# IDENTIFICATION OF NEW URINE BIOMARKERS FOR THE DETECTION OF PROSTATE CANCER


PhD thesis presented by

Marina Rigau Resina


To obtain the degree of

PhD for the UAB

PhD thesis done in the Research Unit in Biomedicine and Translational and Pediatrics Oncology Department in the Vall Hebron Research Institute, under the supervision of Drs.

Jaume Reventós i Puigjaner, Andreas Doll and Joan Morote Robles

Universitat Autònoma de Barcelona, Faculty of Bioscience, Department of Biochemistry, Molecular Biology in the Faculty of Bioscience in the UAB, under the supervision of

Dr. Jaume Farrés

Universitat Autònoma de Barcelona, June 2011

Dr. Jaume Reventós  Dr. Andreas Doll  Dr. Joan Morote

Marina Rigau Resina

This thesis has been supported by the grant “Beca Eduardo Gallego para Jóvenes Licenciados” (2008) from “Fundación Francisco Cobos” and the grant “Talent Empresa” from “Generalitat de Catalunya-AGAUR” (2009-TEM 00066; TransBioMed).

Moreover this work was supported in part for the European Molecular and Biology Organization (EMBO) with a “Short Term Fellowship (ASTF 140-2010)”, a grant for a short stage fellowship from “Ministerio de Ciencia e Inovación (RTICC) and a grant for “AM2c Mobility of Researchers” from “Fundation National de la Research” (FNR) in Luxembourg, (FNR/10/AM2c/2g).

Finally the grant support for the projects presented; “Red de Genómica del Cáncer y Genotipado de tumors “ (Grant number: C03/10), “Instituto de Salud Carlos III, Ministerio de Ciéncia e Inovación” (RTICC RD06/0020/0058), “Fundación para la Investigación en Urología - FIU” (2007-FIU), “Recerca i Societat de la Informació de la Generalitat de Catalunya; Grant number SGR00487” and finally, a grant from AGAUR; Generalitat de Catalunya (Valor 00220).

En realitat tot i que sé, que aquest apartat no és un requisit indispensable, a diferència de l'infinit nombre d'obstacles de superar per poder dipositar una tesi, és poster la part més important com a científic i com a persona. Crec que és necessari plasmar d'alguna manera que el treball que aquí es presenta no hauria sigut possible dur-lo a terme si no hi haguessin hagut un munt de persones al darrera.

En primer lloc vull mostrar la meva gratitut als meus directors de tesi, el Dr. Jaume Reventós, el Dr. Andreas Doll i el Dr. Joan Morote, per brindar-me l'oportunitat de realiztar aquest treball, per confiar en les meves possibilitats i per voler introduir-me en el món de la recerca i en un món "paral.lel", una mica complicat a l'inici (TransBioMed), però que per a mi representa un repte important.

Alhora vull agraïr el support dels clinics del servei d'urologia l'Hospital de la Vall d'Hebrón, així com també agraïr als patòlegs, als auxiliars de quiròfan i a les enfermes de les consultes externes. Tots plegats han fet possible el contacte amb els pacients. De la mateixa manera, vull mostrar el meu agraïment més sincer a tots els pacients que han volgut contribuir amb una "granet de sorra" a que tot aquest treball s'hagi pogut tirar endevant.

Agraeixo d'una forma molt carinyosa la càlida acollida al Laboratori de Proteòmica, sota la direcció del Dr. Francesc Canals, que amb l'ajut de les "floretes de proteòmica" i en especial a la Núria Colomé, han fet que em senti una "més" del seu laboratori, així com la seva dedicació cap a al meu projecte.

Agraeixo també al departament d'estadística de l'Institut, a l'Israel i a l'Àlex Sànchez, per el seu ajut incondicional en tots els anàlisis estadístics, i sobretot per la seva llarga paciència. Així com també, agraeixo a la Paqui, la Mª Ángeles i al Ricardo, per el seu suport tècnic en gran part dels experiments aquí presentats.

De la mateixa manera agraeïxo l'acollida, durant la meva estada al "Clinical Proteomics Laboratory" de Luxembourg, per part del Dr. Bruno Domon i la Dr. Elodie Duriez, així com tots els altres integrants del grup. Aquesta estada ha estat clau per a la elaboració de gran part dels resultats que es presenten en aquesta tesi i alhora ha sigut una experiència plena de nous coneixements, així com una etapa personal important.

Agraeixo al Dr. Lennart Martens i al Dr. Sven Degroeve del grup "Computational Omics and Systems Biology" del "Department of Medical Protein Research" a la Universitat de Ghent (Bèlgica), els quals estan col.laborant en l'anàlisi estadístic de part dels resultats que aquí es presenten.

Finalment, vull agraïr sense noms perquè sé que aquells a qui he d'agraïr el seu support, ajut, constància, dedicació, somriures i plors, "sabis consells", xocolata…. i sobretot l'amistat que s'ha creat al voltant d'aquesta tesi (o bé al marge), no cal que els ennumeri, tots i totes estareu sempre sota aquesta frase.

I, tot i que aquestes persones ja venen ennumerades anteriorment, no puc deixar d'agraïr la creativitat del Borja, pel diseny de la portada, a la meva germana per la idea del dibuix (fet per un amic del meu avi, al meu avi), i bueno…. al Miquel pel seu suport com a parella i ajut durant tot aquest temps, i per acabar, no podria deixar d'anomenar als meus pares, per ser sempre allà.

Gràcies a tots per acompanyar-me en aquesta etapa, que ha sigut bonica, en alguns moments dura, però sobretot la recordaré com a "dolça" per tots els moments que he viscut al vostre costat.

# EL CÀNCER DE PRÒSTATA; MÉS VAL PREVENIR QUE CURAR

*Premi de Comunicació Científica JOAN LLUÍS VIVES, Edició 2009*

*Modalitat de Ciències Bàsiques, Ciències de la Salut, enginyeries i arquitectures*

Fa 4 dies el món sencer estava alertat per una nova variant de la grip porcina, la Grip A, causada pel virus H1N1. L'11 de juny del 2009 l'Organització Mundial de la Salut (OMS) va classificar aquesta nova grip com a pandèmia. Les autoritats sanitàries de tot el planeta vàren posar en marxa protocols d'emergència per por que aquesta nova soca viral provoqués més baixes de les esperades. La Directora general de l'OMS, Margaret Chan, advertia el següent; "*pot ser que en un mes aquest virus desapareix-hi, pot ser que quedi com està o pot ser que s'agreugi la situació actual*".

La taxa de mortalitat de la malaltia, que en un principi va ser elevada, va passar a ser molt baixa després d'iniciar els tractaments amb antivirals. Gairebé 100 anys abans, la "Grip Espanyola", també coneguda com "L'epidèmia de grip de 1918" va causar la mort de 25 a 40 milions de persones a tot el món. Va ser vehiculada per un virus de la mateixa soca que la Grip del 2009, un virus gripal del tipus H1N1.

Una malaltia infecciosa és una manifestació clínica conseqüència d'una infecció provocada per un microorganisme o per prions. Si coneixem quin és el microorganisme que causa la malaltia, en la majoria dels casos, som capaços de combatre'l i, com hem viscut aquest darrer any, pot passar d'ocupar primeres planes dels diaris a desaparèixer de les nostres ments en qüestió de poc temps.

Tots nosaltres coneixem, però, algun amic, amic dels nostres pares o fins i tot hem patit la proximitat d'un familiar afectat per una malaltia, que ha passat a formar part de la nostre societat i que malauradament encara avui en dia no es pot curar, *el càncer*. Què és el càncer? Perquè sembla que tothom deixa aquest món per causa del càncer? Perquè, en un món on la tecnologia i els avenços científics estan a l'ordre del dia, no som capaços de trobar què és allò que fa que la gent mori, a vegades sense adonar-se'n.

El càncer és la segona causa de mort a Catalunya. Segons l'OMS, el nombre de morts per càncer no para d'augmentar, a nivell mundial. D'aquí al 2030 es calcula que el nombre de morts anualment s'haurà pràcticament doblat, passant de quasi 8 milions de persones el 2007 a 11 milions i mig el 2030.

El càncer és un tipus de malaltia en què, un grup de cèl·lules adquireixen un fenotip que els permet créixer de forma descontrolada i dividir-se més enllà del normal, donant lloc al que

anomenem tumors malignes. Alhora aquestes cèl·lules són capaces d'envair cèl·lules veïnes o teixits adjacents i fins i tot d'arribar a altres punts de l'organisme mitjançant el sistema limfàtic o la circulació sanguínia, creant així nous nuclis tumorals, que anomenem metàstasis.

La majoria dels càncers són provocats per anomalies en el material genètic de les cèl·lules canceroses. Les anomalies genètiques poden ser causades per carcinògens externs, els quals es poden adquirir de forma aleatòria, per errors en els sistemes de reparació del material genètic, o bé poden heretar-se, i per tant, estan presents a totes les cèl·lules des del naixement de l'individu.

