Hodges, C. V. Studies on prostatic cancer. I. The effect of castration, of estrogen and androgen injection on serum phosphatases in metastatic carcinoma of the prostate. *CA: a cancer journal for clinicians* **22**, 232-240, doi:10.3322/canjclin.22.4.232 (1972).

102 Robson, M. & Dawson, N. How is androgen-dependent metastatic prostate cancer best treated? *Hematology/Oncology Clinics of North America* **10**, 727-747, doi:10.1016/S0889-8588(05)70364-6 (1996).

103 Immediate versus deferred treatment for advanced prostatic cancer: initial results of the Medical Research Council trial. *British Journal of Urology* **79**, 235-246, doi:10.1046/j.1464-410X.1997.d01-6840.x (1997).

104 Crawford, E. D. *et al.* Androgen-targeted therapy in men with prostate cancer: evolving practice and future considerations. *Prostate cancer and prostatic diseases* **22**, 24-38, doi:10.1038/s41391-018-0079-0 (2019).

105 Gründker, C. & Emons, G. The Role of Gonadotropin-Releasing Hormone in Cancer Cell Proliferation and Metastasis. *Frontiers in endocrinology* **8**, 187, doi:10.3389/fendo.2017.00187 (2017).

106 Kumar, R. J., Barqawi, A. & Crawford, E. D. Adverse events associated with hormonal therapy for prostate cancer. *Reviews in urology* **7 Suppl 5**, S37-43 (2005).

107 Tyrrell, C. J. *et al.* A multicenter randomized trial comparing the luteinizing hormone-releasing hormone analogue goserelin acetate alone and with flutamide in the treatment of advanced prostate cancer. The International Prostate Cancer Study Group. *J Urol* **146**, 1321-1326, doi:10.1016/s0022-5347(17)38080-1 (1991).

108 Kirby, M., Hirst, C. & Crawford, E. D. Characterising the castration-resistant prostate cancer population: a systematic review. *International journal of clinical practice* **65**, 1180-1192, doi:10.1111/j.1742-1241.2011.02799.x (2011).

109 Sathianathen, N. J. *et al.* Taxane-based chemohormonal therapy for metastatic hormone-sensitive prostate cancer. *The Cochrane database of systematic reviews* **10**, Cd012816, doi:10.1002/14651858.CD012816.pub2 (2018).

110 James, N. D. *et al.* Addition of docetaxel, zoledronic acid, or both to first-line long-term hormone therapy in prostate cancer (STAMPEDE): survival results from an adaptive, multiarm, multistage, platform randomised controlled trial. *Lancet (London, England)* **387**, 1163-1177, doi:10.1016/s0140-6736(15)01037-5 (2016).

111 Cornford, P. *et al.* EAU-ESTRO-SIOG Guidelines on Prostate Cancer. Part II: Treatment of Relapsing, Metastatic, and Castration-Resistant Prostate Cancer. *Eur Urol* **71**, 630-642, doi:10.1016/j.eururo.2016.08.002 (2017).

112 Paller, C. J. & Antonarakis, E. S. Cabazitaxel: a novel second-line treatment for metastatic castration-resistant prostate cancer. *Drug design, development and therapy* **5**, 117-124, doi:10.2147/dddt.s13029 (2011).
113 Kantoff, P. W. *et al.* Sipuleucel-T immunotherapy for castration-resistant prostate cancer. *The New England journal of medicine* **363**, 411-422, doi:10.1056/NEJMoa1001294 (2010).
114 Le, D. T. *et al.* Mismatch repair deficiency predicts response of solid tumors to PD-1 blockade. *Science (New York, N.Y.)* **357**, 409-413, doi:10.1126/science.aan6733 (2017).
115 Antonarakis, E. S. *et al.* Pembrolizumab for Treatment-Refractory Metastatic Castration-Resistant Prostate Cancer: Multicohort, Open-Label Phase II KEYNOTE-199 Study. *Journal of clinical oncology : official journal of the American Society of Clinical Oncology* **38**, 395-405, doi:10.1200/jco.19.01638 (2020).
116 Appleman, L. *et al.* 347 KEYNOTE-365 cohort C: pembrolizumab + enzalutamide in patients with abiraterone acetate–pretreated metastatic castration-resistant prostate cancer (mCRPC)—data after minimum of 22 months of follow-up. *Journal for ImmunoTherapy of Cancer* **9**, A374-A374, doi:10.1136/jitc-2021-SITC2021.347 (2021).
117 Sharma, P. *et al.* Nivolumab Plus Ipilimumab for Metastatic Castration-Resistant Prostate Cancer: Preliminary Analysis of Patients in the CheckMate 650 Trial. *Cancer cell* **38**, 489-499.e483, doi:10.1016/j.ccell.2020.08.007 (2020).
118 Parker, C. C. *et al.* Radiotherapy to the primary tumour for newly diagnosed, metastatic prostate cancer (STAMPEDE): a randomised controlled phase 3 trial. *Lancet (London, England)* **392**, 2353-2366, doi:10.1016/s0140-6736(18)32486-3 (2018).
119 Slovin, S. F. *et al.* Ipilimumab alone or in combination with radiotherapy in metastatic castration-resistant prostate cancer: results from an open-label, multicenter phase I/II study. *Annals of oncology : official journal of the European Society for Medical Oncology* **24**, 1813-1821, doi:10.1093/annonc/mdt107 (2013).
120 Scher, H. I. *et al.* Increased survival with enzalutamide in prostate cancer after chemotherapy. *The New England journal of medicine* **367**, 1187-1197, doi:10.1056/NEJMoa1207506 (2012).
121 de Bono, J. S. *et al.* Abiraterone and increased survival in metastatic prostate cancer. *The New England journal of medicine* **364**, 1995-2005, doi:10.1056/NEJMoa1014618 (2011).
122 Chung, C. & Abboud, K. Targeting the androgen receptor signaling pathway in advanced prostate cancer. *American journal of health-system pharmacy : AJHP : official journal of the American Society of Health-System Pharmacists* **79**, 1224-1235, doi:10.1093/ajhp/zxac105 (2022).
123 Pommier, Y., O'Connor, M. J. & de Bono, J. Laying a trap to kill cancer cells: PARP inhibitors and their mechanisms of action. *Science translational medicine* **8**, 362ps317, doi:10.1126/scitranslmed.aaf9246 (2016).
124 Maughan, B. L. & Antonarakis, E. S. Olaparib and rucaparib for the treatment of DNA repair-deficient metastatic castration-resistant prostate cancer. *Expert opinion on pharmacotherapy* **22**, 1625-1632, doi:10.1080/14656566.2021.1912015 (2021).
125 van der Zande, K., Oyen, W. J. G., Zwart, W. & Bergman, A. M. Radium-223 Treatment of Patients with Metastatic Castration Resistant Prostate Cancer: Biomarkers for Stratification and Response Evaluation. *Cancers* **13**, doi:10.3390/cancers13174346 (2021).
126 Galletti, G., Leach, B. I., Lam, L. & Tagawa, S. T. Mechanisms of resistance to systemic therapy in metastatic castration-resistant prostate cancer. *Cancer Treatment Reviews* **57**, 16-27, doi:10.1016/j.ctrv.2017.04.008 (2017).
127 Hussain, M. *et al.* Cabozantinib (XL184) in metastatic castration-resistant prostate cancer (mCRPC): Results from a phase II randomized discontinuation trial. *Journal of Clinical Oncology* **29**, 4516-4516, doi:10.1200/jco.2011.29.15_suppl.4516 (2011).
128 Syndikus, I., Morgan, R. C., Sydes, M. R., Graham, J. D. & Dearnaley, D. P. Late gastrointestinal toxicity after dose-escalated conformal radiotherapy for early prostate cancer: results from

