Universidade do Minho
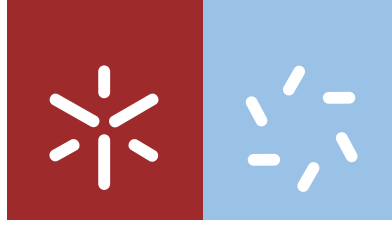Escola de Ciências

Tânia Soraia Vieira da Silva

The role of macroH2A1 in prostate carcinogenesis

outubro de 2015

Universidade do Minho
Escola de Ciências

Tânia Soraia Vieira da Silva

The role of macroH2A1 in prostate carcinogenesis

Dissertação de Mestrado
Mestrado em Genética Molecular

Trabalho efetuado sob a orientação da
Professora Doutora Carmen de Lurdes Fonseca Jerónimo
e da
Professora Doutora Ana Arminda Lopes Preto de Almeida

outubro de 2015

DECLARAÇÃO

Nome

Tânia Soraia Vieira da Silva

Endereço electrónico: soraia.silva92@gmail.com **Telefone:** 253382241

Número do Bilhete de Identidade: 14140514

Título tese

The role of macroH2A1 in prostate carcinogenesis

Orientadores:

Professora Doutora Carmen de Lurdes Fonseca Jerónimo

Professora Doutora Ana Arminda Lopes Preto de Almeida

Ano de conclusão: 2015

Designação do Mestrado: Genética Molecular

É AUTORIZADA A REPRODUÇÃO INTEGRAL DESTA TESE APENAS PARA EFEITOS DE INVESTIGAÇÃO, MEDIANTE DECLARAÇÃO ESCRITA DO INTERESSADO, QUE A TAL SE COMPROMETE.

Universidade do Minho, 21 de Outubro de 2015

Assinatura: _____

"I've missed more than 9000 shots in my career. I've lost almost 300 games. Twenty-six times I've been trusted to take the game winning shot and missed. I've failed over and over and over again in my life. And that is why I succeed."
Michael Jordan

AGRADECIMENTOS

Os últimos dois anos representaram uma fase muito importante na minha formação profissional, por todos os conhecimentos que adquiri e as técnicas que desenvolvi. A elaboração desta Tese de Mestrado representou, para mim, a minha evolução, não só a nível profissional como pessoal. E não teria sido possível sem um conjunto de pessoas, na qual sinto a necessidade de sobrescrever o meu profundo agradecimento.

Em primeiro lugar, agradeço à minha orientadora, Professora Doutora Carmen Jerónimo, por ter concordado que eu realizasse a minha Tese de Mestrado no Grupo de Epigenética e Biologia do Cancro (GEBC) do IPO-Porto. Obrigada por ter sempre acreditado em mim e lutado por este trabalho, mesmo quando surgiram contratempos. Agradeço-lhe por todos os conselhos e por se preocupar de forma tão maternal com todos os elementos do Grupo, sem distinções. Tudo o que me ensinou, todos os conhecimentos que partilhou, fez-me a crescer. Obrigada!

À Professora Doutora Ana Preto, que no seu papel de orientadora, sempre foi a minha ligação para a Universidade do Minho no último ano. Agradeço por todo o seu apoio e por toda a disponibilidade que demonstrou.

Ao Professor Doutor Rui Henrique, na qual foi um privilégio aprender tanto com ele. Obrigada pelo tempo que dispensou para acompanhar este projecto. Sinto-me grata pelo meu trabalho ter sido submetido às suas críticas e conhecimentos científicos. E ainda, como Director do Serviço de Anatomia Patológica, não posso deixar de reconhecer o papel crucial que esse grupo prestou para a realização desta tese.

Ao Director do Departamento de Genética e do Centro de Investigação do IPO-Porto, o Professor Doutor Manuel Teixeira, por estarem disponíveis todas as condições necessárias para a concretização deste trabalho no GEBC.

Agradeço ao João, que foi um pilar neste projecto. Foi quem sempre me acompanhou na parte prática, me apresentou a todas as técnicas de laboratório e me ajudou a desenhar as minhas linhas de trabalho. Tendo sido um ano tão atarefado para ele, agradeço-lhe que tenha encontrado tempo, paciência e dedicação para mim. Deixaste-me aprender com os meus erros e ensinaste-me a contornar rapidamente as dificuldades que foram surgindo. E, tal como o vinho, este projecto ganhou qualidade ao desenvolver-se no Carvalho.

Às mulheres da ciência. À Inês, na qual talvez tenha sido abusivo da minha parte todas as vezes que a obriguei a ir para o lado negro, para revelar. Agradeço-lhe por ainda me ter acompanhado no crescimento de linhas celulares. A força que transmites é incalculável e se ri e chorei o teu lado, foi porque foste das pessoas que eu procurava quando mais precisava. Também te agradeço, principalmente, por me teres segurando nos momentos de maior agitação. O teu apoio e a tua amizade não têm um preço. À Sara, por me ter introduzindo à técnica laboratorial na qual tinha menos conhecimento e a toda uma nova gastronomia que o meu estômago desconhecia. Se realizo com eficácia e autonomia o método de detecção da expressão de proteínas em células de tecido, é a ti que o devo. E se estou mais nutrida, também. Obrigada pelo teu bom humor, pelos conselhos e a alegria que espalhas. E à Ana Luís, pelos seus conhecimentos de anatomia patológica e humor inteligente, agradeço-te por sempre me ajudares a refrescar as ideias.

À Catarina e à Daniela. Foi o vosso afecto que mais me deu forças e ânimo nesta jornada. Sinto-me grata por vos ter conhecido e termos partilhado tantos momentos juntas. Agradeço-vos por terem ouvido todos os meus "Ai, Senhor!", as minhas histórias, os meus problemas e as minhas pequenas vitórias. Por sempre terem acreditado em mim, pelas conversas e pelos risos partilhados à qual apenas nós compreendíamos. É fácil perceber que nossa amizade é para sempre!

Aos restantes elementos do GEBC: à Francisca, à Eva, à Maria João, à Ana Catarina, à Sofia, à Micaela e ao Nuno. Agradeço-vos pelo "bullying", pela boa disposição, por darem vida ao laboratório e pela divisão de tarefas. O trabalho tornou-se muito mais fácil com a vossa presença e só vos desejo sucesso pelos projectos que estão a lutar. Também quero agradecer ao meu colega de secretária, o Jorge, na qual, percorremos lado a lado, a mesma caminhada. E um agradecimento especial ao Pedro que, embora esteja num fuso horário diferente, esteve sempre a horas para responder a qualquer dúvida inesperada que me surgisse.

À Dona Luísa, porque sempre se certificou de que nada faltava no laboratório. O seu carinho, organização, cuidado e dedicação é sempre indispensável.

Às minhas colegas de Ponte de Lima na qual partilhei casa neste último ano: Marina, Laura e Reci. Obrigada pelas noites longas de conversas, pelos jantares deliciosos feitos em cima do joelho, pelas competições e sessões com pipocas. E por todas as vezes que me ouviram a narrar o meu dia no laboratório, por muito pouco que

entendessem. E às outras limianas, Andreia, Valentina e Francisca, que sempre lutaram, mais do que eu, pela distância que nos separa.

Aos meus colegas de Mestrado de Genética Molecular, tenho a agradecer a união que sempre tivemos. À Andrea, pelas longas conversas que partilhámos, onde dividíamos as preocupações e as saudades.

Aos meus irmãos. À Cátia, pelas boleias e por sempre me ouvir quando mais precisava. A tua boa disposição, os teus conselhos e o teu apoio foram indispensáveis. Ao Nuno e à Ana, por serem simplesmente o que são, crianças. Involuntariamente, sempre me ajudaram a descontrair de uma semana de trabalho e a, de vez enquanto, apreciar o mundo através de olhos de criança.

Ao meu pais, que sempre me apoiaram, financeiramente, para que nada me faltasse. São a eles quem eu mais tenho que agradecer por me terem permitido esta experiência. Agradeço-lhes pelo ensinamento, carinho, dedicação, amor, apoio e incentivo que sempre foram incondicionais. Ao meu pai, agradeço-lhe todas as vezes que abandonou a cama mais cedo por minha causa. Ao seu humor singular, paciência, dedicação e esforço que faz por querer estar sempre presente. À minha mãe que é, e sempre será, o pilar da minha vida e na qual, é impossível descrever o quanto lhe devo. Agradeço-lhe por tudo.

A todos vós, o meu profundo agradecimento!

THE ROLE OF MACROH2A1 IN PROSTATE CARCINOGENESIS
ABSTRACT

Prostate cancer (PCa) is the most common noncutaneous malignancy in men and the major cause of cancer-related morbidity and mortality worldwide. Due to genetic and epigenetic deregulations, prostate cancer is characteristically asymptomatic in early stages. Deeper understanding of this mechanisms strength the development of new and improved diagnostic and prognostic tools and, therefore, better treatment strategies.

The shuffle of canonical histones, an epigenetic mechanism, is highly conserved among species and expression alterations of these histones variants, such as macroH2A1, are related to cancer development. *H2AFY* gene codifies two isoforms of the H2A histone variant macroH2A1: macroH2A1.1 and macroH2A1.2. MacroH2A1.1 inhibits cell proliferation and cell migration, whilst macroH2A1.2 has opposite functions. To date, there were studies of this histone variant in several cancer types, but none in PCa. Thus, our aim was to assess whether macroH2A1 is implicated in prostate carcinogenesis.

In a large series of prostate samples from Portuguese Oncology Institute-Porto, we found that macroH2A1.1 transcript levels were downregulated in high-grade prostatic intraepithelial neoplasia (PIN) and primary PCa compared to normal prostatic tissues. Moreover, QKI, a splicing regulator that induces macroH2A1.1 expression, presented similar results. Compared with clinicopathological data, macroH2A1.1 and QKI expression were associated with Gleason Score and PSA blood levels. Both transcripts were able to discriminate cancerous from noncancerous prostate tissues.

MacroH2A1.1 *in vitro* overexpression in a PCa Cell line decreased cell viability. Thus, macroH2A1.1 seems to play a critical role in PCa development.

Keywords: Cancer, epigenetic, isoforms, macroH2A1, prostate, QKI, splicing regulators

O PAPEL DA MACROH2A1 NO CARCINOMA DA PRÓSTATA

RESUMO

O cancro da próstata é, mundialmente, a neoplasia não-cutânea mais comum do sexo masculino e a maior causa de morbilidade e mortalidade associada ao cancro. Com alterações genéticas e epigenéticas, o cancro da próstata é, inicialmente, assintomático. Uma melhor compreensão sobre estes mecanismos oferece o desenvolvimento de novas e aperfeiçoadas análises diagnósticas e, posteriormente, uma melhor aplicação de tratamentos.

A substituição das histonas canónicas, um mecanismo epigenético, encontra-se conservada ao longo da evolução. Alterações da expressão dessas variantes de histonas, como a macroH2A1, correlacionam-se com o desenvolvimento de cancro. O gene *H2AFY* codifica duas isoformas da variante macroH2A1, da família H2A: macroH2A1.1 e macroH2A1.2. Enquanto a macroH2A1.1 inibe a proliferação e a migração celular, a macroH2A1.2 tem consequências opostas. Até hoje, há registos desta variante de histona em diversos estudos de cancro, embora nenhum em cancro da próstata. Com base no que foi descrito, esta tese tem como principal objectivo determinar se a variante macroH2A1 está associada com o desenvolvimento do carcinoma da próstata.

Utilizando uma longa série de amostras de próstata do Instituto Português de Oncologia – Porto, descobrimos que os níveis de transcrito da macroH2A1.1 se encontravam mais baixos em neoplasias intraepiteliais prostáticas (PIN) de alto grau e tecidos primários de cancro da próstata, quando comparados com tecidos não-neoplásicos de próstata. Adicionalmente, o QKI, um regulador de *splicing* que induz a expressão da macroH2A1.1, demonstrou resultados semelhantes. Comparando com os dados clinico patológicos, a expressão dos genes macroH2A1.1 e QKI estão associados com o Gleason Score e níveis de PSA no sangue. Ambos os transcritos também discriminam significativamente tecidos primários de cancro da próstata de tecidos não neoplásicos.

A sobreexpressão de macroH2A1.1 numa linha de cancro da próstata diminuiu a viabilidade celular. Assim, a macroH2A1.1 parece desempenhar um papel relevante no desenvolvimento de cancro da próstata.

Palavras-chave: Cancro, epigenética, isoformas, macroH2A1, próstata, QKI, reguladores de *slincing*

.

TABLE OF CONTENTS

FIGURE LIST

TABLE LIST

ABBREVIATIONS

APS – Ammonium persulfate

AR – Androgen receptor

AS – Active surveillance

AJCC – American Joint Committee on Cancer

ATRX – α-thalassemia/MR, X-linked

APLF – Aprataxin-PNK-like factor

BPH – Benign prostate hyperplasia

BSA – Bovine serum albumin

CDK8 – Cyclin-dependent kinase 8

DNMT – DNA methyltransferase

DDX5 – Deadbox 5 (also known as **p68**)

DDX17 – Deadbox 17 (also known as **p72**)

DRE – Digital Rectal Examination

ES cells – Embryonic stem cells

EMT – Epithelial mesenchymal transition

ERBB2 – Receptor tyrosine-protein kinase

FBS – Fetal bovine serum

FFPE – Formalin-fixed, paraffin embedded

GUSβ – β-glucuronidase,

HAO1 – Hydroxyacid Oxidase 1

H3K04me3 – Tri-methylated histone H3 at lysine 4 (active transcription)

H3K27me – Mono-methylated histone H3 at lysine 27 (repress transcription)

H3K27me3 – Tri-methylated histone H3 at lysine 27 (repress transcription)

HAT – Histone acetyltransferase

HDAC – Histone deacetylase

HDM – Histone demethylase

HGPIN – High-grade PIN

HMT – Histone methyltransferase

LGPIN – Low-grade PIN

MET – Mesenchymal epithelial transition

mH2A1 – MacroH2A1

mH2A1.1 – MacroH2A1.1

mH2A1.2 – MacroH2A1.2

miRNA – MicroRNA

MNPT – Morphologic normal prostate tissue

MRI – Magnetic Resonance Imaging

mRNA – Messenger ribonucleic acid

NAD+ – Nicotinamide adenine dinucleotide

NC – Negative control

NRF-1 – Nuclear respiratory factor 1

PCa – Prostate cancer

PCR – Polymerase chain reaction

PIA – Proliferative inflammatory atrophy

PIN – Prostate intraepithelial neoplasia

PSA – Prostate-specific antigen

PAR – Poly-ADP-ribose

PARP-1 – Poly-ADP-ribose polymerase 1

PTM – Post-translational modification

qRT-PCR – Quantitative reverse transcriptase polymerase chain reaction

QKI – Quaking

RP – Radical Prostatectomy

SDS – Sodium dodecyl sulfate

SOD3 – Superoxide Dismutase 3

SAHF – Senescence-Associated Heterochromatic Foci

SWI/SNF – Switch/sucrose non-fermentable

TEMED – Tetramethylethylenediamine

TNBC – Triple-negative breast cancer

TRUS – Transrectal Ultrasound

TURP – Transurethral resection of the prostate

TSS – Transcription start site

UICC – Union for International Cancer Control

WW – Watchful waiting

I . INTRODUCTION

1. PROSTATE

1.1. PROSTATE ANATOMY, HISTOLOGY AND PHYSIOLOGY

Prostate, along with the seminal vesicles and the bulbourethral glands, constitute the male reproduction's accessory glands [1]. The prostate is a walnut-shaped organ, which the size grow with age, around 28 to 47cm^2 and is localized under the bladder, near the rectum, surrounding the beginning of the urethra [2, 3]. The function of the prostate is to segregate an alkaline fluid, where one of the components is a serine protease of the Kallikrein family, the prostate-specific antigen (PSA) [4].

The prostate is composed by acini and ducts, organized in lobules, and delimited by fibromuscular stroma. Acinus *per se* consists of epithelial (secretory and basal) and neuroendocrine cells, surrounded by fibroblasts and smooth-muscle cells [5]. Stromal and epithelial cells express androgen receptors (ARs), depending on androgens (i.e. testosterone) to proliferate [5]. A thin layer of connective tissue surrounds the prostate, being connect with nerves and other tissues, constituting the prostatic capsule [1].

The model of prostate anatomy has been puzzlingly discussed through time and culminate divided into lobes, based on laboratory animal's analogy [6]. This concept was accepted until the decade of 1960s, when John E. McNeal start describing the most widely accepted anatomic divisions of the prostate: peripheral, central, transition and anterior fibromuscular stroma zones (Figure 1) [7, 8]. The peripheral zone is structured by a disc of tissue with radiated ducts laterally from the urethra lateral and distal, which constitutes 70% of the glandular prostate. The central zone is organized by ducts that follow the ejaculators ducts, constituting 25% of the prostate. The transition zone includes the prostatic urethra and arranges 5% of the glandular prostate. Lastly, the anterior fibromuscular forms the thick surface of the prostate and is responsible for sphincter functions [9, 10]. Cells of the transition zone proliferate dramatically throughout the puberty and later, after the age of 55 years, leading to the increase of the main zone of the glandular prostate, the peripheral zone [5]
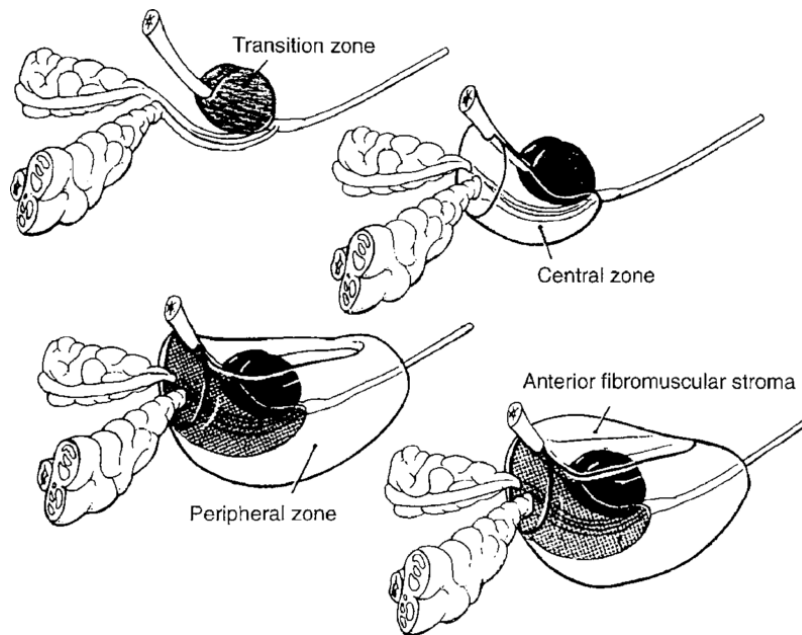
Figure 1. Anatomic zones of the prostate described by McNeal. Adapted from *Hammerich et al*, 2008 [8].

1.2. NON-CANCEROUS PROSTATE DISEASES

Prostate disorders are commonly being more frequent in men with advanced age [11]. The most common non-cancerous prostate diseases include benign prostatic hyperplasia (BPH), proliferative inflammatory atrophy (PIA) and prostatic intraepithelial neoplasia (PIN).

BPH origins in the transition zone of the prostate and is described by an excess of glands and stroma [12]. The possible risk factors for BPH are heredity, gene polymorphisms, diet, metabolic syndrome, exercise and cigarette smoking [13].