El diagnòstic definitiu d'aquesta malaltia acostuma a realitzar-se mitjançant l'anàlisi histològica del teixit afectat, el qual s'obté mitjançant una biòpsia. Alhora, existeixen un seguit de marcadors de cribatge (screening), en sèrum, que ajuden al metge en la decisió de realitzar o no aquesta biòpsia. La majoria de càncers quan són detectats en estadis inicials, és a dir, quan aquest està localitzat dintre de l'òrgan i encara no ha envaït teixits adjacents, es poden tractar eliminant el focus tumoral. Però, si el tumor es detecta en estadis ja avançats, en la majoria dels casos no es coneix encara cap teràpia curativa. En la majoria dels casos, el tractament es basa en una combinació de cirurgia, quimioteràpia (administració de tòxics químics) i radioteràpia (radiacions ionitzants), tots ells ataquen d'una forma més o menys directe les cèl·lules canceroses.

## EL CÀNCER DE PRÒSTATA, EL DILEMA DIAGNÒSTIC

El càncer de pròstata, és la segona causa de mort per malaltia oncològica en els homes del món occidental. S'estima que un de cada sis homes desenvoluparà un càncer d'aquest tipus al llarg de la seva vida. A Catalunya les dades recollides en un estudi realitzat des del 1998 fins al 2002, a càrrec de l' IRAC (International Agency for Cancer Research), demostren que, el càncer de pròstata és el tipus de càncer més freqüent que afecta a la població masculina.

El càncer de pròstata afecta com el seu nom indica, a la pròstata. La pròstata, és un òrgan glandular de l'aparell genital-urinari masculí. Té la mida d'una nou i es localitza sota de la bufeta, envoltant la uretra i davant del recte. La seva funció és secretar productes que s'afegiran al líquid seminal amb la finalitat de nodrir i protegir els espermatozous.

El diagnòstic actual del càncer de pròstata es basa en una triada diagnòstica que consta de; l'anàlisi dels nivells de PSA (Prostate Specific Antigen) en sèrum, el tacte rectal (TR) i finalment la biòpsia prostàtica. El PSA és una proteïna secretada exclusivament per les glàndules prostàtiques, que participa en la liquació del semen ejaculat. Avui en dia, encara no es coneix si aquesta proteïna juga un paper important en el desenvolupament del càncer de pròstata, però si més no, el seu ús com a marcador d'screening, és conegut i utilitzat

mundialment. Des de la seva introducció a finals dels anys 80, el diagnòstic del càncer de pròstata ha millorat significativament i, tot i que la mortalitat deguda a aquest tipus de càncer ha disminuït, segueix sent una de les formes canceroses més mortíferes que ataquen a la nostre societat. Quan els nivells de PSA en sèrum es situen per sobre de 4 ng/mL, l'uròleg pot estimar quina és la probabilitat que el pacient estigui afectat per un càncer. A més a més, juntament amb la palpació de la pròstata a través del tacte rectal, es decideix la necessitat de practicar o no una biòpsia prostàtica, que permetrà establir el diagnòstic definitiu.

El principal problema del PSA com a marcador d'screening és que aquest presenta un nivell d'especificitat baix, un 33%, i alhora, un valor predictiu negatiu baix. L'especificitat d'una prova es defineix com la probabilitat que un individu sa tingui un resultat negatiu i, el valor predictiu negatiu, com la probabilitat que la malaltia no estigui present quan la prova resulta negativa. Exemplificat de forma numèrica, això implicaria que, de cada 100 pacients que es sotmetin a un test de PSA i tinguin valors superiors a 4 ng/mL, mitjançant la biòpsia es diagnosticaran aproximadament, 66 pacients negatius i només 33 pacients positius per càncer de pròstata.

Per tal de detectar aquesta malaltia de forma precoç, s'aconsella que els homes a partir dels 50 anys es facin regularment una anàlisi dels nivells de PSA en sèrum i, quan els nivells del marcador superin els 4 ng/mL o bé quan el pacient tingui antecedents de la malaltia, es realitzarà un tacte rectal per tal de determinar la necessitat de realitzar una biòpsia. El problema és que, així com la societat femenina es sotmet regularment a exàmens ginecològics exploratoris, en la societat masculina existeix un rebuig cap a aquest tipus d'exploracions.

És per tot això, que el càncer de pròstata, es beneficiaria de l'existència de nous marcadors d'screening més específics i alhora d'un diagnòstic menys invasiu. Per altra banda, una millora en el diagnòstic evitaria un gran nombre de biòpsies innecessàries i conseqüentment un important estalvi econòmic en el cost sanitari actual.

La recerca de nous marcadors en el càncer de pròstata suposa un camp de treball important en la detecció precoç d'aquest tipus de càncer. Donada la situació de la pròstata a l'organisme, sota la bufeta i envoltant la uretra, les secrecions i inclús les mateixes cèl·lules prostàtiques, ja siguin normals o malignes, poden trobar-se presents en l'orina. És per això que considerem l'orina com una font important d'informació, a través de la qual es podria arribar a determinar quina situació s'està donant a l'òrgan en qüestió.

Altres estudis evidencien l'existència de potencials biomarcadors en l'orina que podrien ajudar en la millora del diagnòstic del càncer de pròstata. A nosaltres ens ocupa l'estudi

d'aquelles molècules proteiques que es troben a l'orina, ja siguin secretades per les cèl·lules prostàtiques o bé de les mateixes cèl·lules que, per descamació, van a parar a l'orina.

Suposem doncs, que un massatge prostàtic enriqueix la mostra d'orina de tot tipus de molècules proteiques. Així doncs, la nostra hipòtesi de treball recolza que, l'orina després d'un massatge prostàtic pot ser el fluid ideal per a la recerca de nous biomarcadors capaços de discriminar entre pacients amb o sense càncer de pròstata.

## ANÀLISI PROTEÒMICA COMPARATIVA

Al segle XXI ens trobem en plena era de les "òmiques". El terme òmic es refereix a la anàlisis global dels sistemes biològics. Dins un sistema biològic el "tot" guanya importància sobre les "parts". El proteoma és la imatge dinàmica de totes les proteïnes expressades per un organisme, en un moment donat i sota determinades condicions de temps i ambient. La proteòmica es defineix com l'anàlisi del proteoma.

Les proteïnes són compostos orgànics formats d'aminoàcids arranjats en una cadena lineal i units per enllaços pèptids. La seqüència d'aminoàcids que forma una proteïna ve definida per l'ADN, seqüència de nucleòtids del gen que la codifica. Poc després o fins i tot durant la síntesi, els residus d'una proteïna, sovint són modificats químicament per modificacions post traduccionals, que alteren les propietats físiques i químiques, el plegament, l'estabilitat, l'activitat i la funció de la proteïna.

El dogma general de la biologia, fins fa poc temps, establia que un gen (DNA) es transcriu a un RNA missatger (mRNA) i posteriorment es tradueix per donar lloc a una proteïna. Avui en dia es sap que degut al processament alternatiu (splicing), un mateix gen pot donar lloc a diferents proteïnes. Per tant, si el genoma humà consta d'uns 30.000 gens el nombre de proteïnes és exponencialment major. És per això, que avui en dia l'estudi del proteoma, i per tant l'anàlisi proteòmica d'elevat rendiment (High Throughput Analysis), estigui adquirint molts importància en el camp de la recerca.

Els estudis proteòmics es basen principalment amb la tècnica de l'espectrometria de masses amb combinació de mètodes de fraccionament i separació de proteïnes i pèptids (fragments proteics), com l'electroforesi bidimensional (2D-PAGE) o la cromatografia líquida d'elevada resolució (HPLC). L'espectrometria de masses es una tècnica experimental que permet la mesura de ions derivats de molècules (proteïnes i pèptids). Mitjançant aquesta metodologia es poden realitzar anàlisis qualitatives/comparatives però també quantitatives de tot tipus de mostres proteiques.

## DISSENY EXPERIMENTAL I RESULTATS PRELIMINARS

Engeguem el nostre projecte a l'Hospital de la Vall Hebron que ens abasteix de totes les mostres que necessitem, clínics dedicats i gent amb moltes ganes de treballar.

L'objectiu principal és arribar a determinar un perfil proteòmic a l'orina capaç de diferenciar entre pacients amb càncer de pròstata i controls sans.

En primer lloc l'estudi requereix d'una recerca exhaustiva d'un grup de pacients, tots ells ben caracteritzats; homes d'edat avançada, que es sotmeten a un diagnòstic ràpid de càncer de pròstata, amb nivells de PSA per sobre de 4 ng/mL i per tant amb una certa probabilitat d'amagar un càncer de pròstata. Classifiquem els pacients en dos grups: *els problemes*, o dit d'una altra manera aquells que pateixen la malaltia (ja sigui en estadis inicials o avançats) i no menys importants, *els controls*, aquells qui no presenten cap anomalia considerable (inflamació aguda o crònica, atròfia proliferativa o lesió preneoplàsica d'alt grau "HG-PIN"), en resum, controls sans. De tots ells s'obté una mostra d'orina després de la realització d'un massatge prostàtic. La mostra d'orina es processa d'acord amb el protocol estandarditzat per tal d'extreure la fracció proteica que en forma part. Tots els pacients seleccionats per l'estudi es sotmetran posteriorment a una biòpsia prostàtica, de manera que coneixerem amb certesa quin és el seu diagnòstic definitiu.