the UK Medical Research Council RT01 trial (ISRCTN47772397). *International journal of radiation oncology, biology, physics* **77**, 773-783, doi:10.1016/j.ijrobp.2009.05.052 (2010).
129 Wang, Q. Cancer predisposition genes: molecular mechanisms and clinical impact on personalized cancer care: examples of Lynch and HBOC syndromes. *Acta pharmacologica Sinica* **37**, 143-149, doi:10.1038/aps.2015.89 (2016).
130 Greaves, M. & Maley, C. C. Clonal evolution in cancer. *Nature* **481**, 306-313, doi:10.1038/nature10762 (2012).
131 Barabási, A. L., Gulbahce, N. & Loscalzo, J. Network medicine: a network-based approach to human disease. *Nature reviews. Genetics* **12**, 56-68, doi:10.1038/nrg2918 (2011).
132 Barabási, A.-L. & Albert, R. Emergence of Scaling in Random Networks. *Science (New York, N.Y.)* **286**, 509-512, doi:doi:10.1126/science.286.5439.509 (1999).
133 Barabási, A.-L. & Oltvai, Z. N. Network biology: understanding the cell's functional organization. *Nature Reviews Genetics* **5**, 101-113, doi:10.1038/nrg1272 (2004).
134 Alcalá-Corona, S. A., Sandoval-Motta, S., Espinal-Enríquez, J. & Hernández-Lemus, E. Modularity in Biological Networks. *Frontiers in Genetics* **12**, 701331, doi:10.3389/fgene.2021.701331 (2021).
135 Ma'ayan, A., Blitzer, R. D. & Iyengar, R. Toward predictive models of mammalian cells. *Annual review of biophysics and biomolecular structure* **34**, 319-349, doi:10.1146/annurev.biophys.34.040204.144415 (2005).
136 Jacob, F. & Monod, J. Genetic regulatory mechanisms in the synthesis of proteins. *Journal of Molecular Biology* **3**, 318-356, doi:10.1016/S0022-2836(61)80072-7 (1961).
137 Davidson, E. H. *et al.* A genomic regulatory network for development. *Science (New York, N.Y.)* **295**, 1669-1678, doi:10.1126/science.1069883 (2002).
138 Snyder, M. & Gallagher, J. E. Systems biology from a yeast omics perspective. *FEBS letters* **583**, 3895-3899, doi:10.1016/j.febslet.2009.11.011 (2009).
139 Mendoza-Parra, M. A. *et al.* Reconstructed cell fate-regulatory programs in stem cells reveal hierarchies and key factors of neurogenesis. *Genome Res* **26**, 1505-1519, doi:10.1101/gr.208926.116 (2016).
140 Ravasz, E., Somera, A. L., Mongru, D. A., Oltvai, Z. N. & Barabási, A. L. Hierarchical organization of modularity in metabolic networks. *Science (New York, N.Y.)* **297**, 1551-1555, doi:10.1126/science.1073374 (2002).
141 Jeong, H., Mason, S. P., Barabási, A. L. & Oltvai, Z. N. Lethality and centrality in protein networks. *Nature* **411**, 41-42, doi:10.1038/35075138 (2001).
142 Dyer, M. D., Murali, T. M. & Sobral, B. W. The Landscape of Human Proteins Interacting with Viruses and Other Pathogens. *PLOS Pathogens* **4**, e32, doi:10.1371/journal.ppat.0040032 (2008).
143 von Mering, C. *et al.* STRING: a database of predicted functional associations between proteins. *Nucleic acids research* **31**, 258-261, doi:10.1093/nar/gkg034 (2003).
144 Cahan, P. *et al.* CellNet: network biology applied to stem cell engineering. *Cell* **158**, 903-915, doi:10.1016/j.cell.2014.07.020 (2014).
145 Torkamani, A. & Schork, N. J. Identification of rare cancer driver mutations by network reconstruction. *Genome Res* **19**, 1570-1578, doi:10.1101/gr.092833.109 (2009).
146 Mine, K. L. *et al.* Gene network reconstruction reveals cell cycle and antiviral genes as major drivers of cervical cancer. *Nat Commun* **4**, 1806, doi:10.1038/ncomms2693 (2013).
147 Svoboda, M. *et al.* AID/APOBEC-network reconstruction identifies pathways associated with survival in ovarian cancer. *BMC Genomics* **17**, 643, doi:10.1186/s12864-016-3001-y (2016).
148 Lin, C. Y. *et al.* Hubba: hub objects analyzer--a framework of interactome hubs identification for network biology. *Nucleic acids research* **36**, W438-443, doi:10.1093/nar/gkn257 (2008).
149 Hernandez-Toro, J., Prieto, C. & De Las Rivas, J. APID2NET: unified interactome graphic analyzer. *Bioinformatics* **23**, 2495-2497, doi:10.1093/bioinformatics/btm373 (2007).