A PIN lesion take place in the peripheral zone and is commonly characterized by neoplastic resembles with undetectable abnormal changes phenotypically and not raising the PSA levels [14]. PIN lesions were firstly characterized by Bostwick and Brawer [15] in *low* and *high grade* PIN (LGPIN and HGPIN, respectively), which differ by architectural and cytological characteristics (Table 1) [16]. PIN spread through prostatic ducts and is characterized by the conservation of the basal cell layer, while luminal cells are replaced by neoplastic cells [17]. Those neoplastic cells own a hyperchromatic nuclei and nucleoli enlargement [18].

4

Table 1. Criteria for low and high PIN. Adapted from *Bostwick and Cheng*, 2012 [16].

	LGPIN	HGPIN
Architecture	Epithelial cells crowding and stratification, with irregular spacing	Similar to low-grade PIN; More crowding and stratification; four patterns: tufting, micro papillary, cribriform, and flat
Cytology Nuclei	Enlarged, with marked size variation	Enlarged; some size and shape variation
Chromatin	Normal	Increased density and clumping
Nucleoli	Rarely prominent*	Prominent
Basal cell layer	Intact	May show some disruption
Basement membrane	Intact	Intact

*Fewer than 10% of cells have prominent nucleoli.

PIA usually originates in the peripheral zone and is described by the rapidly epithelial cells division without full differentiation [16]. Proliferative cell regeneration is induced by inflammation or external factors, as chemicals or bacteria [16]. PIN and PIA share similar alterations at key molecular pathways and originates in the same prostate glandular zone, suggesting that PIA can be a precursor of PIN [19].

1.3. PROSTATE CANCER

The most prevalent malignant disease in prostate is adenocarcinoma corresponding to approximately 95% of the cases [8]. Prostate cancer (PCa) is characterized by a heterogeneous low proliferate carcinoma, asymptomatic when confined in the organ (latent tumors) [20]. The only recognized putative precursor of PCa is HGPIN, which pre-dates the onset of PCa by 5–10 years and, along with PCa, disrupt both cell layers (Figure 2) [21]. Other prostate diseases keep the basal cell layer intact [16].
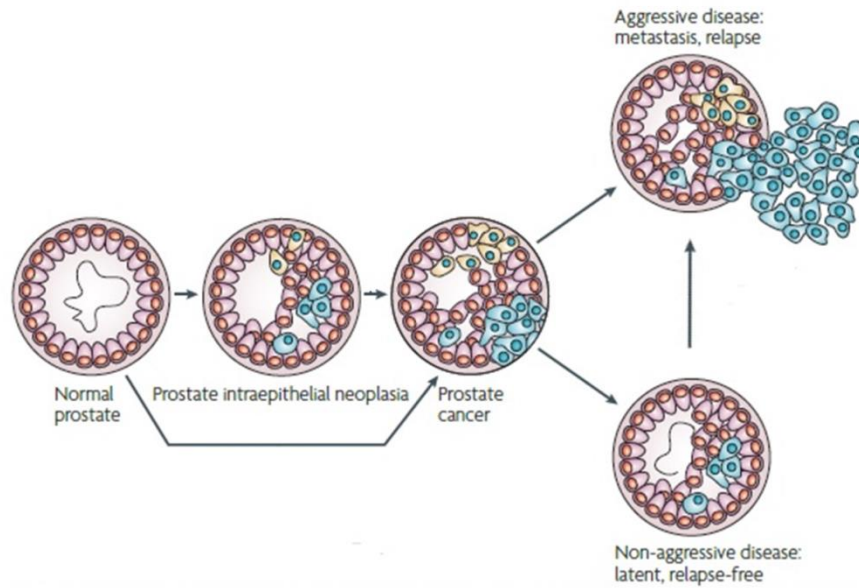
Figure 2. Cellular progression of prostate cancer. Adapted from *Witte*, 2009 *[26]*.

Approximately 9% of isolated HGPIN is found in biopsies, although the prevalence of HGPIN with PCa in biopsies vary with the number of cores and race, not surpassing 45% [16]. Irrefutably, HGPIN and PCa share the location [15] and have similar morphology [22], histology [23] and chromosomal abnormalities [24]. Consequently, HGPIN diagnosis could be used as tool for patients with PCa predisposition [18]. The heterogeneity, slow-growing behavior and no symptoms in early phases, turns PCa in a real challenge for patient management, triggering late diagnosis, and consequently compromising prognosis and target therapies [25]. Hence, it is important to comprehend underlying mechanisms and sequential pathways of PCa initiation and development [26].

1.4. EPIDEMIOLOGY OF PROSTATE CANCER: INCIDENCE AND MORTALITY

PCa is a major health concern due to growth and ageing of global population. PCa is the fourth more frequent cancer considering overall population and, after lung cancer, is the most common cancer in men [27].

PCa incidence diverge drastically worldwide, thought could be related with the median-age and number of cases diagnosed per country. In the early 1990s, there was a dramatically increase of PCa incidence worldwide, due to the introduction and largely

use of transurethral resection of the prostate (TURP) and PSA screening for cancer detection in developed countries [28]. Nonetheless, PCa incidence is higher in North America, Australia and Nordic countries, whereas lower incidence is found in Asia and Northern Africa (Figure 3A) [27]. In 2015, is expected that PCa would represent 25% of all new diagnosis cases in men [28].

Concerning PCa mortality, rates have been more constant through time. PCa mortality has been decreasing due to early diagnosis and therefore, the therapy is provided at earlier stages of the disease [29]. Currently, PCa constitutes the fifth cancer-related mortality worldwide, excluding non-melanoma skin cancer. PCa mortality rate is more prevalent in Africa and South America (Figure 3B) [30]. In Portugal, PCa is currently, the number one in incidence rates and second in mortality rate, among men [27].
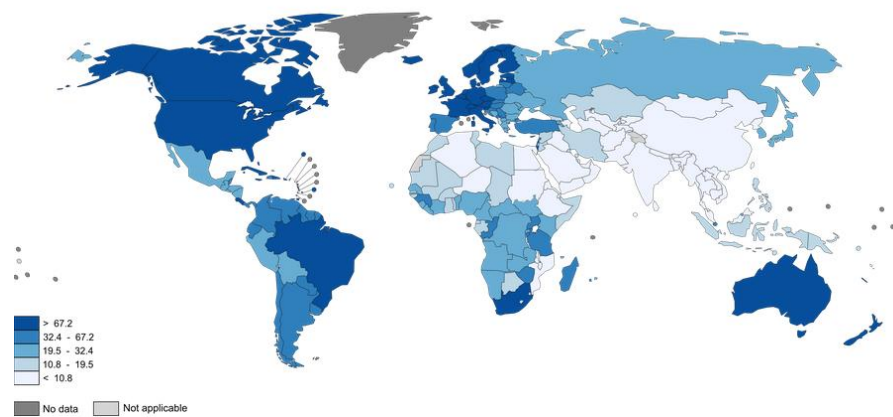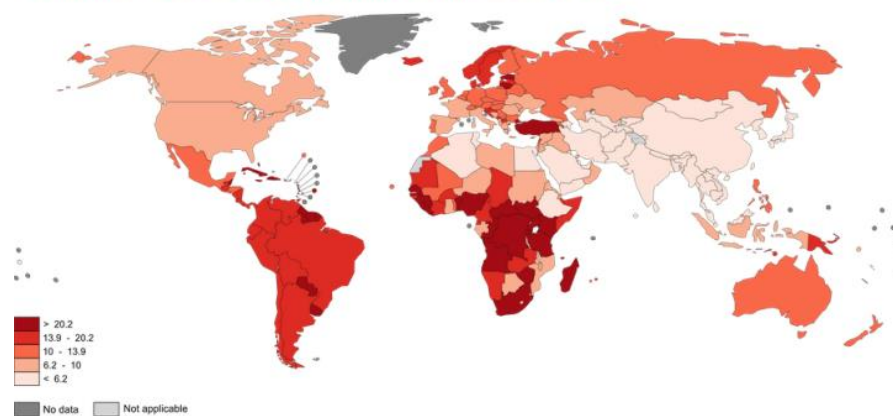
Figure 3. Estimated incidence (A) and mortality worldwide of prostate cancer in 2012. Adapted from *Globocan* [27].

1.5. RISK FACTORS

To date, there are established three risk factors which represent the furthermost main influences that could lead to PCa: age, family history and ethnicity [31].

PCa patients' average age is between 70 and 74 years, and it increases in older men [32]. Indeed, the likelihood of PCa development is 85% in men after 65 years old and higher than 90% in men with more than 90 years old [33]. It should be recalled that precursor lesions and PCa early phases are silence diseases, therefore, it may exist during years or decades before PCa is diagnosed [34].

Family history always represented a risk factor to develop cancer and PCa is no exception. Familial PCa represents 10-15% of all PCa diagnosed cases [35]. Additionally, it was observed that first-degree relatives of PCa patients have a higher risk to develop the disease. Furthermore, the number of affected members in a family and the early-onset of the disease increase even more the risk of prostate cancer [33]. Nevertheless, familial PCa and non-familial PCa are clinically and pathologically similar [31].

Lastly, ethnicity may justify PCa incidence divergence around the world. African-American men present 60% higher probabilities to develop PCa, even in younger ages, comparing to Caucasian American men [36]. However, immigration studies suggest that races with low PCa incidence, as Asians, increase dramatically the probabilities of develop PCa when immigrate to America [33, 34]. Hence, external factors, as environment, dietary habits, exercise, access to medical care and diagnosis tools and others [37, 38], might have an additional role in the likelihood of developing PCa [31].

1.6. DIAGNOSTIC TOOLS FOR PROSTATE CANCER

The efforts for development of tools for PCa detection is to effectively identify this disease while silently confined in the organ and, thus, curable. The two complemental detection tools available nowadays are digital rectal examination (DRE) and PSA screening.

Since PCa develops in the peripheral zone of the prostate gland and knowing the prostate proximity to the distal rectum, about 18% of all PCa can be detected by DRE

[5, 39]. However, DRE lacks in sensibility and depends on professional experience [39]. Alternatively to DRE, PCa can be detected by Transrectal ultrasound (TRUS) or Transrectal magnetic resonance imaging (MRI), though the last is more utilized to verify PCa invasion to nearby tissues [40, 41].

Alternatively, the glycoprotein PSA is segregated in epithelial cells of the prostate and release in the blood circulation. PSA quantification was introduced as a diagnosis tool in the 1980's, providing the identification of prostate diseases, with low levels of specificity and sensibility for PCa [42]. It is expected PSA levels between 0 to 4.0ng/ml in prostate-healthy men under 70 years old and slightly higher through age. Additionally, PSA levels can be influenced by obesity [43], cardiovascular disorders [44], type 2 diabetes [45] and other prostatic diseases besides PCa. This test demonstrates important limitations in specificity and sensibility but, to date, is the only available biomarker used for the detection and monitoring of treatment efficacy for prostate cancer [46].

Regardless limitations, the annual combination of DRE and PSA screening, in fact, diminish the number of advanced PCa patients [47]. If DRE and PSA screening results are PCa abnormal, is recommended a TRUS-guided systemic needle biopsies 3 to 6 months, with 12 or more small tissue cylinders (cores) removed each biopsy for analyzation [5, 48].

Moreover, it is important to take into account that PCa patients are, commonly, older than 50 years old and inaccurate regular diagnosis can lead to over diagnosis and over treatment of latent tumors and damage both physical and psychological. Therefore, it is advocated to avoid PSA screening in men over 75 years old [1]. All the points mentioned above strengthen the importance to develop specific non-invasive diagnosis methods for PCa [49].

1.7. PROGNOSTIC TOOLS FOR PROSTATE CANCER

Prognostics tools are designed to accurately distinguish clinically significant from indolent PCa. Currently, Gleason Score and the TNM systems assist clinicians in decision-making.

The heterogeneity of PCa is the main problem in prostate biopsies, since cores may not represent the complete tumor [50]. To decipher the glandular epithelial

architectural patterns, ignoring cytologic details, in 1966, Donald F. Gleason elaborate a histological grading system based on the sum of the two more frequent glandular histological patterns present in each tumor: the Gleason Score [51]. This system scores well-differentiated pattern as 1 and as 5 the most undifferentiated. Therefore, Gleason grading system increases with the tumor aggressiveness, in a 2 (1+1) to 10 (5+5) combined score scale (Figure 4) [51]. Although limited by the pathologist proficiency and the cores removed, accurate Gleason Score is critical, once, can differ in malignancy based on the most frequent pattern, for example, a Gleason Score 5+3 (n=8) represent a worst prognosis than a Gleason Score 4+4 (n=8) or 3+5 (n=8) [50].
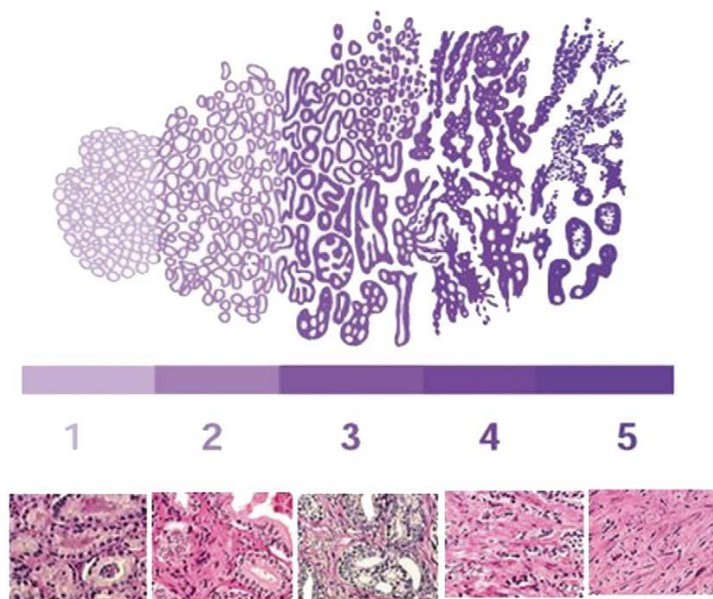
Figure 4. Gleason Score: histological grading for prostate cancer. Grade 1 (well differentiated): closely packed, uniform shaped glands. Grade 2 (well differentiated): infiltration into the surrounding stroma, more variation in gland size and spacing. Grade 3 (moderately differentiated): irregular size and shape, separation of the glands, less defined boundaries and less intervening stroma. Grade 4 (poorly differentiated): fusion of the glands with a ragged invasive edge. Grade 5 (undifferentiated): complete absence of gland formation with sheets or clusters of cells. Adapted from *Harnden et al*, 2007 [52].

To recognize the dimension of PCa and the level of extension, in 1950s was establish a clinical and pathological staging system for solid tumors: the TNM (Tumor Node Metastasis) classification system. The American Joint Committee on Cancer (AJCC) and the Union for International Cancer Control (UICC) systematically update the staging system that allows distinguish primary tumors clinically (T) and pathologically (pT), regional lymph nodes status clinically (N) and pathologically (pN) and distant metastases (M) (Table 2) [53]. Clinical staging is only associated to the evaluation of cancer spread, being firstly obtained during diagnosis, before treatment. Pathological staging is related to histological data and is firstly determined after radical

prostatectomy (RP) thus there is no pT1 classification [54]. Concerning PCa metastasis, tumors frequently spread to bones, lymph nodes, lungs, liver and brain [55].

The TNM classification, along with Gleason Score and PSA screening results, provides a complete PCa stage, classified to I to IV, increasing with the PCa aggressiveness [54].

Table 2. The AJCC/UICC TNM staging classification for PCa. Adapted from *Edge and Compton*, 2010 [54].

PRIMARY TUMOR (T)	
CLINICAL	
Tx	**Primary tumor cannot be assessed**
T0	**No evidence of primary tumor**
T1	**Clinically unapparent tumor neither palpable nor visible by imaging**
T1a	Tumor incidental histologic finding in 5% or less of tumor resected
T1b	Tumor incidental histologic finding in more than 5% of tumor resected
T1c	Tumor identified by needle biopsy
T2	**Tumor confined within prostate gland**
T2a	Tumor involves one half of one side or less
T2b	Tumor involves more than one half of one lobe but not both lobes
T2c	Tumor involves both lobes
T3	**Tumor extends through prostate capsule**
T3a	Extracapsular extension (unilateral or bilateral)
T3b	Tumor invades seminal vesicle(s)
T4	**Tumor is fixed or invades adjacent structures other than seminal vesicles, such as: external sphincter, rectum, bladder, levator muscles, and/or pelvic wall**
PATHOLOGIC (PT)	
pT2	**Organ confined**
pT2a	Unilateral, one half of one side or less
pT2b	Unilateral, involving more than one half of one lobe but not both lobes
pT2c	Bilateral disease
pT3	**Extraprostatic extension**
pT3a	Extraprostatic extension or microscopic invasion of bladder neck
pT3b	Seminal vesicle invasion
pT4	**Invasion of rectum, levator muscles and/or pelvic wall**
REGIONAL LYMPH NODES (N)	
CLINICAL	
Nx	**Regional lymph nodes were not assessed**
N0	**No regional lymph node metastasis**
N1	**Metastasis in regional lymph node(s)**
PATHOLOGIC (pN)	
pNx	**Regional nodes not sampled**
pN0	**No positive regional nodes**
pN1	**Metastasis in regional node(s)**
DISTANT METASTASIS (M)	
M0	**No distant metastasis**
M1	**Distant metastasis**
M1a	Non-regional lymph node(s)
M1b	Bone(s)
M1c	Other site(s) with or without bone disease

1.8. PROSTATE CANCER'S CLINICAL MANAGEMENT

The main goal of treatment of clinically localized PCa (stage I and II) is the cancer eradication, while no curative treatment is available for advanced PCa (stage III and IV) and treatment is only palliative support. Early-stages PCa present about 90% of progression-free survival after 5 to 10 years [56]. In fact, it is more frequent a man die with PCa than from PCa. Nevertheless, all aggressive therapies could lead to different side-effects, as urine or bowel dysfunction, fatigue, increased risk of diabetes or heart attacks and others [57]. Therefore, age, life expectancy, comorbidities and quality of life of the patients are taken in consideration to select the better treatment approach.

To avoid inadequate treatments, PCa patients can be monitored by watchful waiting (WW) or active surveillance (AS). It is suggested WW to patients who are not advised to undergo aggressive treatment. These patients are followed on 6 months and only are treated if PCa progress. AS is recommended for indolent tumors where therapies are pointless: low Gleason Score grade, low PSA screening result and <50% presence of cancer in biopsies [58]. These patients are followed by systematically diagnosis procedures, evaluating the progression of PCa.

Clinically localized PCa can be treated with RP or radiotherapies. For early-stage PCa patients with good general conditions for surgical intervention and with 10 or more years of life expectancy, the most adequate treatment is ablation of the prostate gland and the seminal vesicles by RP [59]. Radiotherapy may be an alternative to RP, showing high rates of disease-free survival, either by noninvasive external-beam radiation therapy or interstitial radiation therapy (brachytherapy), in which radioactive seeds with a life-time of 60 days are placed near the tumor [60, 61].

For advanced PCa patients, the treatment option is suppress the action or inhibit the production of testosterone, decreasing the prostate hormone-response. Androgen-deprivation therapy can be achieved by surgical castration (orchiectomy) or chemical castration, a combination of gonadotropin-releasing hormone analogues with antiandrogens (i.e. bicalutamide) [57]. These therapies may be used along with early-advanced PCa treatments. Unluckily, AR mutations lead to castration-resistance after 18-30 months of treatment [62]. The therapies available for metastatic castration-resistant PCa patients only provide supportive care.