La metodologia escollida per dur a terme l'anàlisi comparativa s'anomena Differential in Gel Electrophoresis (DIGE). Aquesta tècnica permet, mitjançant electroforesi bidimensional, separar les proteïnes pel seu pes molecular i per la seva càrrega (punt isoelèctric), situant-les en un espai de 2D (gel bidimensional). Alhora, permet marcar les mostres control i problema amb diferents fluorocroms, els quals ens permetran visualitzar les diferents mostres en un mateix gel, i per tant, compara-les entre si. Comparem un total de 9 mostres control amb 9 mostres problema. De l'anàlisi comparativa n'extraiem la informació necessària per tal d'establir quines són les proteïnes (punts situats a l'espai de 2D) que més difereixen entre ambdós tipus de mostres. Determinats els punts més rellevants, els retallem del gel i n'identifiquem el seu contingut, mitjançant espectometria de masses, és a dir quina proteïna correspon.

Un llistat, vint-i-quatre proteïnes a destacar, però quina és la més important? No som capaços de senyalar-la amb el dit ja que podríem errar, potser no és una, potser en són varies plegades o potser ninguna.

Arribats a aquest punt, l'estudi s'endinsa en un experiment de validació. Validar; fer vàlid, verificar, donar validesa... Una validació experimental implica l'obtenció del mateix resultat mitjançant un altre mètode experimental. En aquest cas és important que el mètode

seleccionat ens permeti avaluar els nivells de les proteïnes identificades en un gran nombre de pacients. Ha de ser doncs un mètode altament reproduïble i molt sensible, capaç de detectar les proteïnes a baixa concentració. Partirem d'un nombre elevat de mostres d'orina, per tal de donar certesa als resultats. Les mostres s'obtenen i es processen de la mateixa manera que anteriorment, però en aquest cas la població no es selecciona sinó que s'agafa a l'atzar, d'aquesta manera podem saber quin és el comportament d'aquestes proteïnes a la població real.

Els resultats que s'obtinguin d'aquest experiment poden ser doncs, concloents, no concloents, i per tant requerir d'una anàlisi encara més minuciosa, o fins i tot no satisfactoris, de moment no ho sabem. Resultats preliminars senyalen que no només serà una proteïna sola sinó un conjunt de proteïnes que a través de l'orina potser podran, algun dia, a ajudar al facultatiu a prendre una decisió, això sí, creiem que sense deixar de banda l'actual marcador diagnòstic, el PSA.

*"Pot ser que ajudi en el diagnòstic, pot ser que finalment no en traiem res i per tant que quedi com està, però en aquest cas la situació no s'agreujarà, ja que som molts els qui estem lluitant contra aquesta malaltia, ja sigui patint-la o combatin-la"*.

Al meu avi Lluís i a la meva àvia Isabel

# INDEX

# ABREVIATIONS

Active surveillance (AS)

Alpha-methylacyl-CoA racemase (AMACR)

Androgen deprivation therapy (ADT)

Androgen receptor (AR)

Anterior fibromuscular zone (AFZ)

Area Under the Curve (AUC)

Area Under the MultiROC curve (AUCm)

Atypical Small Acinar Proliferation (ASAP)

B-cell lymphoma 2 (Bcl-2)

Benign pathology (BP)

Benign prostatic hyperplasia (BPH)

Calgranulin B (MRP-14)

Capillary electrophoresis coupled to mass spectometry (CE-MS)

Center for Medicaid and Medicare service (CMS)

Central zone (CZ)

Circulating tumor cells (CTCs)

Clusterine (CLU)

Coefficient of variancy (CV)

Culture media (CM)

Differential In-Gel Electrophoresis (DIGE)

Digital rectal exam (DRE)

Immunoglobulin (IgG)

Internal standard peptides (ILP)

Enzyme-Linked ImmunoSorbent Assay (ELISA)

Epithelial to mesenchymal transition (EMT)

European Medicines Agency (EMA)

European Randomized Study of Screening for PCa (ERSPC)

Expressed-prostatic secretion (EPS)

External beam radiation (EBRT)

Extracted ion current (XIC)

Food and Drug Administration (FDA)

Free PSA (fPSA)

Glutathione S-transferase P1 (GSTP1)

GTP-binding proteins (GPCRs)

High Grade Prostatic Intraepithelial Neoplasia (HGPIN)

Human Genome Project (HGP)

Human Proteome Project (HPP)

Human serum albumin (HAS)

Immunohistochemistry (IHC)

Insulin-like growth factor (IGF)

Interleukine-6 (IL-6)

International Agency for Research on Cancer (IARC)

Intraepithelial Neoplasia (PIN)

Kallikrein-related peptidase 3 (KLK3)

Limit of Detection (LOD)

Limit of Quantification (LOQ)

Liquid chromatography (LC)

Loss of heterozygosis (LOH)

Lower urinary tract symptoms (LUTS)

Low grade PIN (LGPIN)

Mass spectrometry (MS)

Mass-to-charge ratio (m/z)

Matrix Assisted Laser Desorption Ionization (MALDI)

Messenger RNA (mRNA)

MicroRNAs (miRNAs)

Multiple Reaction Monitoring (MRM)

Multivariate ROC (MultiROC)

Negative predictive value (NPV)

Non-conding RNA (ncRNA)

Odorant receptor (OR)

Peptide mass fingerprint (PMF)

Peripheral zone (PZ)

Positive predictive value (PPV)

Post-translational modifications (PTMs)

Proliferative inflammatory atrophy (PIA)

Prostate acid phosphatase (PAP)

Prostate biopsy (PB)

Prostate cancer (PCa)

Prostate cancer antigen 3 (PCA3)

Prostate Lung Colorectal and Ovarian Cancer (PLCO)

Prostate massage (PM)

Prostate specific antigen (PSA)

Prostate specific G-protein coupled receptor (PSGR)

Prostate specific membrane antigen (PSMA)

Proteotypic peptides (PTP)

PSA density (PSAD)

PSA doubling time (PSADT)

PSA velocity (PSAV)

Quality Assessment (QA)

Quality Control (QC)

Reactive oxygen species (ROS)

Real Time Polymerase Chain Reaction (RTqPCR)

Radical Prostatectomy (RP)

Receiver Operating Characteristics (ROC)

Reduction by DUtasteride of Prostate Cancer Events (REDUCE)

Selected Reaction Monitoring (SRM)

Single nucleotide polymorphisms (SNPs)

Small interfering RNAs (siRNAs)

Surface-enhanced Laser Desorption Ionization-Time-of-Flight (SELDI-TOF)

Time-of-Flight Mass Spectrometry (TOF-MS)

Transcription-mediated amplification (TMA)

Transforming growth factor ß3 (TGFB3)

Transition zone (TZ)

Transrectal ultrasound guided prostate biopsies (TRUS)

Triple quadrupole mass spectometers (QQQ-MS)

Tumoral node metastasis (TNM)

Tumor initiating cells (TICs)

Two-dimensional (2D)

Two-dimensional polyacrylamide gel electrophoresis (2DE)

Watchful waiting (WW)

5-alpha-dihydrotestosterone (DHT)

# INTRODUCTION

# 1. THE PROSTATE

## 1.1 ANATOMY, MORPHOLOGY AND FUNCTION

The human prostate (from the Greek word, prostates, meaning "one who stands before," "protector," or "guardian") is a walnut-sized tissue. In a young adult, the prostate weighs approximately 20gr and measures 3cm in length. Its main function is to store and secrete a slightly alkaline fluid and proteins that are supposed to provide nutritional support to the seminal fluid. Prostate compounds constitute approximately 25% of the volume of semen, together with spermatozoa and seminal vesicle fluid. The alkalinity of semen helps to neutralize the acidity of the vaginal tract, prolonging the lifespan of sperm. Prostatic fluid is expelled in the first ejaculate fraction together with spermatozoa and the fluid of the seminal vesicle. In addition, prostatic fluid improves the motility of spermatozoa, promotes their longer survival and provides better protection to the DNA. The prostate is located in the pelvis, under the urinary bladder and in front of the rectum. The prostate surrounds part of the urethra, the tube that carries urine from the bladder during urination and semen during ejaculation (Figure 1a). Because of its location, prostate diseases mainly affect the processes of urination and ejaculation, though rarely defecation.

The prostate can be divided into two different classifications: zone and lobe.

The zone classification is used more often in pathology and was first proposed by McNeal in 1968 [1]. According to this classification, the prostate can be divided into 4 zones: the peripheral zone (PZ) and the central zone (CZ), which together comprise <95% of the prostate mass in the prostate of a normal man, the periurethral transition zone (TZ), and finally, the anterior fibromuscular zone (AFZ) or stroma and the periurethral glandular zone [2] (Figure 1b). The PZ zone constitutes the bulk of the apical, posterior, and lateral prostatic tissue and accounts for most of the glandular tissue (70%). It is the zone in which 70% of prostate cancers (PCa) emerge. The TZ accounts for 5-10% of the glandular tissue of the prostate. Cellular proliferation in the TZ results in benign prostatic hyperplasia (BPH). In addition, 20% of the cases of PCa arise in this zone. The CZ surrounds the ejaculatory ducts. Only 2.5% of the reported cases of PCa appear in this zone, but these cancers tends to be more aggressive and more likely to invade the seminal vesicles [3].