150 Chuang, H. Y., Lee, E., Liu, Y. T., Lee, D. & Ideker, T. Network-based classification of breast cancer metastasis. *Molecular systems biology* **3**, 140, doi:10.1038/msb4100180 (2007).
151 Assenov, Y., Ramírez, F., Schelhorn, S. E., Lengauer, T. & Albrecht, M. Computing topological parameters of biological networks. *Bioinformatics* **24**, 282-284, doi:10.1093/bioinformatics/btm554 (2008).
152 Scardoni, G. *et al.* Biological network analysis with CentiScaPe: centralities and experimental dataset integration. *F1000Research* **3**, 139, doi:10.12688/f1000research.4477.2 (2014).
153 Cholley, P.-E. *et al.* Modeling gene-regulatory networks to describe cell fate transitions and predict master regulators. *npj Systems Biology and Applications* **4**, 29, doi:10.1038/s41540-018-0066-z (2018).
154 Zhao, Y. *et al.* Molecular and genetic inflammation networks in major human diseases. *Molecular BioSystems* **12**, 2318-2341, doi:10.1039/C6MB00240D (2016).
155 Takeshima, H. & Ushijima, T. Accumulation of genetic and epigenetic alterations in normal cells and cancer risk. *npj Precision Oncology* **3**, 7, doi:10.1038/s41698-019-0079-0 (2019).
156 Nussinov, R., Jang, H., Tsai, C. J. & Cheng, F. Review: Precision medicine and driver mutations: Computational methods, functional assays and conformational principles for interpreting cancer drivers. *PLoS computational biology* **15**, e1006658, doi:10.1371/journal.pcbi.1006658 (2019).
157 Jaeger, S. *et al.* Quantification of Pathway Cross-talk Reveals Novel Synergistic Drug Combinations for Breast Cancer. *Cancer research* **77**, 459-469, doi:10.1158/0008-5472.can-16-0097 (2017).
158 Piñeiro-Yáñez, E. *et al.* PanDrugs: a novel method to prioritize anticancer drug treatments according to individual genomic data. *Genome medicine* **10**, 41, doi:10.1186/s13073-018-0546-1 (2018).
159 Krzyszczyk, P. *et al.* The growing role of precision and personalized medicine for cancer treatment. *Technology* **6**, 79-100, doi:10.1142/s2339547818300020 (2018).
160 Gonzalgo M, S. K., Meeker A. Part IV. Molecular and cellular biology. . *In: Wein A, Kavoussi L, Partin A, Peters C, eds. Campbell Walsh Urology. Eleventh. Philadelphia: Elsevier Inc.* , p. 459 (2016).
161 Feitelson, M. A. *et al.* Sustained proliferation in cancer: Mechanisms and novel therapeutic targets. *Seminars in cancer biology* **35 Suppl**, S25-s54, doi:10.1016/j.semcancer.2015.02.006 (2015).
162 Luo, J., Solimini, N. L. & Elledge, S. J. Principles of cancer therapy: oncogene and non-oncogene addiction. *Cell* **136**, 823-837, doi:10.1016/j.cell.2009.02.024 (2009).
163 Nenclares, P. & Harrington, K. J. The biology of cancer. *Medicine* **48**, 67-72, doi:10.1016/j.mpmed.2019.11.001 (2020).
164 Thangavel, C. *et al.* RB Loss Promotes Prostate Cancer Metastasis. *Cancer research* **77**, 982-995, doi:10.1158/0008-5472.can-16-1589 (2017).
165 Nyquist, M. D. *et al.* Combined TP53 and RB1 loss promotes prostate cancer resistance to a spectrum of therapeutics and confers vulnerability to replication stress. *Cell reports* **31**, 107669, doi:10.1016/j.celrep.2020.107669 (2020).
166 Ko, H.-K. *et al.* A transgenic mouse model for early prostate metastasis to lymph nodes. *Cancer research* **74**, 945-953, doi:10.1158/0008-5472.CAN-13-1157 (2014).
167 Ku, S. Y. *et al.* Rb1 and Trp53 cooperate to suppress prostate cancer lineage plasticity, metastasis, and antiandrogen resistance. *Science (New York, N.Y.)* **355**, 78-83, doi:10.1126/science.aah4199 (2017).
168 Mu, P. *et al.* SOX2 promotes lineage plasticity and antiandrogen resistance in TP53-and RB1-deficient prostate cancer. *Science (New York, N.Y.)* **355**, 84-88, doi:10.1126/science.aah4307 (2017).
169 Jamaspishvili, T. *et al.* Clinical implications of PTEN loss in prostate cancer. *Nature Reviews Urology* **15**, 222-234, doi:10.1038/nrurol.2018.9 (2018).

170 Krohn, A. *et al.* Genomic deletion of PTEN is associated with tumor progression and early PSA recurrence in ERG fusion-positive and fusion-negative prostate cancer. *The American journal of pathology* **181**, 401-412, doi:10.1016/j.ajpath.2012.04.026 (2012).
171 Sarker, D., Reid, A. H., Yap, T. A. & de Bono, J. S. Targeting the PI3K/AKT pathway for the treatment of prostate cancer. *Clinical cancer research : an official journal of the American Association for Cancer Research* **15**, 4799-4805, doi:10.1158/1078-0432.ccr-08-0125 (2009).
172 Bismar, T. A. *et al.* PTEN genomic deletion is an early event associated with ERG gene rearrangements in prostate cancer. *BJU international* **107**, 477-485, doi:10.1111/j.1464-410X.2010.09470.x (2011).
173 Han, B. *et al.* Fluorescence in situ hybridization study shows association of PTEN deletion with ERG rearrangement during prostate cancer progression. *Modern pathology : an official journal of the United States and Canadian Academy of Pathology, Inc* **22**, 1083-1093, doi:10.1038/modpathol.2009.69 (2009).
174 Li, J. *et al.* PTEN, a Putative Protein Tyrosine Phosphatase Gene Mutated in Human Brain, Breast, and Prostate Cancer. *Science (New York, N.Y.)* **275**, 1943-1947, doi:10.1126/science.275.5308.1943 (1997).
175 Netto, G. J. Molecular Updates in Prostate Cancer. *Surgical pathology clinics* **8**, 561-580, doi:10.1016/j.path.2015.08.003 (2015).
176 Carver, B. S. *et al.* Aberrant ERG expression cooperates with loss of PTEN to promote cancer progression in the prostate. *Nature genetics* **41**, 619-624, doi:10.1038/ng.370 (2009).
177 King, J. C. *et al.* Cooperativity of TMPRSS2-ERG with PI3-kinase pathway activation in prostate oncogenesis. *Nature genetics* **41**, 524-526, doi:10.1038/ng.371 (2009).
178 Kim, J., Eltoum, I.-E. A., Roh, M., Wang, J. & Abdulkadir, S. A. Interactions between Cells with Distinct Mutations in c-MYC and Pten in Prostate Cancer. *PLOS Genetics* **5**, e1000542, doi:10.1371/journal.pgen.1000542 (2009).
179 Kim, M. J. *et al.* Cooperativity of Nkx3.1 and Pten loss of function in a mouse model of prostate carcinogenesis. *Proceedings of the National Academy of Sciences of the United States of America* **99**, 2884-2889, doi:10.1073/pnas.042688999 (2002).
180 Chen, Y. *et al.* ETS factors reprogram the androgen receptor cistrome and prime prostate tumorigenesis in response to PTEN loss. *Nature Medicine* **19**, 1023-1029, doi:10.1038/nm.3216 (2013).
181 Wang, S. *et al.* Prostate-specific deletion of the murine Pten tumor suppressor gene leads to metastatic prostate cancer. *Cancer cell* **4**, 209-221, doi:10.1016/S1535-6108(03)00215-0 (2003).
182 Millis, S. Z. *et al.* Phosphatidylinositol 3-kinase pathway genomic alterations in 60,991 diverse solid tumors informs targeted therapy opportunities. *Cancer* **125**, 1185-1199, doi:10.1002/cncr.31921 (2019).
183 Crumbaker, M., Khoja, L. & Joshua, A. M. AR Signaling and the PI3K Pathway in Prostate Cancer. *Cancers* **9**, 34, doi:10.3390/cancers9040034 (2017).
184 Pearson, H. B. *et al.* Identification of Pik3ca mutation as a genetic driver of prostate cancer that cooperates with Pten loss to accelerate progression and castration-resistant growth. *Cancer discovery* **8**, 764-779, doi:10.1158/2159-8290.CD-17-0867 (2018).
185 Roudsari, N. M. *et al.* Inhibitors of the PI3K/Akt/mTOR Pathway in Prostate Cancer Chemoprevention and Intervention. *Pharmaceutics* **13**, 1195, doi:10.3390/pharmaceutics13081195 (2021).
186 Derelanko, M. J., & Hollinger, M.A. Carcinogenesis. *CRC Press* **Second Edition**, doi:10.1201/9781420042078 (2001).
187 Jan, R. & Chaudhry, G. E. Understanding Apoptosis and Apoptotic Pathways Targeted Cancer Therapeutics. *Advanced pharmaceutical bulletin* **9**, 205-218, doi:10.15171/apb.2019.024 (2019).