2. EPIGENETICS

The nucleus of a human cell compacts the three billion base pair genome: DNA bonded to proteins, forming the chromatin [63]. When chromatin is strongly compacted is named heterochromatin and when is lesser condensed is designated euchromatin which is associated with transcription, DNA replication or repair and recombination processes [64]. Epigenetic mechanisms play a key role in chromatin dynamics and therefore in expression regulation.

The term "epigenetic" use the Greek prefix epi- which means over, beyond – genetics and was defined by Conrad Waddington, in the 1940s, as the branch of science of embryonic development studies, through experimental analysis. [65]. The "epigenetic landscape" was the explanation of cellular differentiation: how totipotent cells develops into all the different cells types in an organism with the same genome [66].

Epigenetic definition has been changing through time and currently, is defined as the heritable changes that occur in a gene regulation/function without alter the DNA sequence [67]. Epigenetics studies explain, for example, the differences among monozygotic twins or, in females, the silence of one X chromosome [66].

Epigenetic mechanisms are divided in four different main groups: DNA methylation, non-coding RNAs, post-translational modifications (PTMs) of histones and histone variants, which will be slightly described below. Alterations in epigenetic mechanisms affect innumerous cells processes, being implicated in several diseases, including cancer.

2.1. DNA METHYLATION

In mammals, DNA methylation refers to the addition of a methyl group, by DNA methyltransferases (DNMT), in a cytosine next to a guanine, known as CpG dinucleotides. CpG dinucleotides clusters are designed as "CpG islands" and are generally found in promoters, introns, repetitive sequences or untranslated sequences of the genome [66]. The latter are globally methylated in the genome being important to maintain DNA stability [68].

Promoters with low level of methylation are related with active gene expression, whereas heavy hypermethylated promoters are associated with stable silenced genes, as in the inactive female X chromosome [69]. DNA methylation inhibits gene expression directly by avoiding the binding of transcription factors [70] or indirectly by the recruitment of chromatin remodeling complexes [71].

DNA methylation is the most studied epigenetic mechanism in cancer. A cancer cell is characterized by hypermethylation of tumor suppressor genes promoters and by global hypomethylation of the genome [72].

2.2. NON-CODING RNAS

Nearby 90% of all RNAs transcribed are non-coding RNAs that do not codifiy proteins [73]. Non-coding RNAs, as ribosomal RNAs, are grouped according to size; microRNAs (miRNAs) are 18-30 nucleotides, 30–300nt are denominated small RNAs and non-coding RNAs with larger 300nt are considered long RNAs [73]. Non-coding RNAs are described as key players in gene regulation [73]. From these, miRNAs are the most well studied in cancer [73, 74].

MiRNAs are synthetized and processed in the nucleus and are transported to the cytoplasm to bind complementary mRNAs, repressing their function by degradation or by translation inhibition [74]. Interestingly, miRNAs could also be involved in the up-regulation of translation during the cell cycle [75].

Different mRNAs can be regulated by the same miRNA, the same way as different miRNA can target the same mRNAs [74]. About 30% of the human genes are regulated by time and tissue-specific miRNAs [76], interfering with several cellular pathways as differentiation, proliferation, apoptosis, and stress response [77].

In cancer, upregulated miRNAs target tumor suppressor genes and downregulated miRNAs target oncogenes [74]. Gene amplification, deletion, mutation and other epigenetic mechanisms can alter the miRNAs expression [74].

2.3. HISTONE POST-TRANSLATIONAL MODIFICATIONS

Eukaryotic DNA is packaged by histones, positively-charged proteins that easily bind with the negatively-charge DNA [78]. Eight histones, one pair of each H2A, H2B,

15

H3 and H4, constitute a protein complex designed nucleosome that is wrapped by a core DNA 1.7 times and sealed by one H1 [79, 80], along with numerous hydrogen, electrostatic and hydrophobic bonds [81].

Histones are dynamic proteins responsible for DNA support and chromosomal remodel [82]. All histones share a similar structural architecture with α-helices bonded by short loops and a flexible undefined N-terminal tail where is more susceptible to occur covalent histone modifications (post-translational modifications), such as acetylation, methylation, phosphorylation or ubiquitination which impact on chromatin condensation and globally constitute the so-called histone code (Figure 5). [79, 82, 83]. These modifications are "written", "read" and "erased" by different histone modulating enzymes [84, 85].
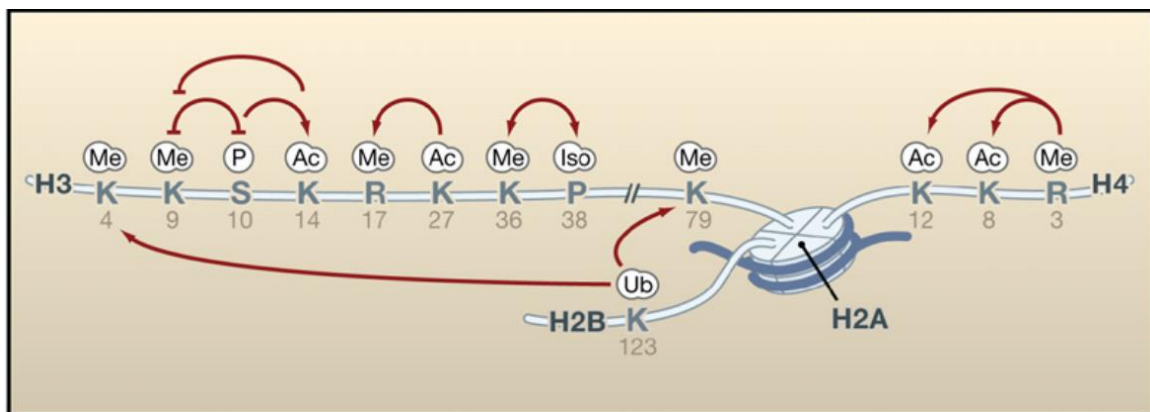
Figure 5. The negative and positive crosstalk between histone post-translational modifications. *Adapted from Kouzadaries, 2007 [79].*

Regarding histone acetylation, gene transcriptional activity is balanced due to alterations of electrostatic charge in the nucleosomes [86]. Therefore, hyperacetylation is characteristic of euchromatin by decreasing the histone-DNA affinity and allowing gene transcription, whereas hypoacetylation is related with heterochromatin [79]. Histone acetylation is "written" by histone acetyltransferases (HATs) and "erased" by histone deacetylases (HDACs) [84].

Histone methylation promotes transcription activation or repression depending on the residue and the number of methylation molecules added (mono-, di- or tri-) [87]. Indeed, tri-methylation of lysine 4 of H3 (H3K4me3) promotes active transcription while mono- and tri-methylation of lysine 27 of H3 (H3K27me and H3K27me3) inhibits gene transcription. The writers of histone methylation are histone methyltransferases (HMT) and the erasers are histone demethylases (HDM) [79, 82].

Histone modifying enzymes expression are disrupted in cancer and the imbalance between writers and erasers affect the PTMs' profile [88]. Moreover, DNMTs are directly recruited by HMTs to inhibit genes' expression and recruit HDACs to increase the gene silencing. This interplay between DNA methylation and PTMs is also impaired in cancer [88].

2.4. HISTONES VARIANTS

The less studied epigenetic mechanism is the shift of canonical histones by sequential similar non-allelic histones variants [89]. Among species, histone variants are the mostly conversed proteins and have been considered functionally irreplaceable [90, 91].

On one side, canonical histones are genomically organized by clusters lacking introns [92]. The transcription is DNA replication-dependent and therefore, exclusive to the S phase of the cell cycle, and the mRNA obtained contains a unique 3′ stem loop [93]. On the other side, histone variants are orphan genes with introns and the mRNA translated holds a polyadenylated tail [81]. Although they are present throughout the cell cycle, variants are tissue and temporal-specific [94, 95]. Variants are named "replacement histones" because they substitute the canonical histones during development and differentiation, establishing cell identity [81].

The slightly sequential differences, along with unique PTMs of histone variants, result in nucleosome-DNA stability differences [96] and alters the efficiency of protein complexes responsible from histone deposition and displacement in the nucleosome. These adjustments change the accessibility of transcription factors into the chromatin, regulating the gene expression.

To date, histone variants have been described for all canonical histones, excluding H4 (Figure 6) [97]. H2A family is the largest histone family with the most structurally diverse histone variants: H2A.X, H2A.Z, macroH2A, H2A.Bbd [98]. Variants of H2A are described by distinguish length, sequence and genome distribution [81]. Mis-regulation or mutations in these H2A histone atypical variants have been implicated in cancer initiation and progression [89].

17
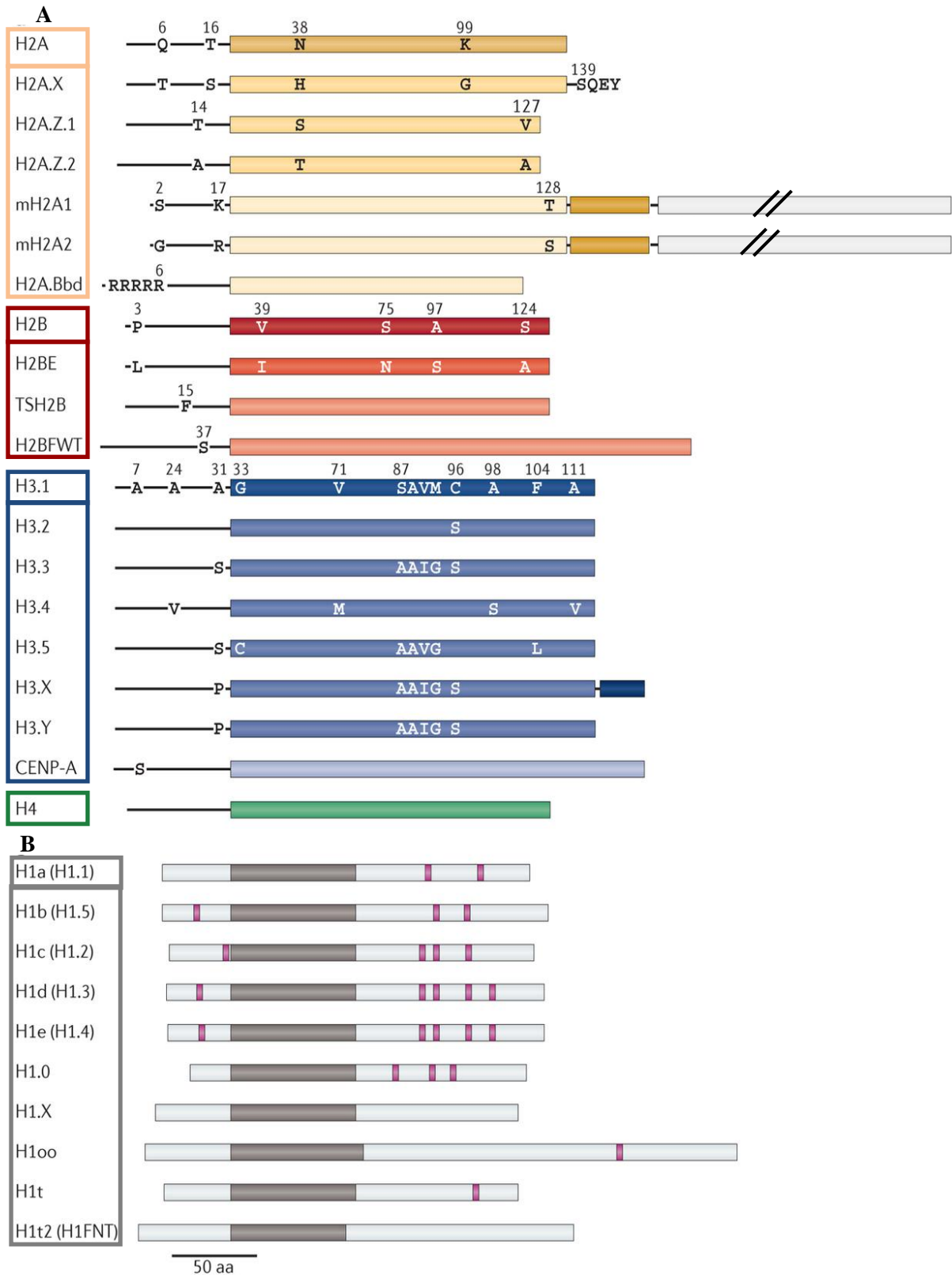
Figure 6 A. Human canonical and histone variants of H2A (yellow), H2B (red), H3 (blue), H4 (green). Unstructured amino- terminal tails are shown as black lines. Specific amino acid residues are depicted at key differences among variants of a common histone protein family. Different shades of color are used to indicate protein sequences that are highly divergent between canonical histones. B. Human canonical and histone variant linker H1. Unstructured amino- terminal tails are shown as light grey. Globular domains are shown in brown. Serine/threonine PXK phosphorylation sites targeted by cyclin-dependent kinases are indicated in magenta. Alternative names of variants are given in parentheses. aa, amino acid; mH2A1, macroH2A1. *Adapted from Maze and al, 2014 [97].*

2.4.1. MACROH2A1: THE SUBDOMAINS

Currently, there are two described macroH2A histones, macroH2A1 and macroH2A2, which are encoded by *H2AFY* and *H2AFY2* genes, respectively [99]. The vertebrates-exclusive macroH2A1 is one of the most distinctive histone variant [81], composed by 370 amino acids with 40kD of molecular weight in a tripartite structural organization, three times larger than canonical H2A.

The first 122 amino acids of the N-terminal histone-like region shares 64% homology with canonical H2A [99]. The last 209 amino acids from the C-terminal histone domain establish a strongly folded macro domain of 30kDa with a random coil without any similarity to other histones [99]. This non-histone region prolongs out from the asymmetric nucleosome, disturbing the transcription factors binding, and contains a putative leucine-zipper motif responsible for nucleosomes interactions [90]. A lysine-rich linker sequence connects the H2A-region with the macro domain that increases nucleosome stability (Figure 7) [100].
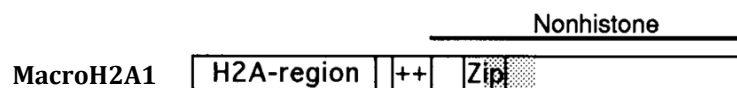
Figure 7 Structure and subdomains of macroH2A1. ++ indicates a lysine-rich linker region that resembles part of the C-terminal domain of histone H1, Zip indicates a region that resembles a leucine zipper, and the gray region shows the location of the region that is different between macroH2A1.1 and 1.2. The region C-terminal to the lysine-rich region is referred to as the non-histone region. *Adapted from Pehrson and Fuji, 1998 [90].*

Chakravarthy et al. demonstrated, by *in vitro* studies, that the presence of macroH2A in the nucleosomes increases DNA-nucleosome stability [101]. MacroH2A1 is responsible for chromatin compaction by affecting the dynamic properties of nucleosomes and even by strengthen the H1-DNA interactions [100, 102]. Overall, nucleosomes containing macroH2A1 are more stable and establish DNA-histone interactions in the entry and exit of the DNA into the nucleosomes leading to formation of heterochromatin regions.

H2AFY (5q31.1) codifies for macroH2A1.1 and macroH2A1.2 isoforms by alternative splicing of a single exon. The difference of these isoforms is a sequence of 10 amino acids in the macro domain of macroH2A1.1 and an alternative exon of 11 amino acids in the macroH2A1.2 [103].

2.4.2. MACROH2A1: EXPRESSION AND NUCLEOSOME DEPOSITION REGULATIONS

MacroH2A1 can be regulated, post-translationally, by alternative splicing or by nucleosome deposition and chromatin localization.

Concerning expression regulation, macroH2A1 isoforms may occur by pre-mRNA splicing regulators, as the Quaking (QKI) or RNA helicases deadbox 5 (DDX5 or p68) and deadbox 17 (Ddx17 or p72) [104, 105]. The RNA-binding protein QKI is a 37kDa RNA-binding protein known to upregulate macroH2A1.1 through binding at the intron upstream of macroH2A1.2-specific exon [104]. RNA helicases DDX5 and DDX17, which share a highly degree of homology, were also reported to control macroH2A1 splicing [105].

Regarding nucleosome regulation, ATP-dependent nucleosome remodeling complexes, such as switch/sucrose non-fermentable (SWI/SNF), rearrange or mobilize nucleosomes and allow the shuffle of histones [106]. Specifically, SWI/SNF helicase ATRX (α-thalassemia/MR, X-linked) acts as negative regulator of macroH2A1 in chromatin-free state, monitoring macroH2A1 deposition. Indeed, increased nucleosomes deposition of macroH2A has been reported for ATRX syndrome patients [107]. Moreover, the Aprataxin-PNK-like factor (APLF), a poly(ADP-ribosyl)ation-regulated protein, acts as DNA damage-specific histone chaperone, increasing the macroH2A1 nucleosome deposition [96, 108].

For nucleosome stability, nuclear respiratory factor 1 (NRF-1) interacts with macroH2A1, avoiding unnecessary gene expression [109]. Nonetheless, macroH2A1 plays a complex role in transcription by either positive or negative gene's regulation [110]. The switch key between repressor/activator roles of macroH2A1 remains uncertain, though PTMs, splicing alternative or the transcription factors nature may be predict the macroH2A1 function [110].

2.4.3. MACROH2A1 TARGETS

Firstly found in the female X inactivation chromosome, by Costanzi and Pehrson in 1998, macroH2A1 were generally described as an autosomal transcriptional repressive histone [111, 112] implicated in female X chromosome inactivation [113-115]. Nonetheless, macroH2A1 is found in about 25% of the genome, being similarly expressed in female and male mammals [116].

MacroH2A1 is present in Senescence-Associated Heterochromatic Foci (SAHFs) [117] and upstream and downstream of transcription start sites (TSSs) of genes implicated in cell cycle [110], pluripotency [118] and development [119]. As previously described, a subset of genes can be positively regulated by this histone variant [120] or even be recruited for DNA repair [121].

MacroH2A1.1 expression is generally limited to differentiated cells, whereas macroH2A1.2 is highly expressed in proliferating cells [122]. Embryonic stem (ES) cells exclusively express macroH2A1.2 and macroH2A1.1 only during development increases [123].