The *lobes classification* is used more often in anatomy. According to this classification, the prostate can be divided into several lobes: the anterior lobe, or the anterior portion of the gland lying in front of the urethra, the median lobe that is situated between the two ejaculatory ducts and the urethra, the lateral lobes (right and left) that are separated by the

prostatic urethra constituting the main mass of the gland, and finally, the posterior lobe, which is used to describe the posteromedial part of the lateral lobes and which can be palpated through the rectum during the digital rectal exam (DRE) (Figure 1c).

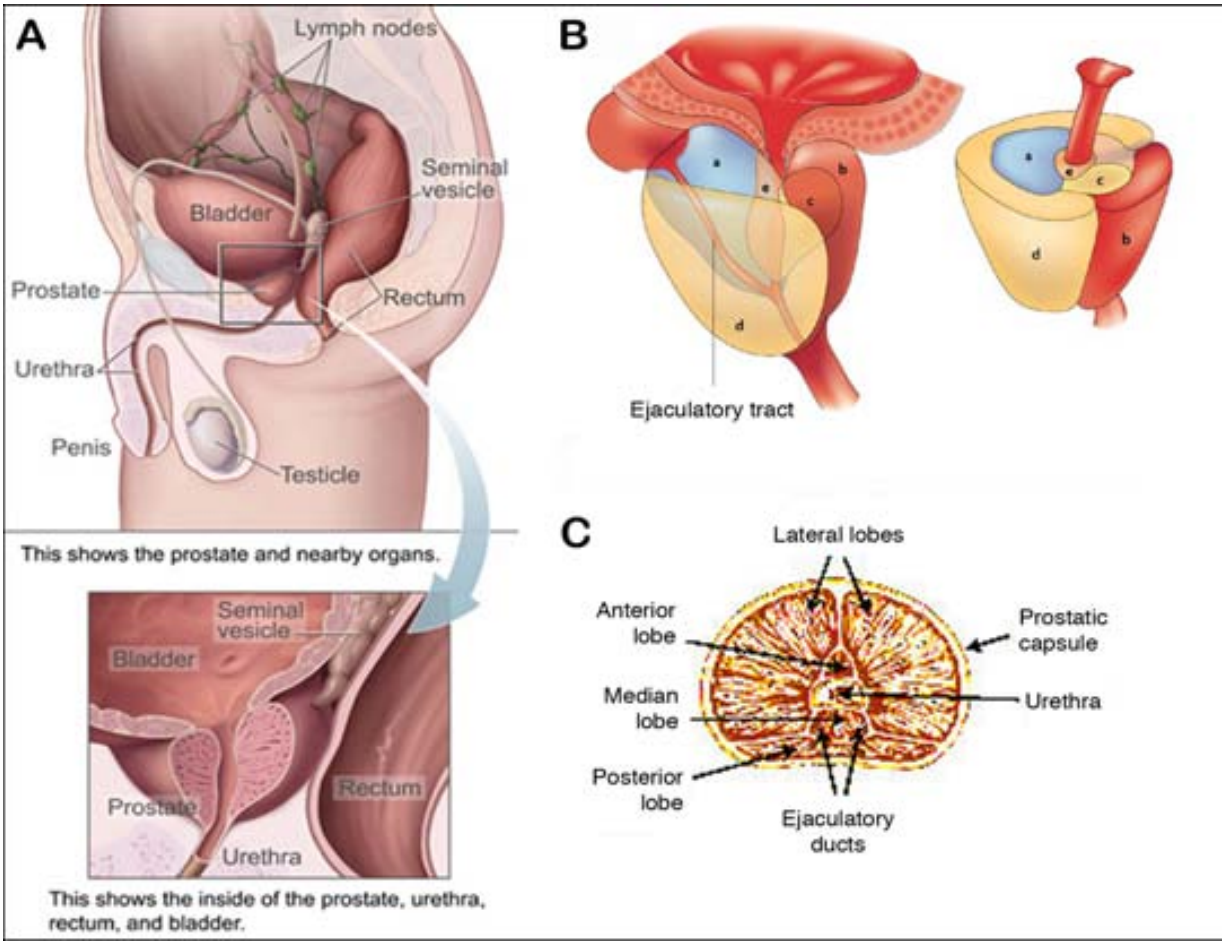



**Figure 1. The human prostate.** (A) Localization of the prostate in humans, adapted from http://en.wikipedia.org/. (B) Prostate zones: a, central zone (CZ); b, fibromuscular zone (AFZ); c, transitional zone (TZ); d, peripheral zone (PZ); e, periurethral region, adapted De Marzo *et al,* 2007 [2] (C) Prostate lobes; lateral lobes, anterior lobes, median lobe and posterior lobe, adapted from http://en.wikipedia.org/.

At the histological level, the human prostate contains a pseudostratified epithelium that consists of tubuloalveolar glands, which are embedded in the fibromuscular stroma. The prostate glands are assembled by a well-defined basement membrane covered by low cuboid, basal epithelial and mucous-secreting columnar cells that rest on top of basal cells (Figure 2).

The mature prostate contains five inter-related cell types organized to form prostatic glands:

1. The predominant epithelial cell type is the *secretory luminal* cell. Luminal epithelial cells form a continuous layer of polarized columnar cells. These cells are differentiated androgen-dependent cells that produce protein secretions and express characteristic markers, such as cytokeratin 8 and 18, the cell surface marker CD57 as well as Androgen Receptor (AR) [4].

These cells also secrete Prostate Specific Antigen (PSA) and Prostate Acid Phosphatase (PAP) into the glandular lumen.

2. The second major epithelial cell type corresponds to the *basal cells*. Basal cells form a single layer on the basement membrane underlying the normal prostate epithelium of each prostatic gland. These cells are relatively undifferentiated and conform to the proliferative compartment. They express p63 and high-molecular cytokeratin 5 and 14 as well as CD44. However, they express low levels of AR and they lack secretory activity [4].

3. The third epithelial cell type is the *neuroendocrine cell*. These cells comprise only a small percentage of the normal prostatic epithelium of uncertain embryological origin. The neuroendocrine peptides secreted by these cells support the growth or differentiation of the surrounding epithelial cells. Neuroendocrine cells express endocrine markers, such as chromogranin A and serotonin [4]. However, these cells do not express AR.

4. Transit-amplifying or intermediate cells express features of the basal and secretory and/or neuroendocrine cells. They express cytokeratin 5 and 8/18 and they are negative in p 63. It has been proposed that these proliferative cells are stem-like cells, or progenitor cells, that are in a process of transition to a differentiated secretory phenotype [5].

5. Stem cells are believed to reside in the basal epithelial layer of the gland. The presence of stem cells in the prostate was first proposed to explain the capacity of the organ to regenerate during androgen cycling experiments [6].

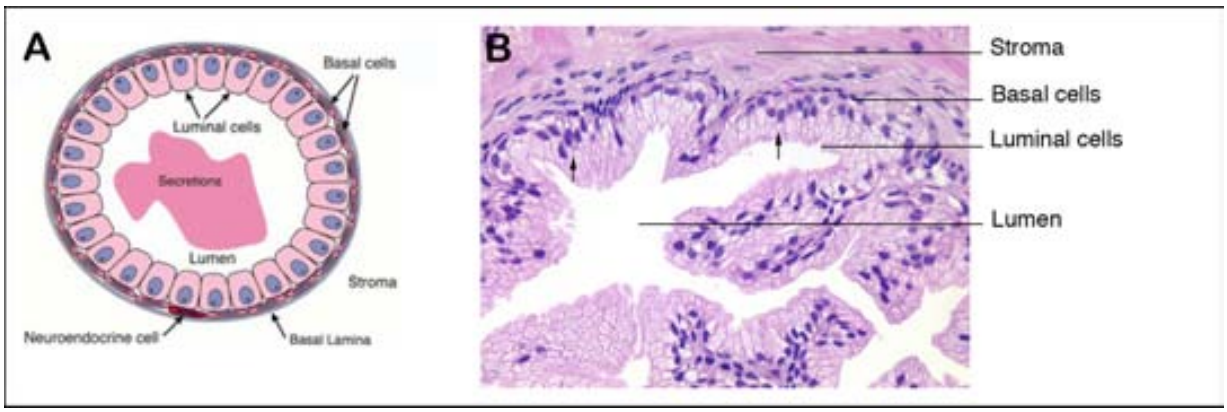



**Figure 2. Prostate gland.** (A) Schematic representation of the prostate gland. Cell types within a human prostatic duct are shown. Note that rare neuroendocrine cells are morphologically indistinguishable from basal cells [4]. (B) Hematoxilina-eosin staining of normal prostatic tissue adapted from Epstein, JI and Netto GJ [7].

The prostate also contains several types of stromal cells, including fibroblasts, myofibroblasts, and smooth muscle cells that guide the growth and differentiation of the epithelium. Nerves and blood cell elements, as well as vascular and stromal endothelial cells, are also present in the gland.