188 Abel, E. L., Angel, J. M., Kiguchi, K. & DiGiovanni, J. Multi-stage chemical carcinogenesis in mouse skin: fundamentals and applications. *Nature protocols* **4**, 1350-1362, doi:10.1038/nprot.2009.120 (2009).
189 Guerra, C. *et al.* Tumor induction by an endogenous K-ras oncogene is highly dependent on cellular context. *Cancer cell* **4**, 111-120, doi:10.1016/s1535-6108(03)00191-0 (2003).
190 Mainardi, S. *et al.* Identification of cancer initiating cells in K-Ras driven lung adenocarcinoma. *Proceedings of the National Academy of Sciences of the United States of America* **111**, 255-260, doi:10.1073/pnas.1320383110 (2014).
191 Kandoth, C. *et al.* Mutational landscape and significance across 12 major cancer types. *Nature* **502**, 333-339, doi:10.1038/nature12634 (2013).
192 Hawksworth, D. *et al.* Overexpression of C-MYC oncogene in prostate cancer predicts biochemical recurrence. *Prostate cancer and prostatic diseases* **13**, 311-315, doi:10.1038/pcan.2010.31 (2010).
193 Eagle, L. R. *et al.* Mutation of the MXI1 gene in prostate cancer. *Nature genetics* **9**, 249-255, doi:10.1038/ng0395-249 (1995).
194 Boutros, P. C. *et al.* Spatial genomic heterogeneity within localized, multifocal prostate cancer. *Nature genetics* **47**, 736-745, doi:10.1038/ng.3315 (2015).
195 Beltran, H. *et al.* Divergent clonal evolution of castration-resistant neuroendocrine prostate cancer. *Nature Medicine* **22**, 298-305, doi:10.1038/nm.4045 (2016).
196 Kumar-Sinha, C., Tomlins, S. A. & Chinnaiyan, A. M. Recurrent gene fusions in prostate cancer. *Nature reviews. Cancer* **8**, 497-511, doi:10.1038/nrc2402 (2008).
197 Aurilio, G. *et al.* Androgen Receptor Signaling Pathway in Prostate Cancer: From Genetics to Clinical Applications. *Cells* **9**, 2653, doi:10.3390/cells9122653 (2020).
198 Cunha, G. R. The role of androgens in the epithelio-mesenchymal interactions involved in prostatic morphogenesis in embryonic mice. *The Anatomical Record* **175**, 87-96, doi:10.1002/ar.1091750108 (1973).
199 Salvesen, H. B. *et al.* PTEN methylation is associated with advanced stage and microsatellite instability in endometrial carcinoma. *International journal of cancer* **91**, 22-26, doi:10.1002/1097-0215(20010101)91:1<22::aid-ijc1002>3.0.co;2-s (2001).
200 Kallio, H. M. L. *et al.* Constitutively active androgen receptor splice variants AR-V3, AR-V7 and AR-V9 are co-expressed in castration-resistant prostate cancer metastases. *British journal of cancer* **119**, 347-356, doi:10.1038/s41416-018-0172-0 (2018).
201 Hu, R. *et al.* Ligand-Independent Androgen Receptor Variants Derived from Splicing of Cryptic Exons Signify Hormone-Refractory Prostate Cancer. *Cancer research* **69**, 16-22, doi:10.1158/0008-5472.can-08-2764 (2008).
202 Quigley, D. A. *et al.* Genomic Hallmarks and Structural Variation in Metastatic Prostate Cancer. *Cell* **174**, 758-769.e759, doi:10.1016/j.cell.2018.06.039 (2018).
203 Behrens, J. *et al.* Functional interaction of an axin homolog, conductin, with beta-catenin, APC, and GSK3beta. *Science (New York, N.Y.)* **280**, 596-599, doi:10.1126/science.280.5363.596 (1998).
204 Gabata, R. *et al.* Anti-tumor Activity of the Small Molecule Inhibitor PRI-724 Against β-Catenin-activated Hepatocellular Carcinoma. *Anticancer research* **40**, 5211-5219, doi:10.21873/anticanres.14524 (2020).
205 El-Khoueiry, A. B. *et al.* A phase I first-in-human study of PRI-724 in patients (pts) with advanced solid tumors. *Journal of Clinical Oncology* **31**, 2501-2501, doi:10.1200/jco.2013.31.15_suppl.2501 (2013).
206 Patel, R. *et al.* Activation of β-Catenin Cooperates with Loss of Pten to Drive AR-Independent Castration-Resistant Prostate Cancer. *Cancer research* **80**, 576-590, doi:10.1158/0008-5472.can-19-1684 (2020).
207 Bland, T. *et al.* WLS-Wnt signaling promotes neuroendocrine prostate cancer. *iScience* **24**, 101970, doi:10.1016/j.isci.2020.101970 (2021).