Regarding PTMs, ubiquitylation of macroH2A1.2 at lysines115 and 116 (K115 K116) were associated with the X chromosome inactivation and phosphorylation at serine 137 (S137) of both macroH2A1.1 and macroH2A1.2 were implicated in cell cycle regulation (Figure 8) [89]. Remarkably, macroH2AS137ph is excluded from female X chromosome inactivation and enriched during mitosis, suggesting phosphorylation as a key regulation for critical interactions of macroH2A1 with effector molecules [115, 120]. Additionally, PTMs without a known function were also described for macroH2A1.2, specifically, methylation at lysines 17, 122 and 238 (K17, K122, K238) and phosphorylation at threonine 128(T128) (Figure 8) [98].
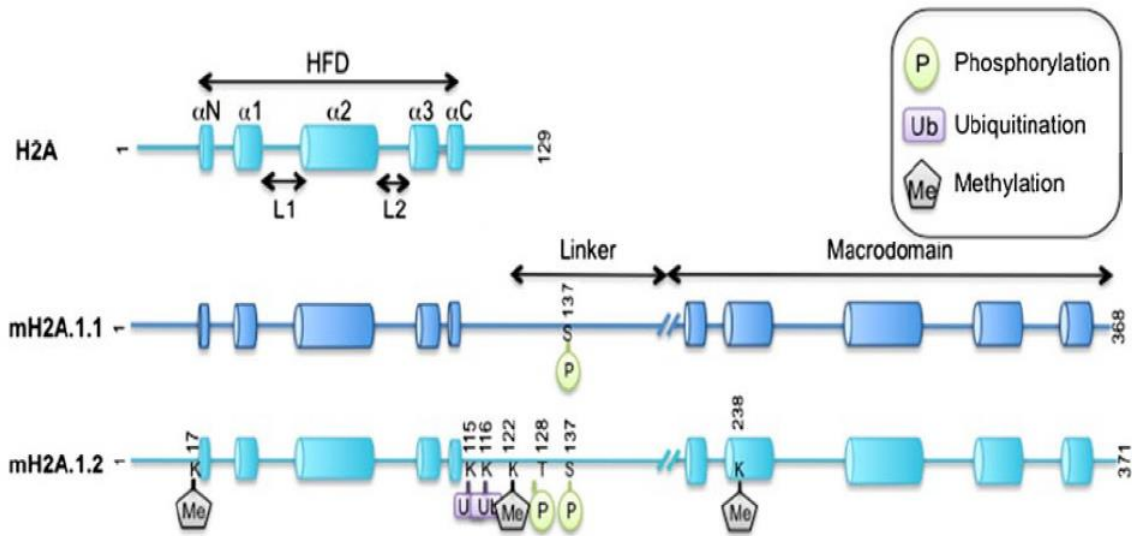
Figure 8. Schematic representation of the canonical histone H2A and histone variants macroH2A1 isoforms with relative PTMs. Specific amino acids are depicted when they are found to be post-transnationally modified (PTMs are indicated by symbols as shown in the legend). Cylinders depict alpha-helical structures. mH2A1.1, macroH2A1.1. mH2A1.2, macroH2A1.2. *Adapted from Vardabasso et al, 2013 [89].*

MacroH2A1.1-specific exon allows macroH2A1.1 to exclusively bind NAD+-derived ligands, including the poly-ADP-ribose (PAR) produced by poly-ADP-ribose polymerase 1 (PARP-1), disturbing several pathways [124]. Indeed, while macroH2A1.2 deposition is related with the repressive H3K27me3, the splicing isoforms exchange, leads PARP-1 to recruit HATs, through macroH2A1.1-PAR binding, promoting H2B K12ac and H2B K120ac thus, activating gene transcription [124].

In addition to PARP-1downregulation, macroH2A1.1 negatively regulates cellular proliferation and cell cycle progression genes, such as cyclin-dependent kinase 8 (CDK8) and *c-Fos*, [104, 110, 125] and upregulates redox metabolism genes [98]. On the other hand, macroH2A1.2 is known for promoting the receptor tyrosine-protein kinase (ERBB2) expression, inducing proliferation [126].

3. THE ROLE OF MACROH2A1 IN CARCINOGENESIS

Expression alterations of macroH2A1 has been studied in several types of cancer, including those of breast [127], lung [128] colon [103], bladder [129] and melanoma [125].

MacroH2A1 expression levels alterations in cancer might occur due to altered pre-mRNA alternative splicing. QKI is reported to be downregulated in several tumors, leading to macroH2A1.1 decreased expression levels [104]. Consequently, macroH2A1.2 expression increases in relation to macroH2A1.1. In breast cancer, DDX5/DDX17 inhibits the expression of redox metabolism related-genes as SOD3 (Superoxide Dismutase 3) and HAO1 (Hydroxyacid Oxidase 1), through macroH2A1.1 downregulation [98, 105].

In non-metastatic breast cancer cells, overexpression of macroH2A1.1 activates SOD3 gene that inhibits cell migration/invasion, whereas in metastatic breast cancer cells, high levels of macroH2A1.2 represses SOD3, increasing cell migration/invasion [105].

Moreover, in breast cancer, macroH2A1.2 overexpression induces ERBB2 overexpression, promoting uncontrolled cell growth and proliferation [126]. However, in triple-negative breast cancer (TNBC), without estrogen and progesterone receptors and absence of amplification of ERBB2, macroH2A1.1 overexpression associated with poor survival [127].

Overall, macroH2A1.1 variant is commonly accepted as a pleiotropic tumor suppressor, by repressing cell proliferation and cell migration, whereas macroH2A1.2 seems to play an oncogenic role, contributing to a metastatic phenotype, and promoting cell proliferation through ERBB2 [98, 105, 121].

II . AIMS

AIMS

PCa is one of the most prevalence cancers in men and cause of cancer-related morbidity and mortality worldwide. Characteristically asymptomatic in early stages, diagnostic and prognostic tools for PCa are fallible and for PCa at late-stage, the therapeutic options are limited. To avoid unnecessary treatments in indolent tumors and choose the better options for PCa patients is important to fully understand the biology of this neoplasm.

Histone variants, the less studied epigenetic mechanism, have been implicated in cancer, regulating several important cellular pathways. Specifically, the histone variant isoforms macroH2A1.1 and macroH2A1.2 were found to be deregulated in some neoplasms; however no data is yet available for prostate cancer. Thus, the major goal of this Master Thesis is to evaluate whether macroH2A1 is deregulated in prostate cancer and whether it plays a critical role in prostate tumorigenesis. To achieve this, several objectives were established:

One. Determine the *H2AFY* expression levels, both total and isoforms, in a cohort of primary prostate tumors (n=197), prostate intraepithelial neoplasia (n=45) and in normal prostatic tissue (n=14).

Two. Study putative regulators of H2AFY.

Three. Correlate the H2AFY transcripts with clinical pathological variables.

Four. Assess the protein expression of *H2AFY* in prostatic tissues.

Five. Determine the *H2AFY* transcripts expression levels (total and isoforms) in a benign prostatic cell line (RWPE-1) and several PCa cell lines (22Rv1, DU145, LNCaP, PC-3 and VCaP).

Six. Evaluate the phenotypic impact of *H2AFY* shRNA or forced expression in PCa cell lines.

III . MATERIALS AND METHODS

1. CLINICAL SAMPLES

1.1 PATIENTS AND CLINICAL SAMPLES COLLECTION

Prostate samples of 197 primary tumors and 45 HGPIN (from here simply designated PIN) lesions were prospectively collected from patients diagnosed with the disease and primarily treated with RP, form 2001 and 2006, at the Portuguese Oncology Institute, Porto, Portugal. Samples of 15 morphological normal prostate tissues (MNPT), used as control, were collected from the peripheral zone of prostates not harboring PCa, obtained from radical cystoprostatectomy for bladder cancer. Immediately after surgery, all tissue specimens were frozen at -80ºC. Thick frozen sections were obtained from frozen tissues for stain identification and after, the tissue block was trimmed to maximize the yield of target cells (>70% of target cells). Histological slides from formalin-fixed, paraffin embedded (FFPE) tissue fragments were also obtained from the same surgical specimens for histopathological examination: Gleason Score and pathological staging evaluations. Relevant clinical data were acquired from clinical registers and these studies were approved by the institutional review board (Comissão de Ética para a Saúde – CES 019/2008) of Portuguese Oncology Institute - Porto, Portugal.

1.2 RNA EXTRACTION AND QUANTIFICATION

Samples were homogenized in Trizol® Reagent (Invitrogen, Carlsbad, CA, USA) and the total RNA were extracted from all 257 samples using PureLink™ RNA Mini Kit (Invitrogen, Carlsbad, CA, USA), following manufacturer's instructions. All genomic DNA were eliminated with TURBO DNA-free (Ambion, Applied Biosystems), according to manufacturer's instructions. The concentration, purity ratios and quality of each sample were determined using a Nanodrop ND-1000 (ThermoScientific, Wilmington, DE, USA) and by an agarose gel electrophoresis. RNA samples were then stored at -80ºC.

1.3 QUANTITATIVE REVERSE TRANSCRIPTION PCR (RT-qPCR)

For each tissue sample, first strand synthesis was performed using the TransPlex® Whole Transcriptome Amplification Kit (Sigma-Aldrich®, Schnelldorf, Germany) and QIAquick PCR Purification Kit (QIAGEN, Germany) for purification.

Expression of the target genes were quantified using Fast SYBR Green® Gene Expression Assay (Applied Biosystems®, Life Technologies™, Foster City, CA, USA), and normalized to the expression of the endogenous control β-glucuronidase (*GUSβ*), a housekeeping gene (Table 3 and Figure 9). In each well, 0.1μL of cDNA samples were mixed with 5μL of 2x KAPA SYBR® FAST qPCR Master Mix Universal (Applied Biosystems®), 0.2μL of 50x ROX low and optimized for 0.2-0.4μL of 10μM primers (Sigma-Aldrich®), completed with sterile bidistilled water (B. Braun, Melsunger, Germany) for a total of 10μl. Each 96-well plate included 2 negative controls and, for standard curve, five sequential dilutions of a cDNA from human prostate RNA (Ambion®, Invitrogen, Carlsbad, CA, USA). PCR were programmed for 3 minutes at 95ºC, followed by 40 cycles of 3 seconds at 95ºC and 30 seconds at 60ºC. Relative expression was obtained by the ratio of the target mean quantity/reference gene mean quantity. All samples were analyzed in triplicate in 7500 Real-Time PCR system (Applied Biosystems®), and the mean value was used for data analysis.

Table 3. Primers sequences for macroH2A1 isoforms and total, splicing regulators and control primers [104, 105].

Gene	Forward Primer 5' -> 3'	Reverse Primer 5' -> 3'
MacroH2A1.1	GGCTTCACAGTCCTCTCCAC	GGTGAACGACAGCATCACTG
MacroH2A1.2	GGCTTCACAGTCCTCTCCAC	GGATTGATTATGGCCTCCAC
MacroH2A1	TCCATTGCATTTCCATCCATCGGC	ACACGAAGTAACTGGAGATGGCCT
QKI	ATTAAACGGTCCCCTGAAGC	ATCAACAGCCCAAG TGTGAC
DDX5	GTAGCTCAGACTGGATCTGG	TCTCTAGGAATGGCTGGTGG
DDX17	AGAAGTAGCAAGACTGACTCC	CCCCCTCTCACTGTAATCTC
GUSβ	CTCATTTGGAATTTTGCCGATT	CCGAGTGAAGATCCCCTTTTA
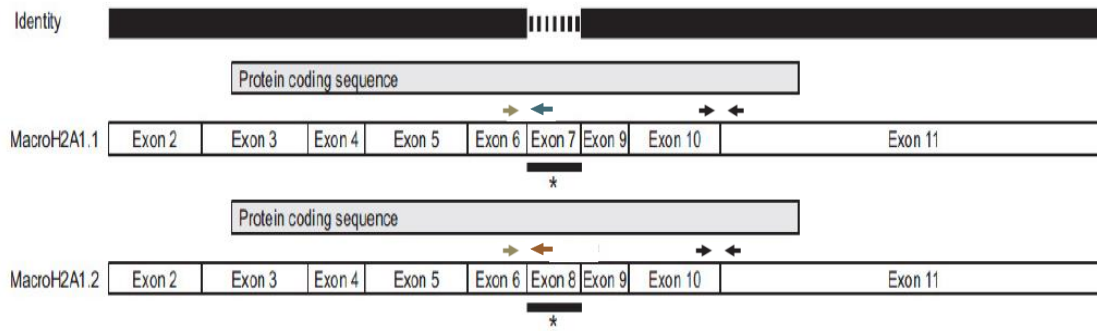
Figure 9 Primers used for expression quantification of total macroH2A1 and isoforms. Specific-reverse primers for macroH2A1.1 and macroH2A1.2 are in blue and orange, respectively, and the forward primer used for both isoforms is in grey. The set of primers for macroH2A1 are in black.

2. IMMUNOHISTOCHEMISTRY

Histological slides from FFPE tissue fragments were also obtained from the same surgical specimens and assessed for Gleason Score and TNM stage. Firstly, slides were deparaffinized in xylene (Sigma-Aldrich®, St. Louis, MO, USA) and then hydrated in a decreasing series of ethanol solutions (Merck, Darmstadt, Germany). Epitope retrieval was performed with ethylenediaminetetraacetic acid (EDTA) buffer (Thermo Scientific, Waltham, MA, USA) for 30 minutes, in a microwave at 700W. Endogenous peroxidase activity was neutralized for 20 minutes with 0.6% hydrogen peroxide (Merck). Protein detection was performed using the NovolinkTM Max Polymer Detection System (Leica Biosystems, Nussloch, Germany), according to manufacturer instructions. Slides were incubated with a rabbit monoclonal antibody specific for macroH2A1.1 (#12455; Cell Signaling Technology, Inc., Danvers, MA, USA) in a 1:1000 dilution at 4ºC, overnight inside a humid chamber. Subsequent washing steps were performed with tris-buffered saline with Tween® 20 (TBS-T) (Sigma-Aldrich®). Antigen-antibody binding reaction was revealed through the slides incubation for 7 minutes, in the dark, in a 0.05% (m/v) 3,3'-diaminobenzidine (DAB) solution (Sigma-Aldrich®) in phosphate-buffered saline (PBS) (Biochrom Ltd., Cambridge, United Kingdom) previously activated with a 0,1% hydrogen peroxide solution. Counterstaining of the slides was obtained with hematoxylin (Merck) for about 5 seconds and then slides were washed for 1 minute in a 0.25% ammonium solution (Merck). Lastly, the slides were dehydrated in an increasing series of ethanol content and diaphanized in xylene. After the coverslip was mounted, slides were dried. The FFPE tissues fragments that were not incubated with the antibody were used as negative

control of the immunohistochemistry (IHC) reaction. For positive control, FFPE tissue from a normal testis was also included. Slides were observed at the optical microscope and evaluated for macroH2A1.1 immunoexpression by an experienced uro-pathologist. Scoring criteria were adapted from a previous publication of our research group [130]: samples with ≤10% of positive cells were considered "negative expression"; >10-50% of positive cells were considered "intermediated expression"; >50% of positive cells, samples were categorized as "positive expression".

3. CELL LINES STUDIES

3.1 PROSTATE CANCER CELL LINES

RWPE-1 were generously provided by Professor Margarida Fardilha from the University of Aveiro, Portugal; 22Rv1 cells were kindly provided by Dr. David Sidransky at the Johns Hopkins University School of Medicine, Baltimore, MD, USA; DU145 was obtained from the American Type Culture Collection (ATCC, Lockville, MD, USA) whereas LNCaP, PC-3 and VCaP cells were kindly ceded by Prof. Ragnhild A. Lothe from the Department of Cancer Prevention at the Institute for Cancer Research, Oslo, Norway. Regarding metastatic cell lines, LNCaP and VCaP are non-invasive and hormone-sensitive, while DU145 and PC-3 are invasive and hormone-refractory. For further *in vitro* studies for macroH2A1.1 overexpression, LNCaP were selected. The six cell lines used in this study were treated as optional growth medium and supplemented as recommended (Table 4) with 1% of Penicillin-Streptomycin (GIBCO®, Invitrogen, Carlsbad, CA, USA).

Cells were maintained in an incubator at 37ºC with 5% CO_2. To harvest the cells for subculture, TrypLE™ Express (GIBCO®) dissociation reagent was used. All prostate cell lines were karyotyped by G-banding (for validation purposes) and routinely tested for *Mycoplasma spp*. contamination (PCR Mycoplasma Detection Set, Clontech Laboratories).

Table 4. PCa cell lines used and the growth conditions.

PCa cell line	Description	Androgen dependence	Growth medium	Supplements
RWPE	Mimic normal prostate cells	Dependent	K-SFM	Bovine pituitary extract (BPE) + Human recombinant epidermal growth factor (EGF)
22Rv1	Mimic prostate cancer cells	Dependent	RPMI-1640	10% Fetal bovine serum (FBS)
DU145	Derived from brain metastasis	Independent	MEM	
LNCaP	Derived from lymph node metastasis	Dependent	RPMI-1640	
PC-3	Derived from bone metastasis	Independent	50% RPMI-1640 + 50% F-12	
VCaP	Derived from vertebral metastasis	Dependent	MEM	

3.2 RNA EXTRACTION

Three cell culture flasks (75 cm^3) at 100% confluence for each cell line were dissociated in TrypLE™ Express (GIBCO®) and centrifuged at for 5 minutes. After removed the supernatant, cell pellets were ressuspended in 1 mL of PBS (GIBCO®), and centrifuged again. Total RNA from normal prostate cell line and PCa cell lines were extracted by TRIzol® Reagent (Invitrogen), according to manufacturer recommendations. In summary, cell pellets were resuspended in 500μL of TRIzol® Reagent and, after homogenization with a sterile syringe with a 0.9mm needle, 500μL of TRIzol® Reagent were again added. After 5 minutes of incubation at room temperature, 200μL of chloroform (Merk, Darmstadt, Germany) were added and tubes vortexed for 15 seconds. Incubation of 3 minutes at room temperature followed a centrifugation at 11,900g during 15 minutes at 4ºC. The upper phase were transferred to a new RNase-free tube with 500μL of 100% isopropanol, mixed by inversion. Tubes were incubated during 10 minutes at room temperature for RNA precipitation and then,

centrifuged at 11,900g during 10 minutes at 4ºC. The supernatant were discarded and the pellet resuspended in 1mL of 75% ethanol, vortexed and centrifuged during 5 minutes at 8,600g at 4ºC. To efficiently wash the pellets, the last step were performed two times. After supernatant removal, RNA pellets were allowed to air dry for 15 to 20 minutes.

The pellets were eluted in RNA storage solution (1mM sodium citrate, pH 6.4) (Ambion®, Life Technologies™, Foster City, CA, USA). The concentration, purity and integrity of RNA samples were determined using a NanoDrop ND-1000 spectrophotometer (NanoDrop, Technologies, Wilmington, DE, USA) and electrophoresis. RNA samples were stored ate -80ºC.

3.3 cDNA SYNTHESIS

In order to evaluate mRNA expression in prostate cell lines, 1000 ng of cDNA was synthesized from total extracted RNA. A cDNA synthesis was performed using the High Capacity cDNA Reverse Transcription Kit (Applied Biosystems®) according manufacturer instructions. In summary, RNA samples were adjusted to 100ng/μL, completed with nuclease-free water (Exiqon) for a final volume of 10μL. For each RNase-free PCR tube, were added, on ice: 2μL of 10x Reverse Transcript (RT) Buffer, 0.8μL of 25x dNTP Mix (100mM), 2μL of 10x RT Random Primers, 1μL of MultiScribe™ Reverse Transcriptase, 1μL of RNase Inhibitor and 3.2μL of nuclease-free water. Tubes were gently centrifuged and reverse transcription was performed in a Veriti® Thermal Cycler (Applied Biosystems®) at 25ºC during 10 minutes, followed by an incubation of 37ºC for 120 minutes and 5 minutes at 85ºC. Samples were then stored at -20ºC.

3.4 QUANTITATIVE REVERSE TRANSCRIPTION PCR (RT-qPCR)

Expression of macroH2A1 total and isoforms and splicing regulators were assessed using Fast SYBR Green® Gene Expression Assay (Applied Biosystems®, Life Technologies™, Foster City, CA, USA), and normalized to *GUSβ*. In each well, 4.5μL of cDNA samples were mixed with 5μL of 2x KAPA SYBR® FAST qPCR Master Mix Universal (Applied Biosystems®), 0.2μL of 50x ROX low and optimized for 0.2-0.3μL

of 10μM primers (Sigma-Aldrich®), completed with sterile bidistilled water (B. Braun, Melsunger, Germany) for a total of 10μl. No cDNA were added in two wells of the 96-well plate, as negative controls. All samples were analyzed in biological and experimental triplicate in 7500 Real-Time PCR system (Applied Biosystems®). Results were normalized for reference gene and relative expression was obtained by the ΔΔCt method.