BOX I: FORMATION AND MORPHOGENESIS OF THE PROSTATE

The development of the prostate gland is largely controlled by sex hormones and is initiated in the uterus at week 12 of fetal development. Prostate formation occurs thorugh epithelial budding from urogenital sinus (endodermal origin). During midgestation, the primitive urogenital sinus is separated from the terminal region of the hindgut through the division of the cloaca by the urorectal septum. The most frontal region of the urogenital sinus forms the urinary bladders, whereas the most caudal region forms the urethra. The prostate gland originates from the intermediate region (pelvic part or urogenital sinus) [4].

## 1.2 PROSTATE DISORDERS

### 1.2.1. PROSTATITIS OR CHRONIC INFLAMMATION

The term, *prostatitis*, or chronic inflammation refers to the histological inflammation of the tissue of the prostate gland that may be associated with a large, indefinite number of lower urinary tract symptoms (LUTS) and sexual discomfort and dysfunction (Figure 3a). This condition affects 5-10% of the male population and is the most common urologic diagnosis in men younger than 50 years of age. It can be associated with an appropriate response of the body to an infection, but it can also occur in the absence of infection. Prostatitis is classified into four categories, including acute and chronic bacterial forms, a chronic a-bacterial form, and an asymptomatic form.

Emerging evidence suggests that among the many risk factors for developing PCa and its progression to metastasis, inflammation represents a major risk [2]. Chronic inflammation of longstanding duration has been linked to the development of carcinoma in several organ systems. The proposed mechanism of carcinogenesis involves repeated tissue damage and regeneration in the presence of highly reactive oxygen and nitrogen species. These reactive molecules are released from the inflammatory cells and can interact with DNA in the proliferating epithelium to produce permanent genomic alterations such as point mutations, deletions, and rearrangements [8].

### 1.2.2. BENIGN PROSTATIC HYPERPLASIA

BPH is one of the most common age-related disorders affeting men, representing the most frequent proliferative abnormality of the human prostate. The development of BPH takes place in the TZ [9]. BPH produces a progressive obstruction of the urethra that leads to urinary retention, bladder function impairment and eventually renal failure [10], and it affects 80% of

men by age 80. This implies that one in three of these patients will require treatment to alleviate the obstructive symptoms caused by the disease. Regardless of the obvious importance of BPH as a major health issue, which significantly affects the quality of life in aging men, its etiology and pathogenesis remain unclear [11]. The architectural development of BPH is well characterized by multifocal expansive nodules originating from a budding and branching of the epithelial glandular ducts and acini, as well as a proliferation of the prostatic stromal elements [12] (Figure 3b).

The cellular and molecular events that contribute to BPH are not well characterized, though recent data from various studies support a shift in the balance between cellular growth and apoptosis and senescence [13, 14]. A higher proliferative-to-apoptotic ratio has been observed in BPH than in normal tissue and an increased expression of the anti-apoptotic factors has also been noted in BPH epithelial cells. Recently, our group have demonstrated the accuracy of the beagle dog animal model, similar to the human clinics, for developing a transcriptome analysis that characterizes a gene expression pattern associated with the onset of BPH [14]. Our results pointed to a number of pathways altered during the initial steps of BPH. These pathways could be involved in both the promotion and the tissue response to the hyperplastic growth. In addition to the genes involved in calcification, matrix remodeling, detoxification, cell movement, and mucosal protection, the up-regulation of transforming growth factor ß3 (TGFB3) and Clusterine (CLU) indicated a complete adjustment of the transdifferentiation, senescence and apoptosis programs. In addition to contributing new clues towards an understanding of hyperplastic growth, this study characterizes a number of genes that could represent new potential therapeutic targets, such as CLU. Considering that BPH is a major health problem that significantly affects the quality of life in aging men, the development of new therapeutic strategies for improving the outcome of BPH is imperative.

Although PSA levels may be elevated in these patients, due to increased organ volume and inflammation attributable to urinary tract infections, BPH is not considered to be a premalignant lesion. An estimated 42% of men from 51 to 60 years of age and up to 80% of men between the ages 70-80 have presented histological evidence of this prostate disorder.

### 1.2.3. PROLIFERATIVE INFLAMMATORY ATROPHY

Tissue inflammation is normally associated with an appropriate response of the body to an infection. On the other hand, prostate atrophy is identified as a reduction in the volume of preexisting glands and stroma and can be divided into two major patterns: diffuse and focal [15]. Proliferative inflammatory atrophy (PIA) lesions are characterized by discrete foci of proliferative glandular epithelium with the morphological appearance of simple atrophy, or

post-atrophic hyperplasia, and they occur in association with inflammation (Figure 3c). The key features of this type of lesion are the presence of two distinct cell layers: mononuclear and/or polymorphonuclear inflammatory cells in both the epithelial and stromal compartments and stromal atrophy with variable amounts of fibrosis. The morphology of PIA is consistent with the description of post-inflammatory atrophy, with that of chronic prostatitis, and with the lesion referred as "lymphocytic prostatitis." As well as showing an increased staining for cell proliferation markers, the key immunophenotypic features of PIA that have been described are an increased staining for glutathione S-transferase P1 (GSTP1) and B-cell lymphoma 2 (Bcl-2) and a decreased staining for the cyclin-dependent kinase inhibitor p27Kip1 [16]. PIA is proposed to be a common proliferative response to environmental damage, such as inflammation and oxidative stress in aging men, which can be caused by repeated infections, dietary factors and/or the onset of autoimmunity.

Regions of PIA are often located in proximity to High Grade Prostatic Intraepithelial Neoplasia (HGPIN) and PCa. About 20% of all human cancers are caused by chronic infection or chronic inflammatory states. It has been proposed that exposure to environmental factors such as infectious agents and dietary carcinogens, and hormonal imbalances lead to injury of the prostate and to the development of chronic inflammation and regenerative 'risk factor' lesions, referred to as PIA [2]. Recent expression array profiling studies have suggested that PIA may only be a very early precursor, or even unrelated to the development of PCa [17].

### 1.2.4 ATYPICAL SMALL ACINAR PROLIFERATION

Atypical Small Acinar Proliferation (ASAP) is a collection of small prostatic glands found through prostate biopsy (PB), whose significance is uncertain and that have not been determined to be either benign or malignant. These glands often show distorted features including partial atrophic-like appearances, a lack of nuclear enlargement, a lack of cells showing prominent nucleoli, and associated inflammations [18, 19]. Foci of ASAP are often < 1mm in size. The incidence of ASAP in PB is approximately 2-3%. Most authors consider ASAP to be a diagnostic risk category and not a valid pathologic entity. However, several studies have demonstrated that in 34% to 60% of the specimens taken from cases subjected to a subsequent biopsy, this diagnosis is predictive of malignancy[20]. Consequently, a repeat biopsy is recommended in cases where an initial diagnosis of ASAP has been proffered.

### 1.2.5. HIGH GRADE INTRAEPITHELIAL NEOPLASIA

Prostate Intraepithelial Neoplasia (PIN) is defined histologically by the presence of nuclear and cytoplasmic features, which are similar to those of PCa in glands with normal architecture [21].

However, unlike cancer, PIN retains the cell layer [22]. As this process is confined to the epithelium, it is, therefore, termed intraepithelial [22]. In addition, PIN lesions generally display a marked elevation of cellular proliferation markers [23] within the pre-existing secretory epithelium, ducts and acini. Notable cytological changes include prominent nucleoli in at least 5% of the cells, nuclear enlargement, nuclear crowding, an increased density of the cytoplasm and a variation in nuclear size [24].

The incidence and extension of PIN increase with age [25], although findings from a study on the autopsies of a few young men revealed that microscopic PIN was detected in the prostate of those men [26]. This prostate disorder is classified into a two-tier classification, based on the cytological characteristics of the secretory cells: low grade PIN (LGPIN) and high-grade PIN (HGPIN) [27] (Table 1).

**Table 1. Characteristics of the to types of PIN lesions [22]**

| PIN | Nucleoli of the cells | Staining for AMACR | Basal cell-layer |
|---|---|---|---|
| LGPIN | Enlarged, vary in size, have a normal or little increase in the chromatin content and possess small or inconspicuous nucleoli | Negative | Intact |
| HGPIN | Large nucleoli rather uniform size, an increased chromatin content and prominent nucleoli that are similar to those of PCa cells | Postivie in the citoplasma [28] | Highly disrupted |

HGPIN is considered most likely to represent a forerunner to PCa, based on several lines of evidence: (1) the incidence and extent of HGPIN in the prostate increase with advancing age [29, 30]; (2) HGPIN lesions are found in the PZ, where most prostate tumors occur [50]; (3) the frequency, severity and extent of HGPIN increase in the presence of PCa; (4) the appearance of HGPIN lesions generally precedes the appearance of carcinoma by at least 10 years, which is consistent with the idea of cancer progression; (5) rates of cell proliferation and deaths are elevated in HGPIN and PCa when compared to the rates for normal prostates; (6) chromosomal abnormalities and allelic imbalance analyses have shown that HGPIN lesions are multifocal, as is the case with carcinomas [31]; (7) the architectural and cytological features of HGPIN closely resemble those of invasive carcinoma, including a disruption of the basal layer and the presence of prominent nucleoli (Figure 3d); (8) differentiation markers that are commonly altered in early invasive carcinoma are also altered in HGPIN lesions [24]; and, (9) the rate of neovascularization is raised in HGPIN and in PCa when compared to the rate found in normal prostates.