208 Pak, S. *et al.* The small molecule WNT/β-catenin inhibitor CWP232291 blocks the growth of castration-resistant prostate cancer by activating the endoplasmic reticulum stress pathway. *Journal of Experimental & Clinical Cancer Research* **38**, 342, doi:10.1186/s13046-019-1342-5 (2019).
209 Koushyar, S., Meniel, V. S., Phesse, T. J. & Pearson, H. B. Exploring the Wnt Pathway as a Therapeutic Target for Prostate Cancer. *Biomolecules* **12**, 309, doi:10.3390/biom12020309 (2022).
210 Baylin, S. B. & Jones, P. A. Epigenetic Determinants of Cancer. *Cold Spring Harbor perspectives in biology* **8**, doi:10.1101/cshperspect.a019505 (2016).
211 Jerónimo, C. *et al.* Quantitation of GSTP1 methylation in non-neoplastic prostatic tissue and organ-confined prostate adenocarcinoma. *Journal of the National Cancer Institute* **93**, 1747-1752, doi:10.1093/jnci/93.22.1747 (2001).
212 Feinberg, A. P., Koldobskiy, M. A. & Göndör, A. Epigenetic modulators, modifiers and mediators in cancer aetiology and progression. *Nature reviews. Genetics* **17**, 284-299, doi:10.1038/nrg.2016.13 (2016).
213 Maruyama, R. *et al.* Aberrant promoter methylation profile of prostate cancers and its relationship to clinicopathological features. *Clinical cancer research : an official journal of the American Association for Cancer Research* **8**, 514-519 (2002).
214 Yamanaka, M. *et al.* Altered methylation of multiple genes in carcinogenesis of the prostate. *International journal of cancer* **106**, 382-387, doi:10.1002/ijc.11227 (2003).
215 Wang, X. P. *et al.* Apc inhibition of Wnt signaling regulates supernumerary tooth formation during embryogenesis and throughout adulthood. *Development (Cambridge, England)* **136**, 1939-1949, doi:10.1242/dev.033803 (2009).
216 Bastian, P. J. *et al.* Molecular biomarker in prostate cancer: the role of CpG island hypermethylation. *Eur Urol* **46**, 698-708, doi:10.1016/j.eururo.2004.07.022 (2004).
217 Kang, G. H., Lee, S., Lee, H. J. & Hwang, K. S. Aberrant CpG island hypermethylation of multiple genes in prostate cancer and prostatic intraepithelial neoplasia. *The Journal of pathology* **202**, 233-240, doi:10.1002/path.1503 (2004).
218 Liu, L., Yoon, J. H., Dammann, R. & Pfeifer, G. P. Frequent hypermethylation of the RASSF1A gene in prostate cancer. *Oncogene* **21**, 6835-6840, doi:10.1038/sj.onc.1205814 (2002).
219 Kuzmin, I. *et al.* The RASSF1A tumor suppressor gene is inactivated in prostate tumors and suppresses growth of prostate carcinoma cells. *Cancer research* **62**, 3498-3502 (2002).
220 Merlo, A. *et al.* 5' CpG island methylation is associated with transcriptional silencing of the tumour suppressor p16/CDKN2/MTS1 in human cancers. *Nat Med* **1**, 686-692, doi:10.1038/nm0795-686 (1995).
221 Payer, L. M. & Burns, K. H. Transposable elements in human genetic disease. *Nature Reviews Genetics* **20**, 760-772, doi:10.1038/s41576-019-0165-8 (2019).
222 Urnov, F. D. Methylation and the genome: the power of a small amendment. *The Journal of nutrition* **132**, 2450s-2456s, doi:10.1093/jn/132.8.2450S (2002).
223 Scott, E. C. *et al.* A hot L1 retrotransposon evades somatic repression and initiates human colorectal cancer. *Genome Res* **26**, 745-755, doi:10.1101/gr.201814.115 (2016).
224 Rodić, N. *et al.* Long interspersed element-1 protein expression is a hallmark of many human cancers. *The American journal of pathology* **184**, 1280-1286, doi:10.1016/j.ajpath.2014.01.007 (2014).
225 Lee, E. *et al.* Landscape of somatic retrotransposition in human cancers. *Science (New York, N.Y.)* **337**, 967-971, doi:10.1126/science.1222077 (2012).
226 Yegnasubramanian, S. *et al.* DNA hypomethylation arises later in prostate cancer progression than CpG island hypermethylation and contributes to metastatic tumor heterogeneity. *Cancer research* **68**, 8954-8967, doi:10.1158/0008-5472.can-07-6088 (2008).

227 Florl, A. R. *et al.* Coordinate hypermethylation at specific genes in prostate carcinoma precedes LINE-1 hypomethylation. *British journal of cancer* **91**, 985-994, doi:10.1038/sj.bjc.6602030 (2004).
228 Pakneshan, P., Xing, R. H. & Rabbani, S. A. Methylation status of uPA promoter as a molecular mechanism regulating prostate cancer invasion and growth in vitro and in vivo. *FASEB journal : official publication of the Federation of American Societies for Experimental Biology* **17**, 1081-1088, doi:10.1096/fj.02-0973com (2003).
229 Cho, B. *et al.* Identification and characterization of a novel cancer/testis antigen gene CAGE. *Biochemical and biophysical research communications* **292**, 715-726, doi:10.1006/bbrc.2002.6701 (2002).
230 Cho, B. *et al.* Promoter hypomethylation of a novel cancer/testis antigen gene CAGE is correlated with its aberrant expression and is seen in premalignant stage of gastric carcinoma. *Biochemical and biophysical research communications* **307**, 52-63, doi:10.1016/s0006-291x(03)01121-5 (2003).
231 Kim, J. *et al.* Polycomb- and Methylation-Independent Roles of EZH2 as a Transcription Activator. *Cell Reports* **25**, 2808-2820.e2804, doi:10.1016/j.celrep.2018.11.035 (2018).
232 Varambally, S. *et al.* The polycomb group protein EZH2 is involved in progression of prostate cancer. *Nature* **419**, 624-629, doi:10.1038/nature01075 (2002).
233 Bryant, R. J., Cross, N. A., Eaton, C. L., Hamdy, F. C. & Cunliffe, V. T. EZH2 promotes proliferation and invasiveness of prostate cancer cells. *Prostate* **67**, 547-556, doi:10.1002/pros.20550 (2007).
234 Jia, L. *et al.* Locus-wide chromatin remodeling and enhanced androgen receptor-mediated transcription in recurrent prostate tumor cells. *Molecular and cellular biology* **26**, 7331-7341, doi:10.1128/mcb.00581-06 (2006).
235 Abbas, A. & Gupta, S. The role of histone deacetylases in prostate cancer. *Epigenetics* **3**, 300-309, doi:10.4161/epi.3.6.7273 (2008).
236 Gryder, B. E. *et al.* Selectively targeting prostate cancer with antiandrogen equipped histone deacetylase inhibitors. *ACS chemical biology* **8**, 2550-2560, doi:10.1021/cb400542w (2013).
237 Shang, Y., Myers, M. & Brown, M. Formation of the androgen receptor transcription complex. *Molecular cell* **9**, 601-610, doi:10.1016/s1097-2765(02)00471-9 (2002).
238 Lovén, J. *et al.* Selective inhibition of tumor oncogenes by disruption of super-enhancers. *Cell* **153**, 320-334, doi:10.1016/j.cell.2013.03.036 (2013).
239 Babu, D. & Fullwood, M. J. Expanding the effects of ERG on chromatin landscapes and dysregulated transcription in prostate cancer. *Nature genetics* **49**, 1294-1295, doi:10.1038/ng.3944 (2017).
240 Boccaletto, P. *et al.* MODOMICS: a database of RNA modification pathways. 2017 update. *Nucleic acids research* **46**, D303-d307, doi:10.1093/nar/gkx1030 (2018).
241 Bird, A. P. CpG-rich islands and the function of DNA methylation. *Nature* **321**, 209-213, doi:10.1038/321209a0 (1986).
242 Blanco, S. *et al.* Aberrant methylation of tRNAs links cellular stress to neuro-developmental disorders. *The EMBO journal* **33**, 2020-2039, doi:10.15252/embj.201489282 (2014).
243 Li, Y. *et al.* The Potential Roles of RNA N6-Methyladenosine in Urological Tumors. *Frontiers in cell and developmental biology* **8**, 579919, doi:10.3389/fcell.2020.579919 (2020).
244 Cai, J. *et al.* RNA m(6)A Methyltransferase METTL3 Promotes The Growth Of Prostate Cancer By Regulating Hedgehog Pathway. *OncoTargets and therapy* **12**, 9143-9152, doi:10.2147/ott.s226796 (2019).
245 Li, E., Wei, B., Wang, X. & Kang, R. METTL3 enhances cell adhesion through stabilizing integrin β1 mRNA via an m6A-HuR-dependent mechanism in prostatic carcinoma. *American journal of cancer research* **10**, 1012-1025 (2020).