4. TRANSFECTION STUDIES

4.1 MACROH2A1.1 OVEREXPRESSION IN LNCaP

Overexpression of macroH2A1.1 were successfully stable and achieved through pEZ-Lv105 (GeneCopoeia™, Rockville, MD, USA) in LNCaP cell line using FuGENE® HD Transfection Reagent (Promega, Madison, WI, USA), following manufacturer's recommendations. Briefly, cells were incubated, one day before transfection, in a 6-well cell-plate at optimized concentration ($2x10^4$ cells/mL). At the time of transfection, cells had reached 30-50% of confluence. Opti-MEM™ medium (GIBCO®) were used to dilute 2μg of transfection molecules to a final volume of 100μL. At room temperature, 4μL of transfection reagent were added for a, firstly optimized, 2:1 FuGENE® HD Transfection Reagent:DNA ratio. This mixture were incubated at 15 minutes at room temperature and then added to the cells in growth medium. Cells were maintained in a humidifying chamber at 37ºC with 5% CO_2. Non-transfected cells were used as positive control. Transfection cells with DNA molecules resistant to Puromycin were used as negative control (NC) to regulate transfection reagent toxicity. These DNA molecules were tested and to not produce any effect in human cells. After 48h of transfection, stable clones with the vector were selected with Puromycin dihydrochloride (cat. 631306, Clontech Laboratories Inc.) at a final concentration of 0.5μg/mL.

4.2 MACROH2A1.1 EXPRESSION ASSAY

The efficiency of macroH2A1.1 overexpression in LNCaP cell line were confirmed by the determination of the transcript levels by RT-qPCR. RNA extraction

and cDNA synthesis were performed as previously designed. Expression of macroH2A1 isoforms were performed using Fast SYBR Green® Gene Expression Assay, as previously described for cell lines.

4.3 PROTEIN EXTRACTION AND QUANTIFICATION

Total protein was extracted from cell lines using the Kinexus Lysis Buffer with Lysis Buffer Cocktail (Kinexus Bioinformatics Corporation, Vancouver, British Columbia, Canada). Succinctly, growth medium was removed from cell culture flasks, and cells were washed 2x with PBS. After total PBS removal, 100 μL of Kinexus Lysis Buffer with Lysis Buffer Cocktail were added to each flask. To increase cell detachment and lysis, cells were scrapped with a cell scrapper (Santa Cruz Biotechnology Inc.). Later, cells were transferred to a 1.5 mL tube and sonicated in ice for 6 cycles of 15 seconds, with 15 seconds gap between each cycle. Tubes were then centrifuged for 30 minutes at maximum speed at 4ºC. Supernatant was carefully transferred to a new tube.

Protein concentration was determined using Pierce BCA Protein assay Kit (Thermo Scientific, Waltham, MA, USA) following to manufacturer's instructions and protein samples were stored at -80ºC.

4.4 SDS-PAGE AND WESTERN BLOT

Protein expression of macroH2A1.1 and macroH2A1.2 in transfected LNCaP were assessed by Western Blot analysis. Briefly, loading buffer was added in 30μg of total protein and then denatured for 5 minutes at 95ºC. After centrifugation, samples were loaded in a polyacrylamide gel composed by a 10% running gel [10% (w/v) acrylamide/bis-acrylamide solution, 0.375M Tris-HCl pH=8.8, 0.1% (w/s) SDS, 0.1% (w/s) APS and 0.04% (v/v) TEMED] and a 4% stacking gel [4% (w/v) acrylamide/bis-acrylamide solution, 0.062M Tris-HCl pH=6.8, 0.1% (w/s) SDS, 0.1% (w/s) APS and 0.25% (v/v) TEMED]. Protein separation was performed in a drive Mini-Protean 3 Eletrophoresis System (BioRad, Hercules, CA, USA) at 120V in a running buffer (0.025M Tris, 0.192M glycine and 0.1% SDS, pH=8.3).

After SDS-PAGE, proteins were blotted in PVDF membranes (Bio-Rad Laboratories Inc., Hercules, CA, USA) previously activated in 20% (v/v) methanol.

Membranes and filter papers were incubated for 20 minutes and the gel for 10 minutes at room temperature in transfer buffer [0.025m Tris, 0.192M glycine and 20% (v/v) methanol]. Protein blotting occurred in Trans-Blot® Turbo™ Transfer System (BioRad, Hercules, CA, USA) for 9 minutes at 25V.

After incubation for an hour in a blocking solution [5% (w/v) BSA (ChemCruz™ Biochemicals, Santa Cruz Biotechnology Inc.,) in 0.01M Tris-buffered saline containing 0.1% (v/v) Tween 20], membranes were incubated overnight at 4ºC with primary antibodies macroH2A1.1 (1:500; #12455) or macroH2A1.2 (1:500; #4827; Cell Signaling Technology, Inc., Danvers, MA, USA) in blocking solution. Membranes were washed in TBS with Tween and incubated for 1 hour with Goat Anti-Rabbit IgG (1:4000; BioRad, Hercules, CA, USA).

Endogenous control β-actin was assessed following the previous steps with an incubation of 1 hour of monoclonal mouse antibody β-Actin (1:8000; Sigma-Aldrich®) and Goat Anti-Mouse IgG (1:4000; BioRad, Hercules, CA, USA).

The blots were developed with the Western Bright[TM] ECL-spray (Advansta Corporation, Menlo Park, CA, USA) and exposed to Amersham[TM] Hyperfilm ECL (GE[TM] Healthcare, Buckinghamshire, United Kingdom). All the experiments were performed in triplicate.

4.5 CELL VIABILITY ASSAY

The cell viability was assessed in LNCaP to evaluate the effect of macroH2A1.1 overexpression by 3-(4, 5-dimethylthiazol-2-yl)-2,5-diphenyltetrazolium (MTT; Sigma-Aldrich®) assay. The mitochondrial enzyme succinate-dehydrogenase is responsible for the cleavage of the yellow 3-(4,5-dimethylthiazol-2-yl)-2,5-diphenyltetrazolium-bromide in blue-color formazan. Since mitochondrial activity is constant in viable cells, living cells with functional mitochondria, is possible to measure cell viability differences between non-overexpressed and overexpressed cells.

Briefly, 5×10^4 cells/mL were seeded in 96-well plates (Sarstedt, Numbrecht, Germany) in 200μL of complete medium and incubated in a humidified chamber at 37ºC and 5% CO_2. The viability assay was performed 72h after cells adhered to the plate. Cells were incubated with a solution of 20μL of MTT (Sigma-Aldrich®) diluted in 180μL complete medium at 37ºC for 1h. After MTT solution was discarded,

formazan crystals were dissolved in 100μL of Dimethyl sulfoxide DMSO (Sigma-Aldrich®) and plates were shaken for 15 minutes for complete dissolution.

An automated plate reader GloMax®-Multi Detection System (Promega, Madison, WI, USA) at 560nm with a reference filter of 630nm allowed colorimetric quantification. Blank consisted in the media of nine wells contained DMSO (Sigma-Aldrich®). The optical density (OD) was directly proportional to the number of viable cells. Three biological independent experiments were performed with methodological triplicates for each experiment.

5 STATISTICAL ANALYSIS

Differences in quantitative expression levels of macroH2A1 and splicing regulators among MNPT, PIN, PCa were assessed by the non-parametric Kruskall-Wallis test, followed by pairwise comparisons through Mann-Whitney *U*-test. For assess the differences of gene expressions in pair matched PIN and PCa, were used the Wilcoxon Signed Rank test, another non-parametric test. A Spearman Correlation was used to test an association between transcript levels of different genes.

The relationship between expression levels and standard clinicopathological variables (Gleason score, PSA levels) were assessed using the Kruskall-Wallis or the Mann-Whitney tests, as appropriate. A receiver operator characteristic (ROC) curve was used to assess the performance of macroH2A1.1 and QKI as diagnostic biomarker. Fisher's Exact Test was used to seek for differences in the frequency of immunoreactivity for macroH2A1 protein according to the immunohistochemical scoring, among the prostate tissues analyzed.

All tests were two-sided and statistical significance was set at $p<0.05$. Statistical analysis was performed using GraphPad Prism version 5.0 software for Windows (GraphPad Software Inc., La Jolla, CA, USA).

IV . RESULTS

1. MACROH2A1 ISOFORMS GENE EXPRESSION LEVELS

Transcript levels of macroH2A1.1 and macroH2A1.2 were assessed independently in 15 MNPTs, 45 PINs and 197 PCa samples. Clinical and histopathological data of patients included in this study are described in Table 5. Age was not significantly different among groups.

Table 5.Clinical and histopathological data of patients and controls.

Clinicopathological Features	MNPT	PIN	PCa
Number of patients, n	15	45	197
Age (years)			
Median (range)	64 (45-80)	64 (51-75)	64 (49-75)
PSA levels (ng/mL)			
Median (range)	n.a	n.a.	8.3 (2.9-23)
Pathological stage			
pT2, n (%)	n.a	n.a	110 (55.8)
pT3, n (%)			87 (44.2)
Gleason Score			
< 7	n.a	n.a	68 (34.5)
≥7			129 (65.5)

n.a. - not applicable

MacroH2A1.1 expression levels were significantly different among the three groups (KW, p<0.0001). The lowest transcript levels were found in PCa samples. Regarding macroH2A1.2, significantly lower expression was found in PIN lesions (p<0.001). However, no significant differences were apparent between MNPT and PCa (Figure 10).
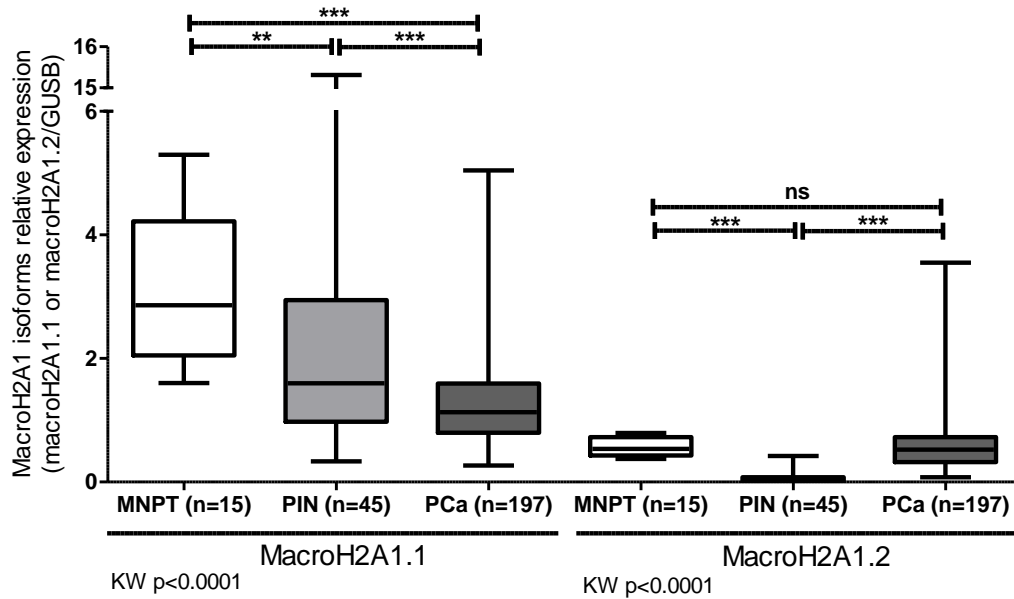
Figure 10. Transcriptional status of macroH2A1.1 and macroH2A1.2 isoforms in clinical samples. MacroH2A1.1 is progressively downregulated through PCa progression and macroH2A1.2 transcript expression was not significant different between MNPT and PCa. Group analysis with Kruskal-Wallis test followed by a pairwise Mann-Whitney U test, **p<0.01 and ***p<0.001, ns = not significant.

2. MACROH2A1 TOTAL GENE EXPRESSION IN PROSTATE

In the same series of clinical samples, total macroH2A1 transcript levels were measured by qRT-PCR. No statistically significant differences were found in total macroH2A1 expression between MNPT and PCa, whereas a significant decrease was observed in PIN (p<0.001) (Figure 11A).

The expression levels of each isoform of macroH2A1 were then normalized for total macroH2A1. The results obtained after this normalization followed the same trend of those obtained after normalization for a housekeeping gene (Figure 11B).
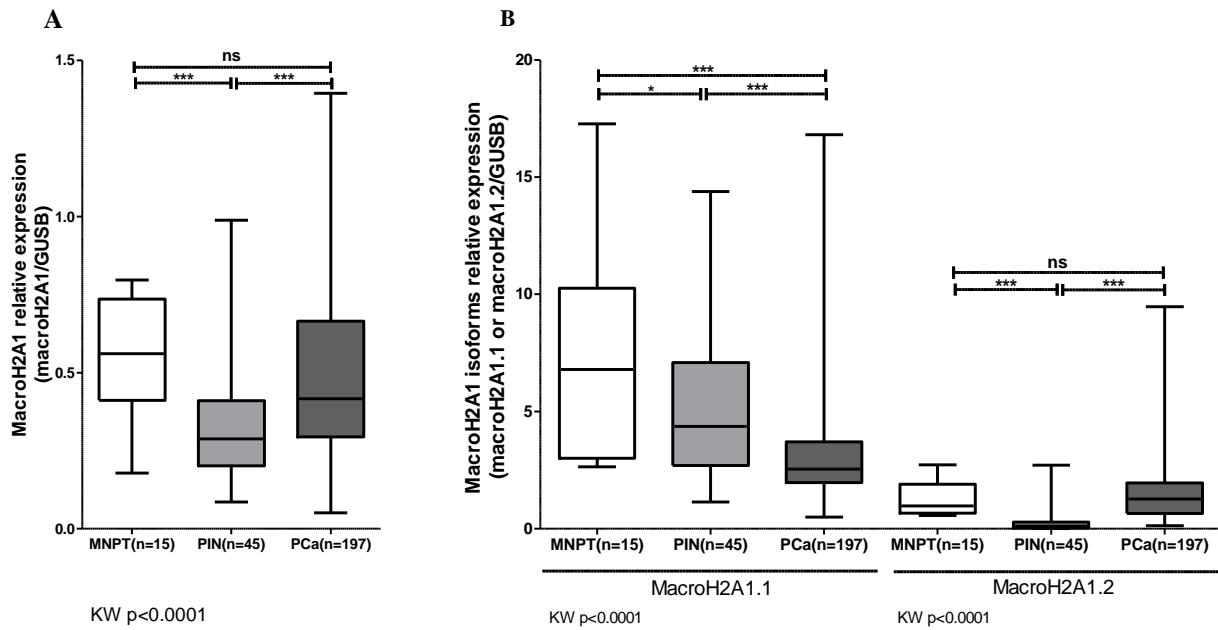
Figure 11. (A) MacroH2A1 gene expression in MNPT, PIN and PCa samples. (B) MacroH2A1 isoforms relative expression relative to total macroH2A1 in clinical samples. Group analysis with Kruskal-Wallis test followed by a pairwise Mann-Whitney *U* test, *p<0.05, ***p<0.001, ns = not significant.

To determine the proportion of macroH2A1 isoforms, macroH2A1.1 was normalized against macroH2A1.2 (Figure 12). In PIN lesions, macroH2A1.1 expression levels were significantly higher than those of macroH2A1.2 ($p < 0.001$), whereas PCa samples displayed a significantly lower ratio compared to MNPT ($p < 0.001$).
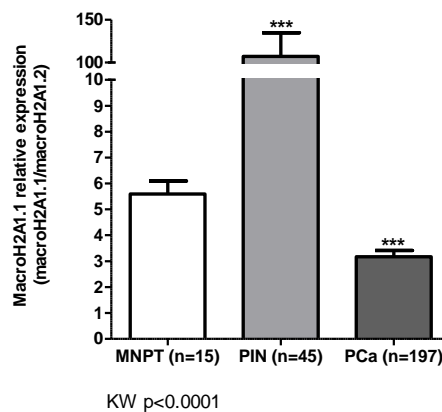
Figure 12. Ratio of isoforms macroH2A1.1 normalized for macroH2A1.2 expression in clinical samples. Group analysis with Kruskal-Wallis test followed by pairwise Mann-Whitney *U* test, ***p<0.001.

45

3. SPLICING REGULATORS OF MACROH2A1 IN PROSTATE TISSUES

To justify the altered ratio between *H2AFY* gene splicing variants, macroH2A1 splicing regulators (QKI, DDX17 and DDX5) transcription levels were also determined in the same sample set (Figure 13).
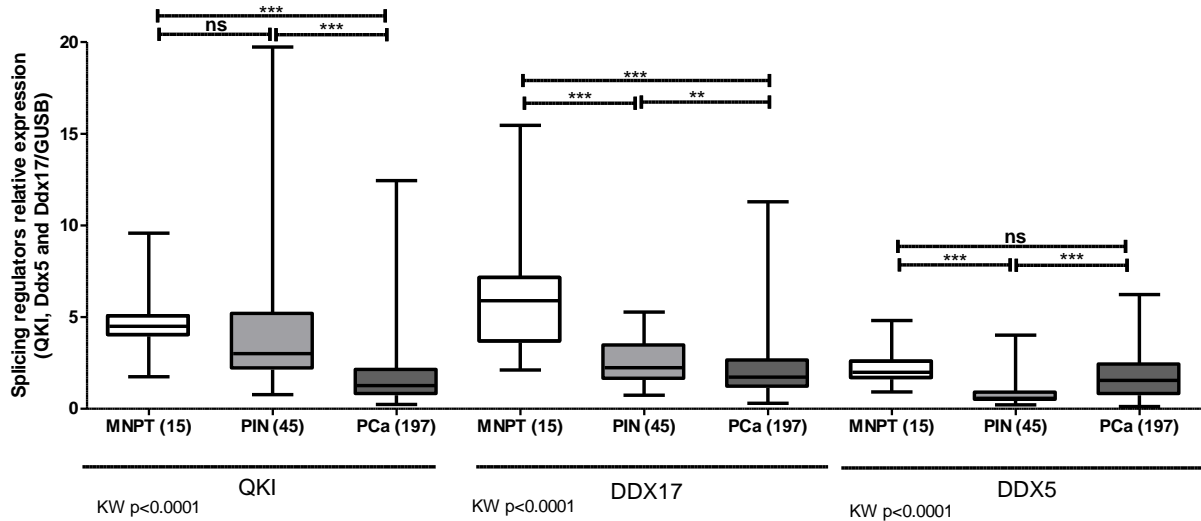
Figure 13. Transcripts levels of splicing regulators of *H2AFY* mRNA in MNPT, PIN and PCa samples. Group analysis with Kruskal-Wallis test followed by pairwise Mann-Whitney *U* test, **p<0.01 and ******p*<0.001, ns = not significant.**

A statistically significant downregulation of QKI and DDX17 was displayed by PCa (p<0.001). Moreover, in PIN lesions, DDX17 expression levels were significantly lower comparing to MNPT (p<0.001). Conversely, no significant differences were apparent between MNPT and PCa samples for DDX5, for which the lowest levels were observed in PIN lesions (p<0.001).

4. CORRELATIONS BETWEEN MACROH2A1 TOTAL AND SPLICE VARIANTS WITH THE THREE MAJOR SPLICING REGULATORS EXPRESSION

We, then, determined the correlation between the expression of the three splicing regulators assessed and the variants and total macroH2A1.1 (Table 6) using Spearman's rho (ρ) correlation [131].

Table 6. Spearman's ρ correlations between total and splice variants of macroH2A1 with three splicing regulators.