Nevertheless, recent studies suggest that the majority of alterations that occur in the progression to PCa take place in the transition from benign epithelium to HGPIN, rather than

from HGPIN to PCa [17]. On the other hand, HGPIN differs from invasive carcinoma in that it normally retains the basal cellular membrane and does not invade the stroma. In addition, HGPIN lesions do not produce high levels of PSA, and consequently, HGPIN can only be detected by biopsy and not through serum PSA testing. Alpha-methylacyl-CoA racemase (AMACR) expression has been found to be negative to weakly positive in biopsy specimens containing HGPIN without carcinoma and weakly positive in radical prostatectomy specimens, while its expression was highly positive in HGPIN lesions adjacent to adenocarcinoma [32]. Similar results were obtained with the immunohistochemical study of GSTP1 expression [33]. Finally, present studies are currently assessing the role of PTOV1 as it may be linked to PCa in detecting the risk of carcinoma in repeated biopsies following a diagnosis of HGPIN [34].

The magnitude of the risk for PCa in men with HGPIN and the optimal follow-up strategies remain controversial. Prostate needle biopsies have provided a reported 1.5% to 31% incidence of HGPIN [35]. Finding an HGPIN through biopsy is a frequent indication for repetition of the biopsy [36]. In addition, after repeated PBs with LGPIN, a 16% incidence of PCa was reported [29], whereas the reported incidence with HGPIN was 24% [37]. In early studies, using limited biopsy schemes, HGPIN was associated with high rates of PCa. It was suggested that its presence indicated an immediate repetition of the biopsy [38]. However, when a more extensive early biopsy scheme was initially used, the cancer detection rate was considerably lower. This was due to the fact that the number of cores sampled during the initial biopsy affected the likelihood of detecting PCa in the subsequent biopsies [39]. For this reason, some investigators believe that repeat biopsies may be unnecessary in the current era and that follow-up for these men can be accomplished using serial DREs and PSA measurements [40]. In fact, HGPIN does not contribute to the serum concentration of PSA or modify the percentage of free PSA (fPSA) [36]; however, PSA velocity (PSAV) helps to identify those men who possess a high likelihood of suffering from PCa and who have a real need for repeating the biopsy [41]. HGPIN either can be detected by DRE alone.

Bostwik and Brawer [42] described a progression model of PIN to carcinoma in which the transition from benign to LGPIN, to HGPIN, and then to PCa is continuous. Nevertheless, Putzi and De Marzo [43] found that LGPIN often coexists with HGPIN, suggesting that both forms arise concomitantly. However, due to the fact that LGPIN is not documented in pathology reports, this data remains controversial.

In summary, HGPIN seems to be a precursor lesion to many peripheral, intermediate to high-grade cases of PCa. However, this lesion is not necessarily a precursor, since many early cancers do not possess adjacent HGPINs. Also, low-grade carcinomas, especially those found

within the TZ, are not closely related to HGPIN [44]. Nevertheless, the recognition of this prostatic disorder is clinically important because of its association with PCa. Although the relationship between PIN and PCa has not yet been demonstrated conclusively, HGPIN has been widely accepted as a precursor lesion to PCa. For this reason, biomarkers for the detection of premalignant prostatic disease can help to identify those patients who will need further monitoring.

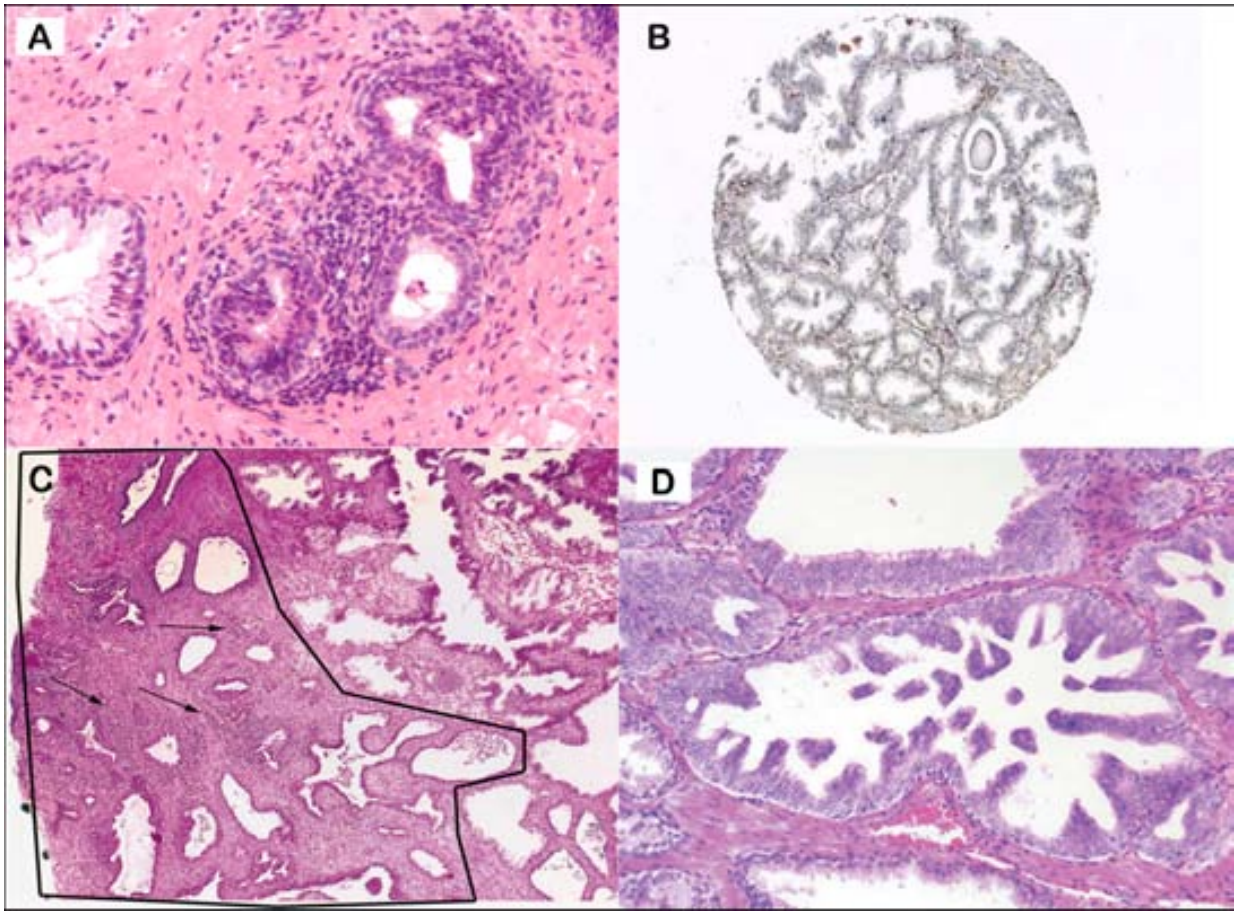



**Figure 3. Architecture of the main disorders afectin prostate.** (A) Prostatitis, (B) BPH, from human tissue microarray, from Arrieta *et al,* [12] (C) PIA (outlined area) occurring adjacent to benign normal appearing glands (lower right). Arrows indicate collections of chronic inflammatory cells (predominantly lymphocytes). This lesion was classified as having marked chronic inflammation. H&E, 40X [16] and (D) HGPIN. A and D are images of hematoxylin-eosin staining of normal prostatic tissue addapted from Epstein, JI and Netto GJ [7].

# 2. PROSTATE CANCER

## 2.1. EPIDEMIOLOGY OF PROSTATE CANCER

Cancer is one of the most important health problems in our society, both in terms of morbidity and social impact. It affects the economy and quality of life of one in every three people throughout their lifetimes [45]. PCa is the most commonly diagnosed cancer among European and American men (24.1% of all cases) [46], and it is the second most common cause of cancer death among men [47] (Figure 4). In 2010, an estimated 217,730 (28%) new cases of PCa were diagnosed in the US, and 32,050 (11%) died as a result of this disease [48]. Even though the introduction of the PSA test in the late 1980s of the past century has led to a dramatic increase its detection [46], the risk of developing this type of cancer during a lifetime is estimated at 1 in 6 men in the US, and the risk of death due to this disease is 1 in 36 [49].