246 Yuan, Y., Du, Y., Wang, L. & Liu, X. The M6A methyltransferase METTL3 promotes the development and progression of prostate carcinoma via mediating MYC methylation. *Journal of Cancer* **11**, 3588-3595, doi:10.7150/jca.42338 (2020).
247 Wen, S. *et al.* Long non-coding RNA NEAT1 promotes bone metastasis of prostate cancer through N6-methyladenosine. *Molecular cancer* **19**, 171, doi:10.1186/s12943-020-01293-4 (2020).
248 Du, C., Lv, C., Feng, Y. & Yu, S. Activation of the KDM5A/miRNA-495/YTHDF2/m6A-MOB3B axis facilitates prostate cancer progression. *Journal of experimental & clinical cancer research : CR* **39**, 223, doi:10.1186/s13046-020-01735-3 (2020).
249 Stockert, J. A., Weil, R., Yadav, K. K., Kyprianou, N. & Tewari, A. K. Pseudouridine as a novel biomarker in prostate cancer. *Urologic oncology* **39**, 63-71, doi:10.1016/j.urolonc.2020.06.026 (2021).
250 Bantis, A. *et al.* Expression of p120, Ki-67 and PCNA as proliferation biomarkers in imprint smears of prostate carcinoma and their prognostic value. *Cytopathology : official journal of the British Society for Clinical Cytology* **15**, 25-31, doi:10.1046/j.0956-5507.2003.00090.x (2004).
251 Kallakury, B. V. *et al.* The prognostic significance of proliferation-associated nucleolar protein p120 expression in prostate adenocarcinoma: a comparison with cyclins A and B1, Ki-67, proliferating cell nuclear antigen, and p34cdc2. *Cancer* **85**, 1569-1576, doi:10.1002/(sici)1097-0142(19990401)85:7<1569::aid-cncr19>3.0.co;2-m (1999).
252 Christofi, T. & Zaravinos, A. RNA editing in the forefront of epitranscriptomics and human health. *Journal of translational medicine* **17**, 319, doi:10.1186/s12967-019-2071-4 (2019).
253 Mo, F. *et al.* Systematic identification and characterization of RNA editing in prostate tumors. *PloS one* **9**, e101431, doi:10.1371/journal.pone.0101431 (2014).
254 Paz-Yaacov, N. *et al.* Elevated RNA Editing Activity Is a Major Contributor to Transcriptomic Diversity in Tumors. *Cell Rep* **13**, 267-276, doi:10.1016/j.celrep.2015.08.080 (2015).
255 Beyer, U. *et al.* Rare ADAR and RNASEH2B variants and a type I interferon signature in glioma and prostate carcinoma risk and tumorigenesis. *Acta neuropathologica* **134**, 905-922, doi:10.1007/s00401-017-1774-y (2017).
256 Ratti, M. *et al.* MicroRNAs (miRNAs) and Long Non-Coding RNAs (lncRNAs) as New Tools for Cancer Therapy: First Steps from Bench to Bedside. *Targeted oncology* **15**, 261-278, doi:10.1007/s11523-020-00717-x (2020).
257 Bolton, E. M., Tuzova, A. V., Walsh, A. L., Lynch, T. & Perry, A. S. Noncoding RNAs in prostate cancer: the long and the short of it. *Clinical cancer research : an official journal of the American Association for Cancer Research* **20**, 35-43, doi:10.1158/1078-0432.ccr-13-1989 (2014).
258 Begolli, R., Sideris, N. & Giakountis, A. LncRNAs as Chromatin Regulators in Cancer: From Molecular Function to Clinical Potential. *Cancers* **11**, 1524, doi:10.3390/cancers11101524 (2019).
259 López-Urrutia, E., Bustamante Montes, L. P., Ladrón de Guevara Cervantes, D., Pérez-Plasencia, C. & Campos-Parra, A. D. Crosstalk Between Long Non-coding RNAs, Micro-RNAs and mRNAs: Deciphering Molecular Mechanisms of Master Regulators in Cancer. *Frontiers in Oncology* **9**, 669, doi:10.3389/fonc.2019.00669 (2019).
260 Borchert, G. M., Lanier, W. & Davidson, B. L. RNA polymerase III transcribes human microRNAs. *Nature structural & molecular biology* **13**, 1097-1101, doi:10.1038/nsmb1167 (2006).
261 Macfarlane, L. A. & Murphy, P. R. MicroRNA: Biogenesis, Function and Role in Cancer. *Current genomics* **11**, 537-561, doi:10.2174/138920210793175895 (2010).
262 Bartel, D. P. MicroRNAs: target recognition and regulatory functions. *Cell* **136**, 215-233, doi:10.1016/j.cell.2009.01.002 (2009).