GENES	MacroH2A1	MacroH2A1.1	MacroH2A1.2
QKI	0.18, p=0.92	0.56, p<0.001	-0.07, p=0.25
DDX17	0.30, p<0.001	0.41, p<0.001	-0.03, p=0.09
DDX5	0.51, p<0.001	0.22, p<0.001	0.32, p<0.001

Coefficient Values	Strength of Correlation
[-1.00 – -0.80]	Strongly negative
[-0.79 – -0.50]	Moderately negative
[-0.49 – -0.20]	Weakly negative
[-0.19 – 0.19]	No association
[0.20 – 0.49]	Weakly positive
[0.50 – 0.79]	Moderately positive
[0.80 – 1.00]	Strongly positive

The strongest correlation was observed between QKI and the splice variant macroH2A1.1 expression levels, whereas DDX5 transcript levels positively associated with MacroH2A1 total expression. Splicing regulator DDX17 expression levels did not significantly associate with any of the MacroH2A1 transcripts.

5. EXPRESSION LEVELS OF MATCHED PIN AND PROSTATE CANCER SAMPLES

In 35 cases, there were paired PIN and PCa samples (both samples derived from the same patient). Although this does not imply a causal relation between them, it may reflect similar carcinogenic processes within the same prostate. A non-parametric signal test was then performed to assess differences in expression levels of QKI and macroH2A1.1 between matched PINs and primary tumors. Statistically significant differences between matched PIN and PCa lesions were observed for macroH2A1.1 and QKI expression levels (p<0.01 and p<0.001, respectively). Twenty-two PCa samples (63%) expressed lower macroH2A1.1 transcript levels and 24 (69%) expressed lower QKI, comparing with matched PIN lesions (Figure 14).
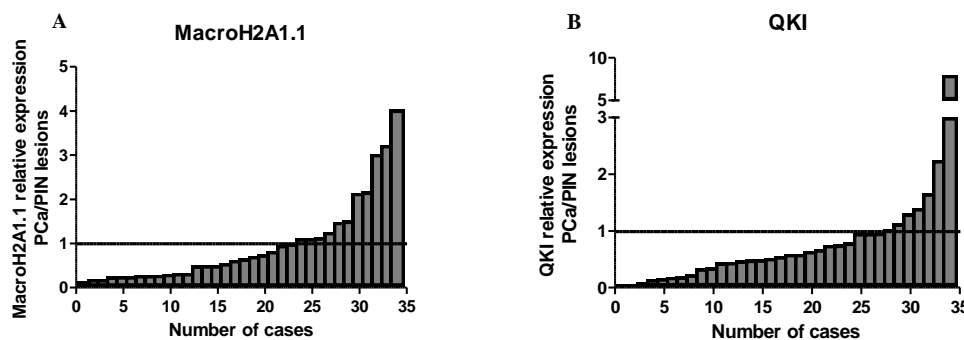
Figure 14. Relative expression of macroH2A1.1 (A) and QKI (B) mRNA levels with matched PCa and PIN lesions.

6. ASSOCIATION BETWEEN *H2AFY* OR SPLICING REGULATORS EXPRESSION AND CLINICO PATHOLOGICAL PARAMETERS

MacroH2A1.1 or QKI transcript levels did not associate with patients' age or pT stage. However, macroH2A1.1 and QKI expression levels significantly associated with Gleason score (Figure 15A e 15B). Indeed, less differentiated tumors displayed lower macroH2A1.1 (p<0.01) and QKI expression levels (p<0.001). Moreover, high macroH2A1.1 expression levels significantly associated with serum PSA levels above 10 ng/mL (p<0.01) (Figure 15C).
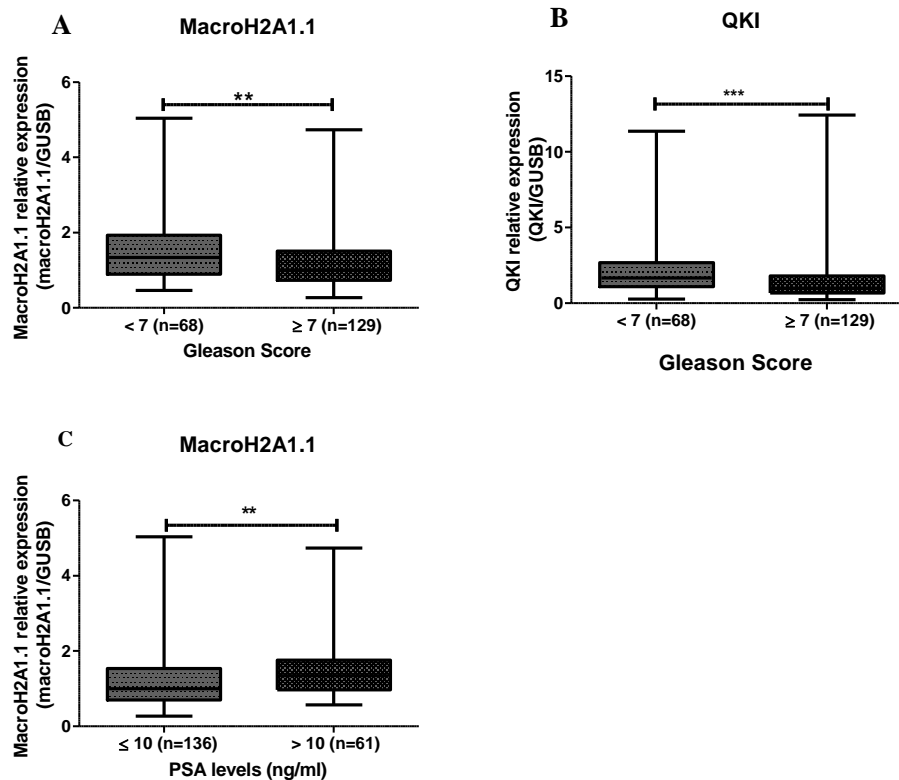
Figure 15. Clinicopathological Gleason Score associations with expressions levels of macroH2A1.1 (A) and QKI (B). Association of PSA levels with macroH2A1.1 transcript levels (C). Pairwise Mann-Whitney *U* test, **p<0.01, ***p<0.001.

7. EVALUATION OF MACROH2A1.1 AND QKI AS DIAGNOSTIC BIOMARKER

Receiver operator characteristic (ROC) curve analysis was performed to assess the diagnostic performance of macroH2A1.1 and QKI in tissue samples (Figure 16 and Table 7). The empirical cut-off values were set to maximize sensitivity and specificity (1.8 for macroH2A1.1 and 2.9 for QKI).
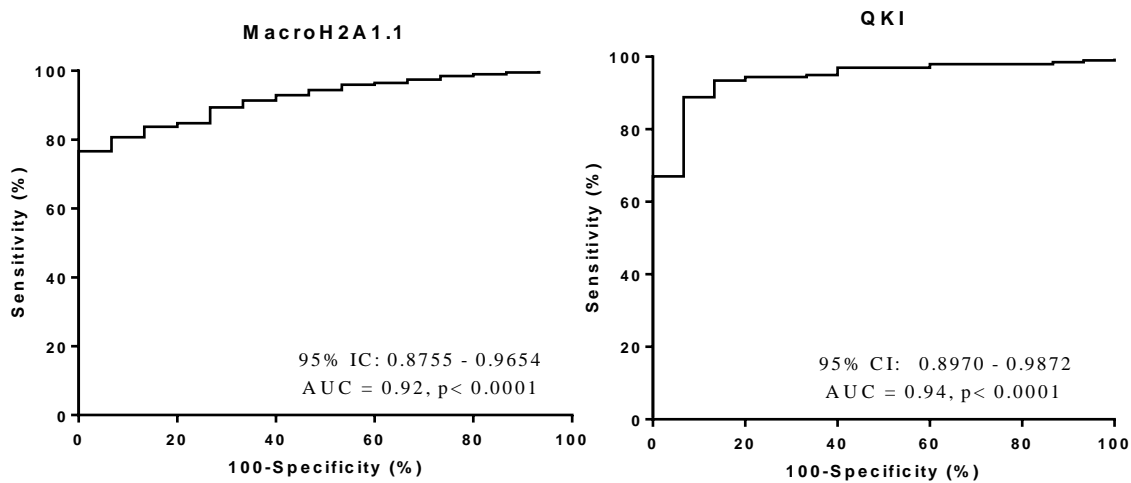
Figure 16. ROC curve analysis of macroH2A1.1 and QKI genes in discriminating PCa from MNPT samples. (AUC, area under curve; CI, confidence interval).

Table 7. Validity estimates for macroH2A1.1 and QKI as diagnostic biomarker for PCa.

Parameter	MacroH2A1.1 performance (%)	QKI performance (%)
Sensitivity	80.7	88.8
Specificity	93.3	93.3
Positive Predictive Value (PPV)	91.9	92.6
Negative Predictive Value (NPV)	97.4	95.8
Accuracy	81.6	88.7

Remarkably, QKI performed better than macroH2A1.1 (sensitivity = 88.8% and specificity = 93.3% for discriminating PCa from MNPT) corresponding to an AUC of 0.94.

8. EVALUATION OF MACROH2A1.1 IMMUNOEXPRESSION IN PROSTATE TISSUES

MacroH2A1.1 protein levels were assessed by immunohistochemistry in a cohort of 243 FFPE prostate tissue samples (Table 8) that corresponds to the same cases analyzed for transcript levels in which representative tissue sections were available.

The distribution of macroH2A1.1 immunostaining in the clinical samples is depicted in Figure 17. Most tissue samples, irrespective of its nature, displayed macroH2A1.1 immunoexpression in over 50% of cells. No statistically significant associations were found between transcript and protein levels of macroH2A1.1.
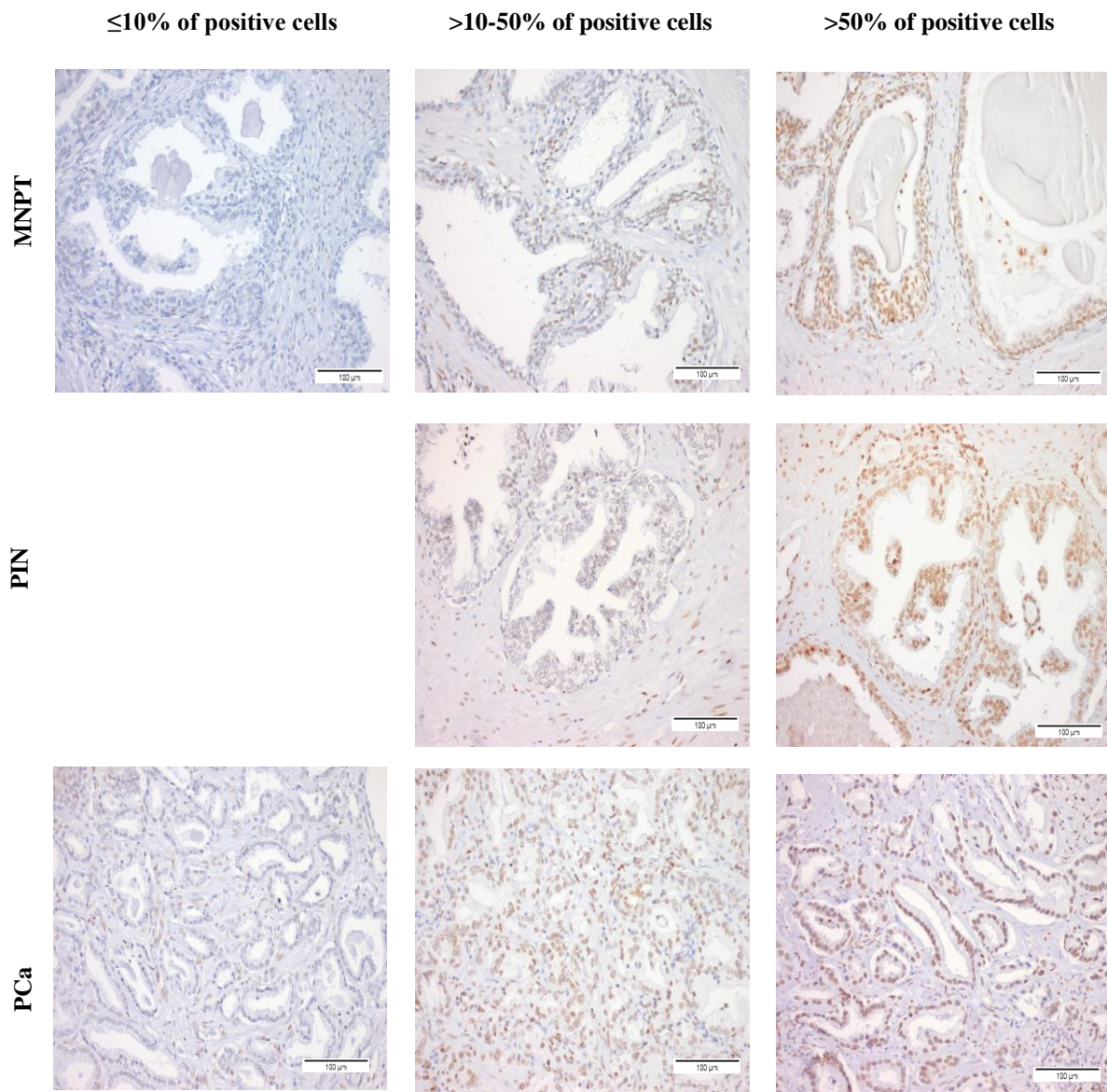
Figure 17. Illustrative images of MacroH2A1.1 immunostaining in MNPT, PIN and PCa samples.

Table 8. Immunohistochemistry of macroH2A1.1 protein levels in MNPT, PIN and PCa clinical samples.

Clinical Sample	Number of samples		
	≤10% of positive cells	>10-50% of positive cells	>50% of positive cells
MNPT (n=14)	1 (7%)	5 (35.7%)	8 (57.3%)
PIN (n=41)	-	2 (4.9%)	39 (95.1%)
PCa (n=188)	35 (18.6%)	39 (20.7%)	153 (60.7%)
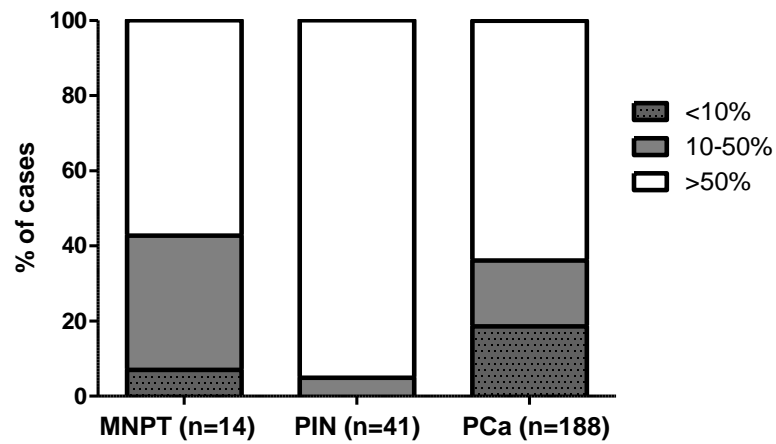
Figure 18. Distribution of macroH2A1.1 protein levels by percentage of positive cells in prostate tissues.

9. MACROH2A1 AND SPLICING REGULATORS EXPRESSION LEVELS IN PROSTATE CANCER CELL LINES

MacroH2A1 and splicing regulators expression levels were assessed, by RT-qPCR, in five PCa cell lines (22Rv1, LNCaP, VCaP, DU145 and PC-3), normalized for one benign prostate cell line (RWPE-1) (Figure 19).
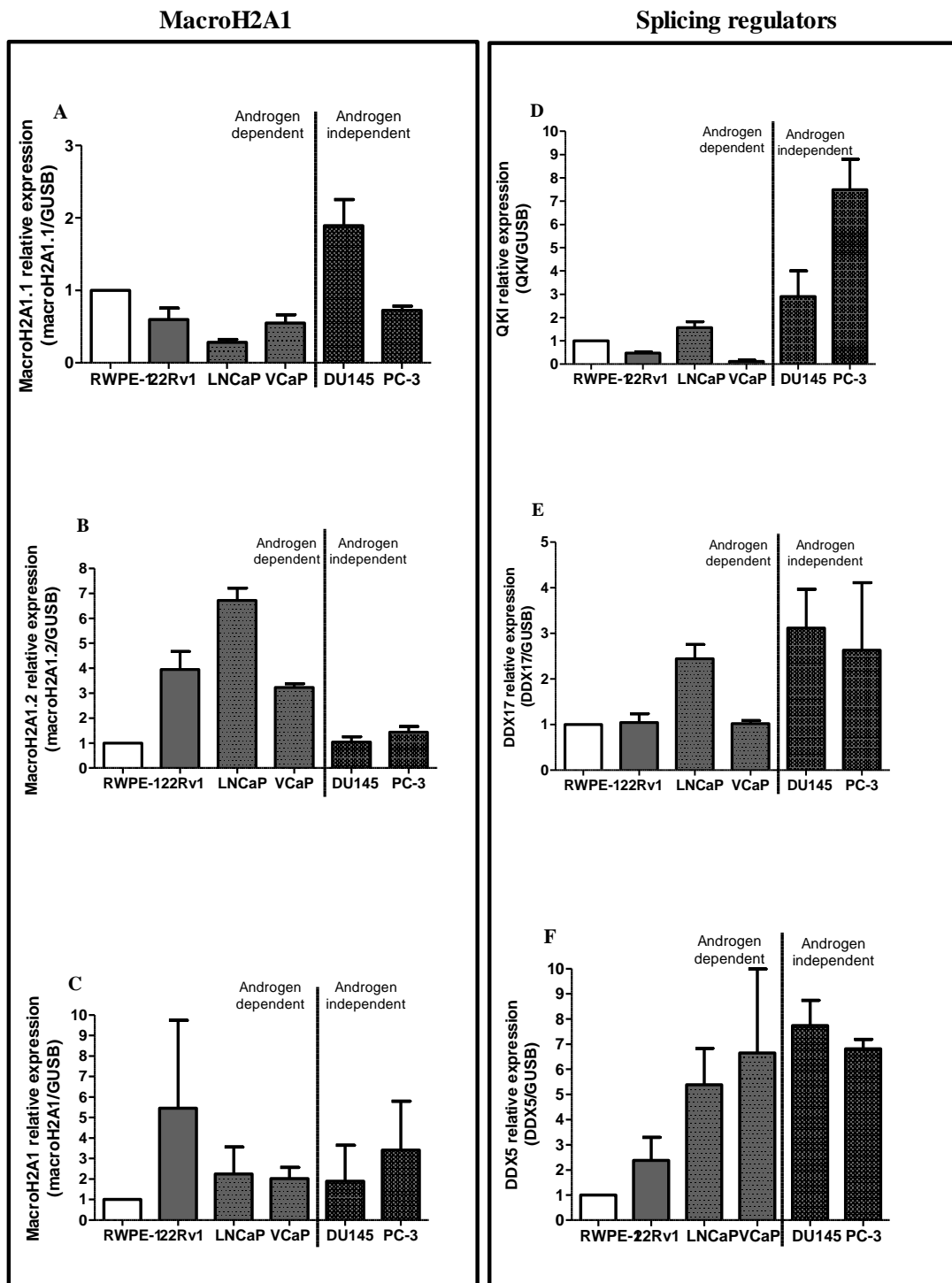
Figure 19. Expression levels of macroH2A1 isoforms (A and B), total (C) and splicing regulators (D, E and F) in prostate cell lines (normalized to RWPE-1).