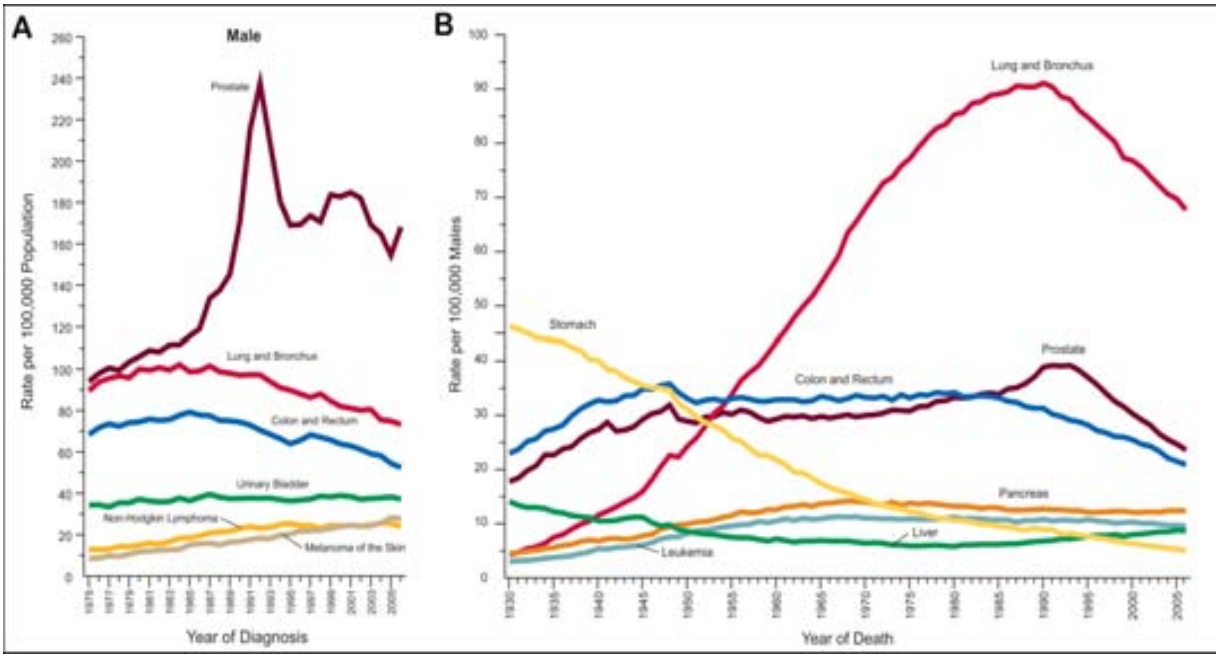



**Figure 4. Cancer Statistics by Jemal et al. 2010** [48]. (A) Annual Age-Adjusted Cancer Incidence Rates* for Selected Cancers, United States, 1975 to 2006. Source: Surveillance, Epidemiology, and End Results (SEER) (www.seer.cancer.gov). (B) Annual Age-Adjusted Cancer Death Rates* Among Males for Selected Cancers, United States, 1930 to 2006. Source: US Mortality Data, 1960 to 2006, US Mortality Vol. 1930 to 1959. National Center for Health Statistics Centers for Disease Control and Prevention. *Rates are age adjusted to the 2000 US standard population and adjusted for delays in reporting.

In Catalonia, PCa is the most frequent cancer ahead of lung and colorectal cancer, and it is the third most common cause of death due to cancer in males (International Agency for Research on Cancer (IARC) in 2007). PCa is the type of cancer that is most frequently reported and it presents a higher incidence in the first period analyzed (1985-1994) than in the last (1995-2002) (Figure 5a). This effect may be associated with a more widespread use of

PSA testing and an overall aging population. It is estimated that, as has occurred in other countries, this increased incidence is stable and is largely attributable to the diagnostic advancements related to the use of PSA testing, which would explain the fall in the last reported mortality period (Figure 5b).

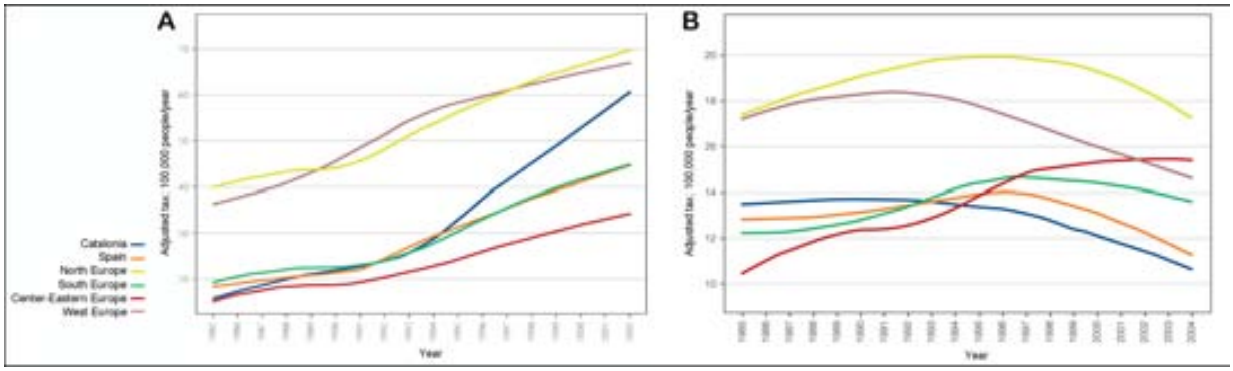



**Figure 5. PCa in Catalonia.** (A) Evolution of PCa in Catalonia, Spain and Europe from 1985 to 2002. [50] [51] [52]. (B) Evolution of the mortality for PCa in Catalonia, Spain and Europe from 1985 to 2004 [53].

PCa is a disease that affects aging men. It is, therefore, a greater concern for developed countries, which possess a larger proportion of elderly men in their populations. For this reason, in developed countries about 15% of the reported cancers in males are PCas, while in undeveloped countries, the percentage is only 4% [54].

## 2.2 RISK FACTORS OF PROSTATE CANCER

Like many other adult cancers, PCa likely represents the accumulation of various genetic insults that have developed over the course of decades. Even though the factors that determine the risk of developing PCa are not well known, a few have been identified. There are three well-established PCa risk factors: increasing age, heredity and ethnic origin. Moreover, dietary and lifestyle factors have been found to be major contributors to population differences in the occurrence of clinical PCa [55]. Dietary (eg. fat in the diet) and lifestyle differences may specifically account for the considerable differences in incidence of clinical PCa between Asian and American populations, reflecting a shift in the 10-year rate of cancer detection [56].

### 2.2.1 AGE

The most important risk factor for PCa is age. The probability of developing PCa varies greatly by age (Figure 6a). PCa incidence rates increase in men until about age 70 and decline thereafter, 70% of PCa are developed in men > 65 years. Moreover, death rates for PCa increase with age [49] (Figure 6a).

## 2.2.2. HEREDITY

Compared to men without a family history, men with one first-degree relative with the disease are 2 to 3 times more likely to develop PCa, while men with more than one affected first-degree relative are 3 to 5 times more likely to be diagnosed with PCa [57]. However, only 9% of the men with PCa have true hereditary PCa. This is defined as 3 or more affected relatives or at least 2 relatives who have developed early-onset disease (before age 55) [58]. Patients with hereditary PCa usually have an earlier onset of 6 to 7 years prior to those patients with spontaneous disease, but they do not differ in other ways [59]. Therefore, patients with a familiar PCa should be screened earier for this disease. An example of a described genetic factor that may play a role in PCa includes BRCA-1 and BRCA-2 mutations. Men with these mutations possess an increased risk of developing a more aggressive form of PCa that develops at a younger age [60-62]. Consistent evidence from genetic studies has also identified locations on chromosome 8 (in a region called 8q24) that are associated with an increased risk of developing PCa and with its more aggressive forms [63, 64].

## 2.2.3 ETHNICAL ORIGIN

Clinical PCa incidence and mortality may, in part, reflect genetic factors that vary in populations originating in different geographical areas. Incidence rates vary by more than 50-fold worldwide, with the majority of cases diagnosed in economically developed countries. The highest incidence rates are observed in North America (where there is a significant difference between African and white American men), Australia, and Northern and Central Europe. The lowest incidence rates are observed in South Eastern and South Central Asia and Northern Africa [49] (Figure 6b). However, if Asian men move from Asia to the USA, their risk of PCa increases. All of these findings indicate that exogenous factors affect the risk of progression from so-called latent PCa to clinical or symptomatic PCa. Factors, such as food and alcohol consumption, sexual behavior patterns, exposure to ultraviolet radiation and occupational exposure have all been discussed as having etiological importance [65].

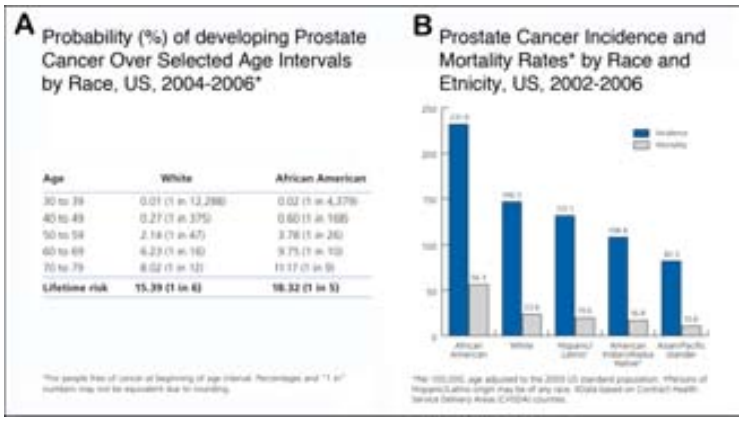



**Figure 6. Age, race and ethnicity as risk factors for PCa.** (A) Probability of developing PCa in US from 2004 to 2006 [66]. (B) Incidence and Mortality of PCa in US from 2002 to 2006 [67].