263 Vasudevan, S., Tong, Y. & Steitz, J. A. Switching from repression to activation: microRNAs can up-regulate translation. *Science (New York, N.Y.)* **318**, 1931-1934, doi:10.1126/science.1149460 (2007).
264 Kozomara, A., Birgaoanu, M. & Griffiths-Jones, S. miRBase: from microRNA sequences to function. *Nucleic acids research* **47**, D155-D162, doi:10.1093/nar/gky1141 (2018).
265 Bartel, D. P. MicroRNAs: genomics, biogenesis, mechanism, and function. *Cell* **116**, 281-297, doi:10.1016/s0092-8674(04)00045-5 (2004).
266 Boehm, M. & Slack, F. J. MicroRNA control of lifespan and metabolism. *Cell cycle (Georgetown, Tex.)* **5**, 837-840, doi:10.4161/cc.5.8.2688 (2006).
267 Carleton, M., Cleary, M. A. & Linsley, P. S. MicroRNAs and cell cycle regulation. *Cell cycle (Georgetown, Tex.)* **6**, 2127-2132, doi:10.4161/cc.6.17.4641 (2007).
268 Calin, G. A. *et al.* Frequent deletions and down-regulation of micro- RNA genes miR15 and miR16 at 13q14 in chronic lymphocytic leukemia. *Proceedings of the National Academy of Sciences of the United States of America* **99**, 15524-15529, doi:10.1073/pnas.242606799 (2002).
269 Calin, G. A. & Croce, C. M. MicroRNAs and chromosomal abnormalities in cancer cells. *Oncogene* **25**, 6202-6210, doi:10.1038/sj.onc.1209910 (2006).
270 Tagawa, H. & Seto, M. A microRNA cluster as a target of genomic amplification in malignant lymphoma. *Leukemia* **19**, 2013-2016, doi:10.1038/sj.leu.2403942 (2005).
271 Hayashita, Y. *et al.* A polycistronic microRNA cluster, miR-17-92, is overexpressed in human lung cancers and enhances cell proliferation. *Cancer research* **65**, 9628-9632, doi:10.1158/0008-5472.can-05-2352 (2005).
272 Mavrakis, K. J. *et al.* Genome-wide RNA-mediated interference screen identifies miR-19 targets in Notch-induced T-cell acute lymphoblastic leukaemia. *Nature cell biology* **12**, 372-379, doi:10.1038/ncb2037 (2010).
273 Iliou, M. S. *et al.* Impaired DICER1 function promotes stemness and metastasis in colon cancer. *Oncogene* **33**, 4003-4015, doi:10.1038/onc.2013.398 (2014).
274 Merritt, W. M. *et al.* Dicer, Drosha, and outcomes in patients with ovarian cancer. *The New England journal of medicine* **359**, 2641-2650, doi:10.1056/NEJMoa0803785 (2008).
275 Pampalakis, G., Diamandis, E. P., Katsaros, D. & Sotiropoulou, G. Down-regulation of dicer expression in ovarian cancer tissues. *Clinical biochemistry* **43**, 324-327, doi:10.1016/j.clinbiochem.2009.09.014 (2010).
276 Melo, S. A. *et al.* A genetic defect in exportin-5 traps precursor microRNAs in the nucleus of cancer cells. *Cancer cell* **18**, 303-315, doi:10.1016/j.ccr.2010.09.007 (2010).
277 Zhang, B., Pan, X., Cobb, G. P. & Anderson, T. A. microRNAs as oncogenes and tumor suppressors. *Developmental biology* **302**, 1-12, doi:10.1016/j.ydbio.2006.08.028 (2007).
278 Lewis, B. P., Burge, C. B. & Bartel, D. P. Conserved seed pairing, often flanked by adenosines, indicates that thousands of human genes are microRNA targets. *Cell* **120**, 15-20, doi:10.1016/j.cell.2004.12.035 (2005).
279 Li, Z., Yang, C. S., Nakashima, K. & Rana, T. M. Small RNA-mediated regulation of iPS cell generation. *The EMBO journal* **30**, 823-834, doi:10.1038/emboj.2011.2 (2011).
280 Zhu, S., Si, M. L., Wu, H. & Mo, Y. Y. MicroRNA-21 targets the tumor suppressor gene tropomyosin 1 (TPM1). *The Journal of biological chemistry* **282**, 14328-14336, doi:10.1074/jbc.M611393200 (2007).
281 Sayed, D. *et al.* MicroRNA-21 is a downstream effector of AKT that mediates its antiapoptotic effects via suppression of Fas ligand. *The Journal of biological chemistry* **285**, 20281-20290, doi:10.1074/jbc.M110.109207 (2010).
282 Lu, Z. *et al.* MicroRNA-21 promotes cell transformation by targeting the programmed cell death 4 gene. *Oncogene* **27**, 4373-4379, doi:10.1038/onc.2008.72 (2008).

283 Hagman, Z. *et al.* miR-205 negatively regulates the androgen receptor and is associated with adverse outcome of prostate cancer patients. *British journal of cancer* **108**, 1668-1676, doi:10.1038/bjc.2013.131 (2013).
284 Bhatnagar, N. *et al.* Downregulation of miR-205 and miR-31 confers resistance to chemotherapy-induced apoptosis in prostate cancer cells. *Cell death & disease* **1**, e105, doi:10.1038/cddis.2010.85 (2010).
285 Ke, X. S. *et al.* Genome-wide profiling of histone h3 lysine 4 and lysine 27 trimethylation reveals an epigenetic signature in prostate carcinogenesis. *PloS one* **4**, e4687, doi:10.1371/journal.pone.0004687 (2009).
286 Majid, S. *et al.* MicroRNA-205-directed transcriptional activation of tumor suppressor genes in prostate cancer. *Cancer* **116**, 5637-5649, doi:10.1002/cncr.25488 (2010).
287 Gandellini, P. *et al.* miR-205 Exerts tumor-suppressive functions in human prostate through down-regulation of protein kinase Cepsilon. *Cancer research* **69**, 2287-2295, doi:10.1158/0008-5472.can-08-2894 (2009).
288 Wan, X. *et al.* Androgen-induced miR-27A acted as a tumor suppressor by targeting MAP2K4 and mediated prostate cancer progression. *The international journal of biochemistry & cell biology* **79**, 249-260, doi:10.1016/j.biocel.2016.08.043 (2016).
289 Lin, J., Huang, H., Lin, L., Li, W. & Huang, J. MiR-23a induced the activation of CDC42/PAK1 pathway and cell cycle arrest in human cov434 cells by targeting FGD4. *Journal of ovarian research* **13**, 90, doi:10.1186/s13048-020-00686-9 (2020).
290 Yan, Y., Liang, Z., Du, Q., Yang, M. & Geller, D. A. MicroRNA-23a downregulates the expression of interferon regulatory factor-1 in hepatocellular carcinoma cells. *Oncology reports* **36**, 633-640, doi:10.3892/or.2016.4864 (2016).
291 Fang, L. L. *et al.* Expression, regulation and mechanism of action of the miR-17-92 cluster in tumor cells (Review). *International journal of molecular medicine* **40**, 1624-1630, doi:10.3892/ijmm.2017.3164 (2017).
292 Porkka, K. P. *et al.* MicroRNA expression profiling in prostate cancer. *Cancer research* **67**, 6130-6135, doi:10.1158/0008-5472.can-07-0533 (2007).
293 Blenkiron, C. *et al.* MicroRNA expression profiling of human breast cancer identifies new markers of tumor subtype. *Genome biology* **8**, R214, doi:10.1186/gb-2007-8-10-r214 (2007).
294 Bader, A. G., Brown, D. & Winkler, M. The promise of microRNA replacement therapy. *Cancer research* **70**, 7027-7030, doi:10.1158/0008-5472.can-10-2010 (2010).
295 Rupaimoole, R. & Slack, F. J. MicroRNA therapeutics: towards a new era for the management of cancer and other diseases. *Nature reviews. Drug discovery* **16**, 203-222, doi:10.1038/nrd.2016.246 (2017).
296 Mayr, C. & Bartel, D. P. Widespread shortening of 3'UTRs by alternative cleavage and polyadenylation activates oncogenes in cancer cells. *Cell* **138**, 673-684, doi:10.1016/j.cell.2009.06.016 (2009).
297 Mayr, C., Hemann, M. T. & Bartel, D. P. Disrupting the pairing between let-7 and Hmga2 enhances oncogenic transformation. *Science (New York, N.Y.)* **315**, 1576-1579, doi:10.1126/science.1137999 (2007).
298 Bartoszewski, R. & Sikorski, A. F. Editorial focus: understanding off-target effects as the key to successful RNAi therapy. *Cellular & molecular biology letters* **24**, 69, doi:10.1186/s11658-019-0196-3 (2019).
299 Trang, P. *et al.* Systemic delivery of tumor suppressor microRNA mimics using a neutral lipid emulsion inhibits lung tumors in mice. *Molecular therapy : the journal of the American Society of Gene Therapy* **19**, 1116-1122, doi:10.1038/mt.2011.48 (2011).
300 Ørom, U. A., Kauppinen, S. & Lund, A. H. LNA-modified oligonucleotides mediate specific inhibition of microRNA function. *Gene* **372**, 137-141, doi:10.1016/j.gene.2005.12.031 (2006).