52

Androgen independent PCa cell lines (DU145 and PC-3) globally displayed the highest macroH2A1.1 and splicing regulators expression levels. Conversely, androgen dependent PCa cell lines (22Rv1, LNCaP and VCaP) demonstrated the highest macroH2A1.2 expression. Considering macroH2A1 as a whole, a clear difference between androgen dependent and androgen independent cell lines was not apparent, although 22Rv1 showed the highest levels.

10. OVEREXPRESSION OF MACROH2A1 IN LNCaP CELL LINE

The cell line that expressed lower transcript levels of macroH2A1.1 was selected for transfection, for further *in vitro* studies of overexpression of this isoform. Effective stable overexpression were achieved in LNCaP and confirmed by RT-qPCR and at protein levels. While macroH2A1.1 is about 52x more expressed in transfected cells, macroH2A1.2 transcript or protein levels were not altered between NC and transfected cells (Figure 20).
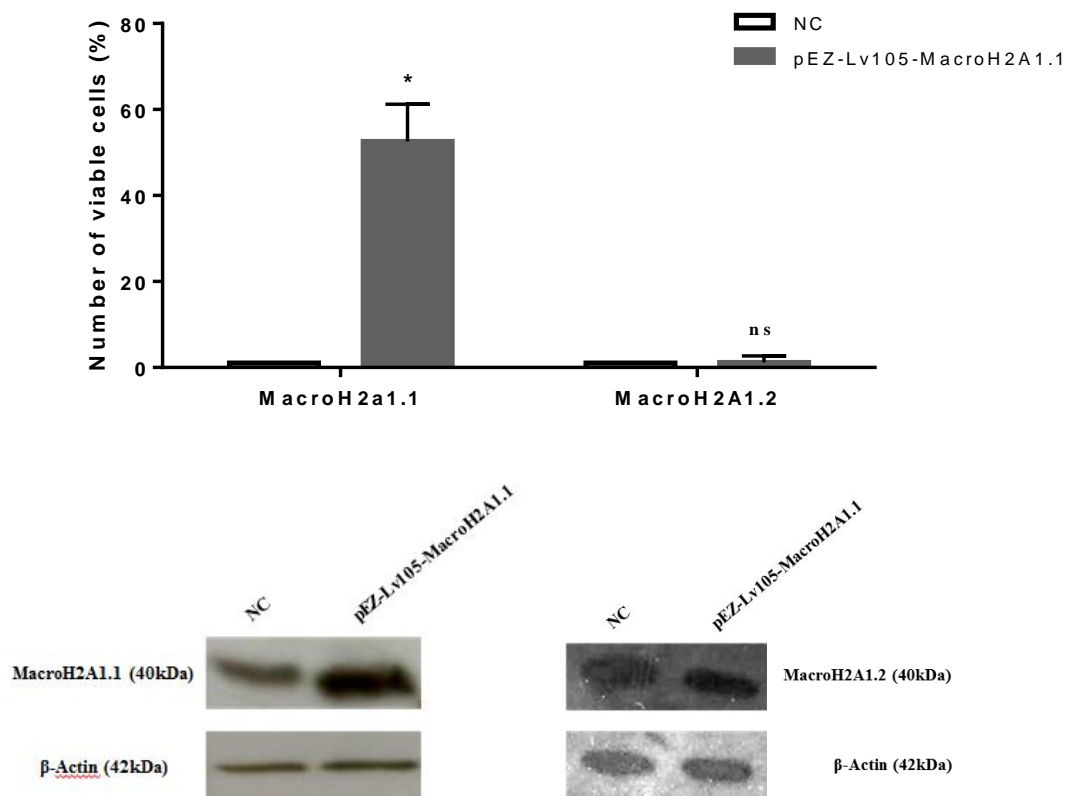
Figure 20. MacroH2A1.1 overexpression in LNCaP was confirmed at mRNA (upper panel), and at protein level (lower panel). MacroH2A1.2 transcript and protein levels were also assessed to confirm specific-variant transfection. *p<0.05, ns = not significant (Mann- Whitney U-test).

11. PRELIMINARY *IN VITRO* STUDIES: IMPACT OF MACROH2A1.1 OVEREXPRESSION IN CELL VIABILITY

Cell viability of LNCaP overexpressing macroH2A1.1 was significantly reduced at 72 hours (p= 0.0003). Results were compared with NC cells and are displayed in Figure 21.
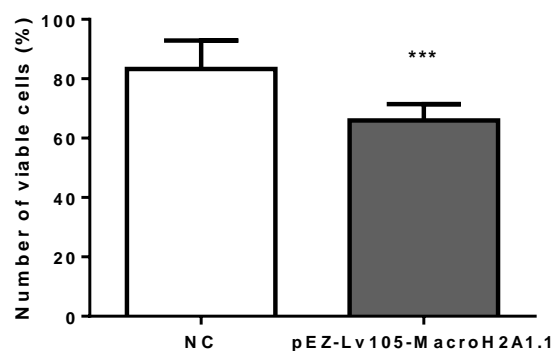
Figure 21. Impact of MacroH2A1.1 overexpression in cell viability of LNCaP at 72h. *p<0.05 (Mann-Whitney U-test).

V . DISCUSSION

DISCUSSION

PCa is the most common tumor in men and a leading cause of mortality and morbidity, worldwide. Both genetic and epigenetic disruption have been implicated in its initiation and progression. Unravel the mechanisms underlying tumor development are key to provide a deeper knowledge of PCa biology, that might be translated into diagnostic and prognostic, as well as provide novel therapeutic targets [20, 88].

Among epigenetic mechanisms, the shuffle of histones has been recently implicated in tumorigenesis [98]. This is apparent for the two macroH2A1 isoforms that have been the recent focus of several studies, attempting to unravel its role in cancer [126, 129]. MacroH2A1.1 is mostly considered a tumor suppressor, inhibiting stem-like properties, counteracting the functions of macroH2A1.2 [118, 123, 132]. Although its role has been previously tackled in breast and lung cancer, it has not been explored in prostate tumorigenesis, to the best of our knowledge. Thus, we aimed to determine the putative role of macroH2A1 isoforms in PCa and evaluate its biomarker performance.

Through the assessment of macroH2A1 isoforms transcript levels by qRT-PCR in prostate tissues, downregulation of macroH2A1.1 in PIN lesions and primary PCa, compared to normal prostate tissues was disclosed, when either $GUS\beta$ or macroH2A1 were used for normalization. This result is in line with previous observations on macroH2A1.1 expression in other primary cancers [103, 129]. Moreover, the intermediate expression levels depicted in PIN lesions is consistent with its putative PCa precursor role. However, macroH2A1.2 expression did not parallel that of macroH2A1.1, as only transcript levels in PIN were significantly lower than those of MNPT and PCa. Nevertheless, this result is in accordance with the lack of altered expression or slight upregulation of macroH2A1.2 in other tumor models [103, 128]. To further illuminate the biological variation of expression of each isoform, macroH2A1.1 transcript levels were normalized against macroH2A1.2. Although PIN displayed the lowest macroH2A1 levels compared to MNPT and PCa, this was mostly due to macroH2A1.2 downregulation. Compared to MNPT, macroH2A1.1 expression levels remained lower in PCa. Thus, our results suggest that sustained macroH2A1.1 downregulation is associated with the emergence of malignant phenotype in the prostate, whereas concomitant macroH2A1.2 decreased expression might be relevant

for the development of PIN lesions only. Notwithstanding, macroH2A1 downregulation seems to play an important role along prostate tumorigenesis, suggestive of a tumor suppressive effect.

Considering these findings, we hypothesized that differential expression of macroH2A1 isoforms might be related with altered splicing regulators, which play an important role in isoforms expression regulation. Thus, transcript levels of the better characterized macroH2A1 splicing regulators (*DDX5*, *DDX17* and *QKI*) were measured in the same tissues. Remarkably, QKI and DDX17 expressions levels paralleled those of macroH2A1.1, whereas DDX5 transcript levels were followed the same pattern of macroH2A1 and macroH2A1.2. When correlation analysis was carried out, a moderately positive correlation between QKI and macroH2A1.1 expression, as well as DDX5 and macroH2A1 expression was apparent, whereas DDX5 and macroH2A1.2 expression displayed a weak correlation, only. Thus, it is tempting to speculate whether variations in macroH2A1 isoform expression in PCa are due to altered expression of splicing regulators, although other factors might be involved. In support of this hypothesis, QKI has been considered a tumor suppressor in various cancers, frequently associated with macroH2A1.1 downregulation [104]. Concerning DDX5 and DDX17 expression, our results are somewhat unexpected as both are considered highly homologous oncogenic RNA-helicases [133]. Nevertheless, lower expression of both DDX5 and DDX17 were reported in non-invasive breast cancer, along with increased macroH2A1.1/macroH2A1.2 ratio [105]. Remarkably, this parallels our observations in PIN lesions, which might also be considered a pre-invasive form of PCa. Interestingly, DDX5, but not DDX17, affect key cellular pathways, including upregulation of AR in PCa and induction of epithelial-mesenchymal transition (EMT) [134, 135], a feature that is associated with tumor invasion.

Subsequently, we focused our attention on QKI and macroH2A1.1 expression in neoplastic lesions of the prostate. For that purpose, a subset of cases with matched PIN and PCa tissues was analyzed. Although lesions were found in the same gland, a direct causal link between them should not be construed, but it might help to elucidate the alterations of QKI and macroH2A1.1 expression along the carcinogenic process. In approximately two thirds of the cases, both QKI and macroH2A1.1 expression was lower in PCa samples compared to matched PIN, a finding that parallels the observed variations in the whole case series. Moreover, this result further supports a causal role of QKI downregulation in macroH2A1.1 decreased expression along prostate

tumorigenesis. Indeed, decreased QKI and macroH2A1.1 expression levels are clearly associated with the emergence of PCa, as supported by its good performance as potential PCa biomarker.

Some interesting associations between QKI and macroH2A1.1 expression and clinicopathological parameters were depicted. Specifically, PCa with higher Gleason score displayed median lower QKI and macroH2A1.1 levels, whereas macroH2A1.1 expression was higher in PCa patients with higher serum PSA levels. The association with the Gleason score seems logical as higher scores correspond to less differentiated and more aggressive PCa. The association with serum PSA levels, on the contrary, is almost counterintuitive. Nevertheless, it should be recalled that cells from less differentiated PCa produce less PSA, which might have a negative impact on global serum PSA levels notwithstanding heavy disease burden and corresponding poor outcome [136]. Overall, these findings suggest that lower QKI and macroH2A1.1 expression levels might be related with worse PCa-related survival, a hypothesis that requires further investigation.

To determine whether alterations in macroH2A1.1 transcript levels affected the respective protein levels, we assessed protein expression using IHC in FFPE sections from the same cases, when available. The variation in the proportion of cases with >50% of immunostained cells roughly followed the trend observed for the transcript, although no apparent statistical correlation between macroH2A1.1 transcript and protein expression was found. This might be influenced by the small number of MNPT analyzed, as well as from the compression of IHC results into three categories, to facilitate statistical analyses. Nevertheless, this result might also be related with the intrinsic biology of macroH2A1.1 as the protein displays a long half-life and variations in its levels might not be easily detectable [96].

In order to understand macroH2A1.1's biological role in PCa, the phenotypic effect of macroH2A1.1 overexpression was evaluated in a cell line with lowest expression levels (LNCaP). An attenuation of malignant phenotype by significantly decrease of cell viability was observed in transfected cells after 72h. MacroH2A1.1 is described as EMT and cancer growth suppressor by reducing PARP1 expression [104, 137, 138]. Further *in vitro* studies of macroH2A1.1 overexpression in PCa cell lines are essential to confirm the role of macroH2A1.1 in primary prostate tumorigenesis.

VI . CONCLUSIONS AND FUTURE PERSPECTIVES

CONCLUSIONS

To the best of our knowledge, this study is the first to report variations in expression of macroH2A1 and its isoforms in prostate tissues, encompassing morphologically normal and neoplastic (pre-invasive and invasive) lesions. Globally, macroH2A1.1 expression is gradually decreased along prostatic tumorigenesis, whereas macroH2A1 and macroH2A1.2 are downregulated in PIN. The variations in macroH2A1 are mostly affected by macroH2A1.2 isoform. These alterations are associated with altered expression of splicing regulators, specifically QKI and macroH2A1.1, as well as DDX5 and macroH2A1 and macroH2A1.2.

Interestingly, less differentiated and more aggressive PCa displays lower QKI and macroH2A1.1 expression, as expected for putative tumor suppressors. Although no significant correlation was observed between macroH2A1.1 transcript and protein expression, the percentage of immunostained cells globally reflected the variations observed in transcript levels.

In vitro, stable macroH2A1.1 overexpression attenuates the malignant phenotype, by decreasing cell viability, probably due to increase of cell differentiation.

FUTURE PERSPECTIVES

To validate macroH2A1.1 role in PCa cell lines is important to:
- ❖ Study the impact of macroH2A1.1 overexpression in prostate cells, by cellular apoptosis and differentiation assays
- ❖ Study the impact of knockout of macroH2A1.1 in prostate cell lines (DU145) by *in vitro* studies (cell viability, apoptosis, invasive, migration and differentiation)

In order to confirm that QKI specific increases macroH2A1.1 expression is vital to:
- ❖ Induction of QKI expression in PCa cell lines to further assess of macroH2A1.1 levels.

VII. REFERENCES

REFERENCES

1.   Dunn, M.W. and M.W. Kazer, *Prostate cancer overview.* Semin Oncol Nurs, 2011. **27**(4): p. 241-50.
2.   Martini, F., Timmons, M.J., Tallitsch, R.B., *Human anatomy.* Eighth edition. ed. 2015, Boston: Pearson. xxxii, 862 pages.
3.   Drake, R.L., Vogl, W., Mitchell, A. W. M., Gray, H., *Gray's anatomy for students.* Third edition. ed. 2015, Philadelphia, PA: Churchill Livingstone/Elsevier. xxv, 1161 pages.
4.   Rosai, J., Ackerman, L. V., *Rosai and Ackerman's surgical pathology.* . Elsevier Health Sciences, ed. 10. 2011.
5.   Longo, D.L., et al., *Harrison's Principles of Internal Medicine.* 18th ed. Vol. 1. 2011. 6.
6.   McNeal, J.E., *Anatomy of the Prostate: An Historical Survey of Divergent Views.* The Prostate, 1980. **1**: p. 11.
7.   Selman, S.H., *The McNeal prostate: a review.* Urology, 2011. **78**(6): p. 1224-8.
8.   Hammerich, K.H., G.E. Ayala, and T.M. Wheeler, *Anatomy of the prostate gland and surgical pathology of prostate cancer.* 2008: Cambridge University Press.
9.   McNeal, J.E., *Normal histology of the prostate.* Am J Surg Pathol, 1988. **12**(8): p. 619-33.
10.   McNeal, J.E., *The zonal anatomy of the prostate.* The Prostate, 1981. **2**: p. 35-49.
11.   Alcaraz, A., et al., *Is there evidence of a relationship between benign prostatic hyperplasia and prostate cancer? Findings of a literature review.* Eur Urol, 2009. **55**(4): p. 864-73.
12.   De Marzo, A.M., et al., *Inflammation in prostate carcinogenesis.* Nat Rev Cancer, 2007. **7**(4): p. 256-69.
13.   Wadhera, P., *An introduction to acinar pressures in BPH and prostate cancer.* Nat Rev Urol, 2013. **10**(6): p. 358-66.
14.   Haggman, M.J.e.a., *The Relationship Between Prostatic Intraepithelial Neoplasia and Prostate Cancer: Critical Issues.* The Journal of Urology, 1997. **158**(1): p. 10.
15.   Bostwick, D. and M. Brawer, *Prostatic lntra-Epithelial Neoplasia and Early Invasion in Prostate Cancer.* Cancer, 1987. **59**(4): p. 7.
16.   Bostwick, D.G. and L. Cheng, *Precursors of prostate cancer.* Histopathology, 2012. **60**(1): p. 4-27.
17.   Lee, C.H., O. Akin-Olugbade, and A. Kirschenbaum, *Overview of prostate anatomy, histology, and pathology.* Endocrinol Metab Clin North Am, 2011. **40**(3): p. 565-75, viii-ix.
18.   Brawer, M.K., *Prostatic Intraepithelial Neoplasia: An Overview.* Reviews in Urology, 2005. **7**(3): p. 8.
19.   De Marzo, A.M.e.a., *Human prostate cancer precursors and pathobiology.* Urology, 1999. **62**(5): p. 8.
20.   Shen, M.M. and C. Abate-Shen, *Molecular genetics of prostate cancer: new prospects for old challenges.* Genes Dev, 2010. **24**(18): p. 1967-2000.

21. Chrisofos, M., et al., *Precursor lesions of prostate cancer.* Crit Rev Clin Lab Sci, 2007. **44**(3): p. 243-70.
22. Sakr, W.A., et al., *The frequency of carcinoma and intraepithelial neoplasia of the prostate in young male patients.* J Urol, 1993. **150**(2 Pt 1): p. 379-85.
23. Bostwick, D.G., et al., *Architectural patterns of high-grade prostatic intraepithelial neoplasia.* Hum Pathol, 1993. **24**(3): p. 298-310.
24. Haggman, M.J., et al., *Allelic loss of 8p sequences in prostatic intraepithelial neoplasia and carcinoma.* Urology, 1997. **50**(4): p. 643-7.
25. Boyd, L.K., X. Mao, and Y.J. Lu, *The complexity of prostate cancer: genomic alterations and heterogeneity.* Nat Rev Urol, 2012. **9**(11): p. 652-64.
26. Witte, J.S., *Prostate cancer genomics: towards a new understanding.* Nat Rev Genetics, 2009. **10**: p. 6.
27. Ferlay J. ; Soerjomataram I. ; Ervik M. ; Dikshit R. ; Eser S. ; Mathers C. , R.M., Parkin D.M. , Forman D. , Bray, F. *Cancer Incidence and Mortality Worldwide: IARC CancerBase No. 11 [Internet].* Lyon, France: International Agency for Research on Cancer 2012 [cited 2015 July]; Available from: http://globocan.iarc.fr.
28. Siegel, R.L., Miller, K.D., Jemal, A. , *Cancer Statistics, 2015.* CA CANCER J CLIN, 2015(65): p. 25.
29. Coldman, A.J., Phillips, N., Pickles, T.A., *Trends in prostate cancer incidence and mortality: an analysis of mortality change by screening intensity.* CMAJ, 2003. **168**(1): p. 5.
30. Center M.M. ; Jemal A., L.-T.J., Ward E., Ferlay J., Brawley O., Bray F., *International variation in Prostate Cancer Incidence and Mortality Rates.* European Urology, 2012. **61**: p. 14.
31. Bostwick, D.G., et al., *Human prostate cancer risk factors.* Cancer, 2004. **101**(10 Suppl): p. 2371-490.
32. Crawford, E.D., *Understanding the epidemiology, natural history, and key pathways involved in prostate cancer.* Urology, 2009. **73**(5 Suppl): p. S4-10.
33. Patel, A.R. and E.A. Klein, *Risk factors for prostate cancer.* Nat Clin Pract Urol, 2009. **6**(2): p. 87-95.
34. Gann, P.H., *Risk Factors for Prostate Cancer.* Reviews in Urology, 2002. **4**(5 Suppl): p. 8.
35. Schaid, D.J., *The complex genetic epidemiology of prostate cancer.* Hum Mol Genet, 2004. **13 Spec No 1**: p. R103-21.
36. Leitzmann, M.F. and S. Rohrmann, *Risk factors for the onset of prostatic cancer: age, location, and behavioral correlates.* Clin Epidemiol, 2012. **4**: p. 1-11.
37. Kasper, J.S. and E. Giovannucci, *A meta-analysis of diabetes mellitus and the risk of prostate cancer.* Cancer Epidemiol Biomarkers Prev, 2006. **15**(11): p. 2056-62.
38. Huncharek, M., et al., *Smoking as a risk factor for prostate cancer: a meta-analysis of 24 prospective cohort studies.* Am J Public Health, 2010. **100**(4): p. 693-701.
39. Gustavo f. Carvalhal, d.s.S., douglas e. Mager, christian ramos and william j. Catalona, *Digital rectal examination for detecting prostate cancer at prostate specific antigen levels of 4 ng./ml. Or less.* 1999.
40. Smith, J.A., Jr., et al., *Transrectal ultrasound versus digital rectal examination for the staging of carcinoma of the prostate: results of a prospective, multi-institutional trial.* J Urol, 1997. **157**(3): p. 902-6.

41.   Schiebler, M.L., et al., *MR imaging in adenocarcinoma of the prostate: interobserver variation and efficacy for determining stage C disease.* AJR Am J Roentgenol, 1992. **158**(3): p. 559-62; discussion 563-4.

42.   Hjertholm, P., et al., *Variation in general practice prostate-specific antigen testing and prostate cancer outcomes: an ecological study.* Int J Cancer, 2015. **136**(2): p. 435-42.

43.   Grubb, R.L., 3rd, et al., *Serum prostate-specific antigen hemodilution among obese men undergoing screening in the Prostate, Lung, Colorectal, and Ovarian Cancer Screening Trial.* Cancer Epidemiol Biomarkers Prev, 2009. **18**(3): p. 748-51.

44.   Açıkgöz Ş., C.M., Doğan S.M., Mungan G., Mustafa Aydın M., Kelek S., Sümbüloğlu V., *Prostate specific antigen levels after acute myocardial infarction.* 2011. **58**(4): p. 5.

45.   Wallner, L.P., et al., *The effects of type 2 diabetes and hypertension on changes in serum prostate specific antigen levels: results from the Olmsted County study.* Urology, 2011. **77**(1): p. 137-41.

46.   Chiam, K., C. Ricciardelli, and T. Bianco-Miotto, *Epigenetic biomarkers in prostate cancer: Current and future uses.* Cancer Lett, 2014. **342**(2): p. 248-56.

47.   DeVita, V.T., T.S. Lawrence, and S.A. Rosenberg, *Cancer: Principles and Practice of Oncology-Advances in Oncology.* 2010: Lippincott Williams & Wilkins.

48.   Marberger, M., *Prostate Cancer 2008: Challenges in Diagnosis and Management.* European Urology Supplements, 2009. **8**(3): p. 89-96.

49.   Costa-Pinheiro P, P.H., Henrique R, Jerónimo C, *Biomarkers and personalized risk stratification for patients with clinically localized prostate cancer.* Expert Rev. Anticancer Ther., 2014. **14**(11): p. 10.

50.   Arora, R., et al., *Heterogeneity of Gleason grade in multifocal adenocarcinoma of the prostate.* Cancer, 2004. **100**(11): p. 2362-6.

51.   Gleason, D.F., *Histologic Grading of Prostate Cancer: A Perspective.* Hum Pathol, 1992. **23**(3): p. 7.

52.   Harnden, P., et al., *Should the Gleason grading system for prostate cancer be modified to account for high-grade tertiary components? A systematic review and meta-analysis.* The Lancet Oncology, 2007. **8**(5): p. 411-419.

53.   Edge, S.B. and C.C. Compton, *The American Joint Committee on Cancer: the 7th edition of the AJCC cancer staging manual and the future of TNM.* Ann Surg Oncol, 2010. **17**(6): p. 1471-4.

54.   Edge S, B.D., Compton C. , *AJCC Cancer Staging Manual.* Springer, New York. Vol. 7. 2010.

55.   Bubendorf, L., et al., *Metastatic patterns of prostate cancer: An autopsy study of 1,589 patients.* Human Pathology, 2000. **31**(5): p. 578-583.

56.   Trewartha, D. and K. Carter, *Advances in prostate cancer treatment.* Nat Rev Drug Discov, 2013. **12**(11): p. 823-4.

57.   Jani, A.B. and S. Hellman, *Early prostate cancer: clinical decision-making.* The Lancet, 2003. **361**(9362): p. 1045-1053.

58.   Heidenreich, A., et al., *EAU guidelines on prostate cancer. Part 1: screening, diagnosis, and treatment of clinically localised disease.* Eur Urol, 2011. **59**(1): p. 61-71.

59.   Zincke, H., et al., *Radical prostatectomy for clinically localized prostate cancer: long-term results of 1,143 patients from a single institution.* J Clin Oncol, 1994. **12**(11): p. 2254-63.

60.	Forman, J.D., et al., *Carcinoma of the prostate in the elderly: the therapeutic ratio of definitive radiotherapy.* J Urol, 1986. **136**(6): p. 1238-41.

61.	Ragde, H., et al., *Interstitial iodine-125 radiation without adjuvant therapy in the treatment of clinically localized prostate carcinoma.* Cancer, 1997. **80**(3): p. 442-53.

62.	Kohli, M. and D.J. Tindall, *New developments in the medical management of prostate cancer.* Mayo Clin Proc, 2010. **85**(1): p. 77-86.

63.	He, C. and P. Cole, *Introduction: epigenetics.* Chem Rev, 2015. **115**(6): p. 2223-4.

64.	Biterge, B. and R. Schneider, *Histone variants: key players of chromatin.* Cell Tissue Res, 2014. **356**(3): p. 457-66.

65.	Waddington, C.H., *The epigenotype. 1942.* Int J Epidemiol, 2012. **41**(1): p. 10-3.

66.	Goldberg, A.D., C.D. Allis, and E. Bernstein, *Epigenetics: a landscape takes shape.* Cell, 2007. **128**(4): p. 635-8.

67.	Sandoval, J. and M. Esteller, *Cancer epigenomics: beyond genomics.* Curr Opin Genet Dev, 2012. **22**(1): p. 50-5.

68.	Suzuki, M.M. and A. Bird, *DNA methylation landscapes: provocative insights from epigenomics.* Nat Rev Genet, 2008. **9**(6): p. 465-76.

69.	Vaissiere, T., C. Sawan, and Z. Herceg, *Epigenetic interplay between histone modifications and DNA methylation in gene silencing.* Mutat Res, 2008. **659**(1-2): p. 40-8.

70.	Bird, A., *DNA methylation patterns and epigenetic memory.* Genes Dev, 2002. **16**(1): p. 6-21.

71.	Portela, A. and M. Esteller, *Epigenetic modifications and human disease.* Nat Biotechnol, 2010. **28**(10): p. 1057-68.

72.	Jones, P.A. and S.B. Baylin, *The fundamental role of epigenetic events in cancer.* Nat Rev Genet, 2002. **3**(6): p. 415-28.

73.	Costa, F.F., *Non-coding RNAs and new opportunities for the private sector.* Drug Discov Today, 2009. **14**(9-10): p. 446-52.

74.	Garzon, R., G.A. Calin, and C.M. Croce, *MicroRNAs in Cancer.* Annu Rev Med, 2009. **60**: p. 167-79.

75.	Vasudevan, S., Y. Tong, and J.A. Steitz, *Switching from repression to activation: microRNAs can up-regulate translation.* Science, 2007. **318**(5858): p. 1931-4.

76.	Guil, S. and M. Esteller, *DNA methylomes, histone codes and miRNAs: tying it all together.* Int J Biochem Cell Biol, 2009. **41**(1): p. 87-95.

77.	Iorio, M.V. and C.M. Croce, *MicroRNAs in cancer: small molecules with a huge impact.* J Clin Oncol, 2009. **27**(34): p. 5848-56.

78.	Annunziato, A., *DNA Packaging: Nucleosomes and Chromatin. .* Nature Education, 2008. **1**(1): p. 26.

79.	Kouzarides, T., *Chromatin Modifications and Their Function.* Cell, 2007. **128**: p. 13.

80.	Marc, F., et al., *Archaeal histone tetramerization determines DNA affinity and the direction of DNA supercoiling.* J Biol Chem, 2002. **277**(34): p. 30879-86.

81.	Kamakaka, R.T. and S. Biggins, *Histone variants: deviants?* GENES & DEVELOPMENT, 2005. **19**: p. 295–310.

82.	Mellor, J., *The Dynamics of Chromatin Remodeling at Promoters.* Mol Cell, 2005. **19**: p. 11.

83.	Cosgrove, M.S., J.D. Boeke, and C. Wolberger, *Regulated nucleosome mobility and the histone code.* Nat Struct Mol Biol, 2004. **11**(11): p. 1037-43.

84.   Allis, T.J.a.C.D., *Translating the Histone Code.* Science, 2001. **293**: p. 1074-80.

85.   Arrowsmith, C.H., et al., *Epigenetic protein families: a new frontier for drug discovery.* Nat Rev Drug Discov, 2012. **11**(5): p. 384-400.

86.   Gibney, E.R. and C.M. Nolan, *Epigenetics and gene expression.* Heredity (Edinb), 2010. **105**(1): p. 4-13.

87.   Lennartsson, A. and K. Ekwall, *Histone modification patterns and epigenetic codes.* Biochim Biophys Acta, 2009. **1790**(9): p. 863-8.

88.   Sharma, S., T.K. Kelly, and P.A. Jones, *Epigenetics in cancer.* Carcinogenesis, 2010. **31**(1): p. 27-36.

89.   Vardabasso, C., Hasson, D., Ratnakumar, K., Chung, C. Y., Duarte, L. F., Bernstein, E., *Histone variants: emerging players in cancer biology.* Cell Mol Life Sci, 2014. **71**(3): p. 379-404.

90.   Pehrson, J.R. and R.N. Fuji, *Evolutionary conservation of histone macroH2A subtypes and domains.* Nucleic Acids Res, 1998. **26**(12): p. 2837–2842.

91.   WU, R.S., D. NISHIOKA, and W.M. BONNER, *Differential Conservation of Histone 2A Variants between Mammals and Sea Urchins.* The Journal of Cell Biology, 1982. **93**: p. 426-431.

92.   Marzluff, W.F., E.J. Wagner, and R.J. Duronio, *Metabolism and regulation of canonical histone mRNAs: life without a poly(A) tail.* Nat Rev Genet, 2008. **9**(11): p. 843-54.

93.   Hake, S.B., A. Xiao, and C.D. Allis, *Linking the epigenetic 'language' of covalent histone modifications to cancer.* Br J Cancer, 2004. **90**(4): p. 761-9.

94.   Volle, C., Dalal, Y., *Histone variants: the tricksters of the chromatin world.* Curr Opin Genet Dev, 2014. **25**: p. 8-14,138.

95.   Bonisch, C. and S.B. Hake, *Histone H2A variants in nucleosomes and chromatin: more or less stable?* Nucleic Acids Res, 2012. **40**(21): p. 10719-41.

96.   Melters, D.P., et al., *Chromatin Dynamics in Vivo: A Game of Musical Chairs.* Genes (Basel), 2015. **6**(3): p. 751-76.

97.   Maze, I., et al., *Every amino acid matters: essential contributions of histone variants to mammalian development and disease.* Nat Rev Genet, 2014. **15**(4): p. 259-71.

98.   L., M.F., et al., *Expression and functionality of histone H2A variants in cancer.* Oncotarget, 2014: p. 16.

99.   Chakravarthy, S., et al., *Structural characterization of histone H2A variants.* Cold Spring Harb Symp Quant Biol, 2004. **69**: p. 227-34.

100.   Chakravarthy, S., A. Patel, and G.D. Bowman, *The basic linker of macroH2A stabilizes DNA at the entry/exit site of the nucleosome.* Nucleic Acids Res, 2012. **40**(17): p. 8285-95.

101.   Chakravarthy, S. and K. Luger, *The histone variant macro-H2A preferentially forms "hybrid nucleosomes".* J Biol Chem, 2006. **281**(35): p. 25522-31.

102.   Muthurajan, U.M., et al., *The linker region of macroH2A promotes self-association of nucleosomal arrays.* J Biol Chem, 2011. **286**(27): p. 23852-64.

103.   Sporn, J.C. and B. Jung, *Differential regulation and predictive potential of MacroH2A1 isoforms in colon cancer.* Am J Pathol, 2012. **180**(6): p. 2516-26.

104.   Novikov, L., et al., *QKI-mediated alternative splicing of the histone variant MacroH2A1 regulates cancer cell proliferation.* Mol Cell Biol, 2011. **31**(20): p. 4244-55.

105.   Dardenne, E., et al., *Splicing switch of an epigenetic regulator by RNA helicases promotes tumor-cell invasiveness.* Nat Struct Mol Biol, 2012. **19**(11): p. 1139-46.

106. Yang, Y.A., Kim, J., Yu, J., *Influence of oncogenic transcription factors on chromatin conformation and implications in prostate cancer.* Appl Clin Genet, 2014. **7**: p. 81-91.

107. Ratnakumar, K., et al., *ATRX-mediated chromatin association of histone variant macroH2A1 regulates alpha-globin expression.* Genes Dev, 2012. **26**(5): p. 433-8.

108. Mehrotra, P.V., et al., *DNA repair factor APLF is a histone chaperone.* Mol Cell, 2011. **41**(1): p. 46-55.

109. Lavigne, M.D., Vatsellas, G., Polyzos, A., Mantouvalou, E., Sianidis, G., Maraziotis, I., Agelopoulos, M., Thanos, D., *Composite macroH2A/NRF-1 Nucleosomes Suppress Noise and Generate Robustness in Gene Expression.* Cell Rep, 2015.

110. Gamble, M.J. and W.L. Kraus, *Multiple facets of the unique histone variant macroH2A: From genomics to cell biology.* Cell Cycle, 2014. **9**(13): p. 2568-2574.

111. Pehrson, J.R., Changolkar, L. N.; Costanzi, C.; Leu, N. A., *Mice Without MacroH2A Histone Variants.* Mol Cell Biol, 2014.

112. Costanzi, C. and J.R. Pehrson, *Histone macroH2A1 is concentrated in the inactive X chromosome of female mammals.* Nature, 1998. **393**(11): p. 599-601.

113. Jeon, Y., K. Sarma, and J.T. Lee, *New and Xisting regulatory mechanisms of X chromosome inactivation.* Curr Opin Genet Dev, 2012. **22**(2): p. 62-71.

114. Tanasijevic, B. and T.P. Rasmussen, *X chromosome inactivation and differentiation occur readily in ES cells doubly-deficient for macroH2A1 and macroH2A2.* PLoS One, 2011. **6**(6): p. e21512.

115. Bernstein, E., et al., *A phosphorylated subpopulation of the histone variant macroH2A1 is excluded from the inactive X chromosome and enriched during mitosis.* Proc Natl Acad Sci U S A, 2008. **105**(5): p. 1533-8.

116. Nora, E.P. and E. Heard, *Chromatin structure and nuclear organization dynamics during X-chromosome inactivation.* Cold Spring Harb Symp Quant Biol, 2010. **75**: p. 333-44.

117. Zhang, R., et al., *Formation of MacroH2A-containing senescence-associated heterochromatin foci and senescence driven by ASF1a and HIRA.* Dev Cell, 2005. **8**(1): p. 19-30.

118. Gaspar-Maia, A., et al., *MacroH2A histone variants act as a barrier upon reprogramming towards pluripotency.* Nat Commun, 2013. **4**: p. 1565.

119. Buschbeck, M., et al., *The histone variant macroH2A is an epigenetic regulator of key developmental genes.* Nat Struct Mol Biol, 2009. **16**(10): p. 1074-9.

120. Gamble, M.J., et al., *The histone variant macroH2A1 marks repressed autosomal chromatin, but protects a subset of its target genes from silencing.* Genes Dev, 2010. **24**(1): p. 21-32.

121. Xu, C., et al., *The histone variant macroH2A1.1 is recruited to DSBs through a mechanism involving PARP1.* FEBS Lett, 2012. **586**(21): p. 3920-5.

122. Pasque, V., et al., *Histone variant macroH2A marks embryonic differentiation in vivo and acts as an epigenetic barrier to induced pluripotency.* J Cell Sci, 2012. **125**(Pt 24): p. 6094-104.

123. Creppe, C., et al., *MacroH2A1 regulates the balance between self-renewal and differentiation commitment in embryonic and adult stem cells.* Mol Cell Biol, 2012. **32**(8): p. 1442-52.

124. Chen, H., et al., *MacroH2A1.1 and PARP-1 cooperate to regulate transcription by promoting CBP-mediated H2B acetylation.* Nat Struct Mol Biol, 2014. **21**(11): p. 981-9.

125. Kapoor, A., et al., *The histone variant macroH2A suppresses melanoma progression through regulation of CDK8.* Nature, 2010. **468**(7327): p. 1105-9.

126. Li, X., et al., *The atypical histone macroH2A1.2 interacts with HER-2 protein in cancer cells.* J Biol Chem, 2012. **287**(27): p. 23171-83.

127. Lavigne, A.C., et al., *Increased macroH2A1.1 expression correlates with poor survival of triple-negative breast cancer patients.* PLoS One, 2014. **9**(6): p. e98930.

128. Sporn, J.C., et al., *Histone macroH2A isoforms predict the risk of lung cancer recurrence.* Oncogene, 2009. **28**(38): p. 3423-8.

129. Park, S.J., Shim, J. W., Park, H. S., Eum, D. Y., Park, M. T., Mi Yi, J., Choi, S. H., Kim, S. D., Son, T. G., Lu, W., Kim, N. D., Yang, K., Heo, K., *MacroH2A1 downregulation enhances the stem-like properties of bladder cancer cells by transactivation of Lin28B.* Oncogene, 2015.

130. Patricio, P., et al., *Deregulation of PAX2 expression in renal cell tumours: mechanisms and potential use in differential diagnosis.* J Cell Mol Med, 2013. **17**(8): p. 1048-58.

131. Kelly H. Zou, P., M. Kemal Tuncali, and M. Stuart G. Silverman, *Correlation and Simple Linear Regression.* Radiology, 2003: p. 617-622.

132. Pasque, V., et al., *Histone variant macroH2A confers resistance to nuclear reprogramming.* EMBO J, 2011. **30**(12): p. 2373-87.

133. Shin, S., et al., *Involvement of RNA helicases p68 and p72 in colon cancer.* Cancer Res, 2007. **67**(16): p. 7572-8.

134. Clark, E.L., et al., *The RNA helicase p68 is a novel androgen receptor coactivator involved in splicing and is overexpressed in prostate cancer.* Cancer Res, 2008. **68**(19): p. 7938-46.

135. Fuller-Pace, F.V., *DEAD box RNA helicase functions in cancer.* RNA Biol, 2013. **10**(1): p. 121-32.

136. McGuire, B.B., et al., *Outcomes in patients with Gleason score 8-10 prostate cancer: relation to preoperative PSA level.* BJU Int, 2012. **109**(12): p. 1764-9.

137. Schiewer, M.J., et al., *Dual roles of PARP-1 promote cancer growth and progression.* Cancer Discov, 2012. **2**(12): p. 1134-49.

138. Pu, H., Horbinski, C., Hensley, P. J., Matuszak, E. A., Atkinson, T., Kyprianou, N., *PARP-1 regulates epithelial-mesenchymal transition (EMT) in prostate tumorigenesis.* Carcinogenesis, 2014. **35**(11): p. 2592-601.