301 Jopling, C. L., Yi, M., Lancaster, A. M., Lemon, S. M. & Sarnow, P. Modulation of hepatitis C virus RNA abundance by a liver-specific MicroRNA. *Science (New York, N.Y.)* **309**, 1577-1581, doi:10.1126/science.1113329 (2005).

302 Brannon-Peppas, L. & Blanchette, J. O. Nanoparticle and targeted systems for cancer therapy. *Advanced drug delivery reviews* **56**, 1649-1659, doi:10.1016/j.addr.2004.02.014 (2004).

303 Wu, X. *et al.* Immunofluorescent Labeling of Cancer Marker Her2 and Other Cellular Targets with Semiconductor Quantum Dots. *Nature biotechnology* **21**, 41-46, doi:10.1038/nbt764 (2003).

304 Nida, D. L., Rahman, M. S., Carlson, K. D., Richards-Kortum, R. & Follen, M. Fluorescent nanocrystals for use in early cervical cancer detection. *Gynecologic oncology* **99**, S89-94, doi:10.1016/j.ygyno.2005.07.050 (2005).

305 Wang, H. Z., Wang, H. Y., Liang, R. Q. & Ruan, K. C. Detection of tumor marker CA125 in ovarian carcinoma using quantum dots. *Acta biochimica et biophysica Sinica* **36**, 681-686, doi:10.1093/abbs/36.10.681 (2004).

306 Choi, K. Y. *et al.* Smart nanocarrier based on PEGylated hyaluronic acid for cancer therapy. *ACS nano* **5**, 8591-8599, doi:10.1021/nn202070n (2011).

PUBLICATIONS

Kasikci Y, et al. (2023) Role of miRNAs in the differential inter-individual gene-regulatory networks in prostate cancer. *Manuscript in preparation.*

Kasikci Y and Gronemeyer H (2021) Complexity against current cancer research - are we on the wrong track? *International journal of cancer.* 10.1002/ijc.33912.

Kumar A, Kasikci Y, et al. (2021) Patient-matched analysis identifies deregulated networks in prostate cancer to guide personalized therapeutic intervention. *American journal of cancer research.* 11,11 5299-5318.

COMMUNICATIONS

Poster – Journées Jeunes Chercheurs (JJC) en Cancérologie meeting. November 18-19, 2021. Paris, France. Personalized Therapeutic Strategies for Prostate Cancer. (Kasikci Y., Kumar A. and Gronemeyer H.)

Presentation – ED days 2022. April 21-22, 2022. Sttasbourg, France. The Complexity of Cancer Research. (Kasikci Y. and Gronemeyer H.)

Poster – EMBO Workshop: The many faces of cancer evolution. May 20-22, 2022. Rimini, Italy. Patient-specific network for personalzied therapy. (Kasikci Y., Kumar A. and Gronemeyer H.)

Yasenya KASIKCI

Analyse intégrative des réseaux dérégulés dans le cancer de la prostate pour un traitement personnalisé

Résumé

Le cancer de la prostate (CP) est le deuxième cancer le plus fréquent chez les hommes, avec une partie significative impliquant la fusion TMPRSS2-ERG. Malgré des données génomiques/transcriptomiques étendues sur le cancer, la compréhension de l'impact des mutations et de la transcription altérée sur les réseaux régulateurs des patients atteints de CP est limitée. Notre étude a utilisé des échantillons de tumeurs appariés chez les patients pour explorer les variations somatiques et les profils transcriptomiques dans les cancers de la prostate ERG+. En intégrant des bases de données d'interactions protéine-protéine et de réseaux de régulation génique, nous avons découvert des altérations de réseaux spécifiques adaptées à chaque patient, avec chaque cas de CP présentant des ensembles uniques de voies altérées. Nous avons également analysé les effets des miARN dérégulés sur les gènes cibles, ajoutant de la complexité à nos découvertes. Ensuite, des expériences de validation ont confirmé l'interaction directe des cibles miARN-ARNm. Cette analyse approfondie améliore notre compréhension des réseaux régulateurs du CP et identifie des facteurs clés pour des thérapies de gestion personnalisées du CP, offrant de nouvelles possibilités de traitements efficaces.

Mots-clés: Biologie des systèmes du cancer, cancer de la prostate, thérapie personnalisée, réseaux dérégulés appariés chez les patients

Abstract

Prostate cancer (PCa) is the second most common cancer in males, with a significant portion involving the TMPRSS2-ERG fusion. Despite extensive cancer genome/transcriptome data, understanding how mutations and altered transcription impact individual PCa patients' regulatory networks is limited. Our study used patient-matched normal and tumor samples to explore somatic variations and transcriptome profiles in primary ERG+ prostate cancers. Integrating protein-protein interaction and gene regulatory network databases, we discovered specific network alterations tailored to each patient and each PCa case exhibited unique sets of altered pathways. We also analyzed the effects of deregulated miRNAs on target genes, adding complexity to our findings. Then, validation experiments confirmed the direct interaction of miRNA-mRNA targets. Consequently, this comprehensive analysis enhances our understanding of PCa regulatory networks and identifies key factors for personalized PCa management therapies, offering new possibilities for effective treatments.

Keywords: Cancer systems biology, prostate cancer, personalized therapy, patient-matched deregulated