### 2.2.4 DIET

In large prospective cohort studies, dietary and nutritional factors have been suggested as ways of altering the risk of PCa, though the results have been inconsistent among the studies [49]. At present, the best dietary advice for reducing the risk of PCa is to eat a wide variety of fruits and vegetables each day, limit the intake of red meats, avoid the excessive consumption of dairy products (e.g. > 3 servings/day), maintain an active lifestyle, and consume foods that help maintain a healthy weight [68]. Moreover, since the development of the prostate gland is largely controlled by sex hormones and is initiated at week 12 of fetal development (see Box I), when the secretion of testosterone from the embryonic testis stimulates prostate morphogenesis, sex steroids, insulin-like growth factor (IGF) and vitamin D axes appear to influence the occurrence and progression of PCa. It also has been suggested that antioxidant carotenoid lycopene, resulting from a high intacke of tomatoes, have been associated with a reduced risk of PCa [69]. Other antioxidants, such as vitamin E and selenium, may also reduce risk for PCa [70].

## 2.3 CHARACTERISTIC FEATURES OF PROSTATE CANCER

The term, *carcinogenesis*, describes a concept whereby clinically hidden, multifocal pre-neoplastic foci emerge within the epithelium of an anatomic region exposed to the same carcinogen. Normal cells can progressively evolve to a neoplastic state. In order to do that, those cells need to acquire a succession of capabilities. These capabilities are known as the hallmarks of cancer and include the ability to sustain proliferative signaling, to evade growth suppressors, to resist cell death, to enable replicative immortality (including angiogenesis), and to activate invasion and metastasis [71]. This multistep process towards human tumor pathogenesis may be understood by examining the necessity for incipient cancer cells to acquire those qualities that enable them to become tumorigenic and finally malignant [72]. Tumors are more than masses of proliferating cells. Rather, they are complex tissues composed of multiple, distinct cell types that participate in heterotypic interactions with one another. The biology of tumors can no longer be understood by simply characterizing the traits of cancer cells; their descriptions must encompass the contributions of the *tumor microenvironment* to the process of tumorigenesis [72]. This is the reason that, recently, two new hallmarks have been included in the biology of cancer: the constitution and the signaling interactions of the tumor microenvironment [72].

PCa is generally regarded as multifocal, since its primary tumors often contain multiple, independent histological foci cancers that are often described as genetically distinct, even to those in close proximity [23]. Notably, about 80% of all radical prostatectomy specimens show

more than one neoplastic focus, as well as 70% of all HGPIN cases [24]. These findings suggest that multiple neoplastic foci may emerge and evolve independently, which has significant implications for the molecular mechanisms of disease progression.

The heterogeneity of PCa is potentially relevant to understanding the distinction between latent and clinical disease, as well as the strong correlation between PCa progression and aging [73]. Although PCa is a disease of older men, studies on healthy specimens obtained from healthy men in their 20s to 40s show the frequent presence of the histological foci of PCa [26], suggesting that cancer initiation has already taken place at a relatively early age. Combined with the evidence that PCa is multifocal, it appears that this organ may be the site of multiple, neoplastic transformation events, many of which give rise only to latent PCa, which does not progress to clinically detectable disease. It is thought that clinical PCa follows a different pathogenic program than that of latent PCa. Alternatively, most latent PCa foci may not undergo the critical activating events that lead to clinical disease, or it is possible that many may remain under active suppression [73]. These lesions may not be apparent at histological examination, though molecular techniques have found evidence of such changes in a variety of epithelial neoplasms, which are suggestive of carcinogenic changes (e.g., p53 loss, loss of heterozygosis (LOH) and microsatellite instability). In contrast, despite the phenotypic heterogeneity of metastatic PCa [74], molecular and cytogenetic analyses show that multiple metastases in the same patient are clonally related, which would indicate that advanced PCa is monoclonal [75].

Most prostate tumors are adenocarcinomas (Figure 7), which present a typical luminal phenotype [73] that shares numerous common features with other prevalent epithelial cancers, such as breast and colon cancers. In biopsy specimens, prostate adenocarcinomas present a typical disruption of the basal cell layer and can be confirmed by the absence of p63 and cytokeratin 5/14 immunostaining, both of which are basal cell markers [76, 77]. Additionally, its diagnosis is supported by AMACR immunostaining, a luminal marker overexpressed in prostate adenocarcinomas [76, 78]. While evident PCa subtypes are lacking at the histopathological level (see Box II for PCa subtypes), recent genomic analyses have provided increasing evidence for molecularly defined subtypes. Oncogenomic pathway analyses that integrate gene expression analysis, copy number alterations, and exon resequencing may provide a unified approach for distinguishing PCa subtypes and stratifying patient outcomes [79].

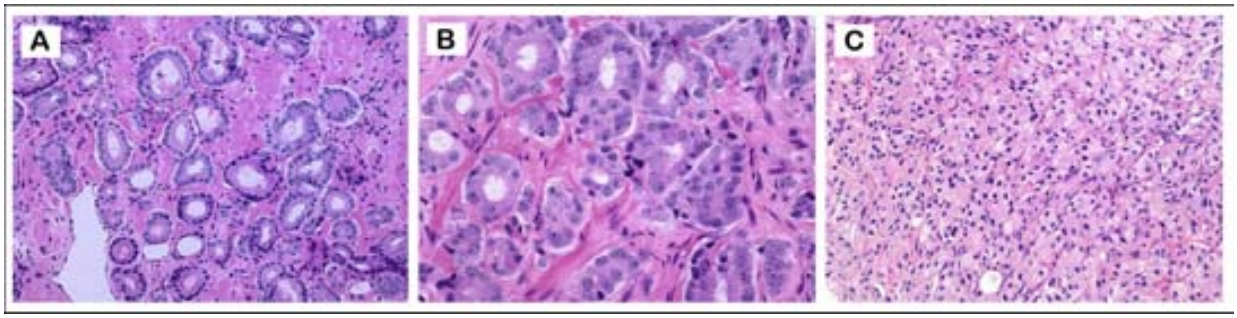



**Figure 7**. **Prostate Cancer features.** (A) Low grade adenocarcinoma where where the majority of the glands are relatively uniform in size. (B) Medium grade adenocarcinoma shows abundant amphophilic cytoplasm, enlarged nuclei with prominent nucleoli. (C) High grade adenocarcinoma shows fused glands, no intervening stroma and disruption of the basal cell layer. All images are have been obtained from web page; http://webpathology.com.

BOX II: PROSTATE CANCER SUBTYPES

*Histopathological subtypes:* The vast majority of PCa corresponds to acinar adenocarcinomas that express AR, while other types of PCa, such as ductal adenocarcinoma, mucinous carcinoma, and signet ring carcinoma, are extremely rare. Perhaps the most significant variant is neuroendocrine PCa, which is generally classified as either small cell carcinoma or carcinoid tumor and represents <2% of all PCa cases [80]. Neuroendocrine carcinoma has differentiation prevalence after recurrence, due to the lack of AR expression by neuroendocrine cells, which are inherently castration resistant. Other types are squamous cell, urothelial and sarcomatoid carcinomas, which commonly occur in association with acinar adenocarcinoma.

## 2.4 PROSTATE CANCER INITIATION AND PROGRESSION

The molecular pathways that contribute to the genesis of subclinical, microscopic PCa precursor lesions, their progression to invasive cancer, and their androgen-independence remain largely unknown, although certain molecular candidates have been implicated in the overall process of disease progression. For instance, in some prostate carcinomas aberrations in specific signaling molecules have been indicated, such as extracellular growth factors, protein tyrosine kinase cell surface receptors, intracellular anti-apoptotic or transcription factors, nuclear receptors and their ligands, growth suppressors, cell cycle regulators and others [81].

PCa lesions can develop in the complete disarrangement of both luminal and basal cells with concomitant loss of basal cell lamina. In its initial stages, when confined to the prostatic capsule, PCa is essentially curable by surgical intervention and/or radiation therapy. In fact, most cases of prostate carcinoma are relatively indolent, and the majority of men diagnosed with PCa will die of other causes instead. However, if not detected early or in the more aggressive forms of the disease, PCa can advance to stages that are characterized by a local

invasion of the seminal vesicles, followed by metastasis, primarily to the bone, which is usually lethal [4] (Figure 8). Consequently, a major clinical challenge is posed by the current inability to readily distinguish between indolent and aggressive tumors in PCa. The absence of this prognostic information may be addressed by a better understanding of the molecular basis of PCa [82].

As a starting point for the PCa progression pathway, it is important to consider chromosomal abnormalities as indicators for the first stages of PCa (Figure 8). Presumably, patterns of consistent allelic loss reflect the reduction or loss-of-function of putative tumor suppressor genes in PCa. Despite the significance of allelic loss to prostate carcinogenesis, no single candidate tumor suppressor gene has been definitively assigned a role in PCa progression [73].

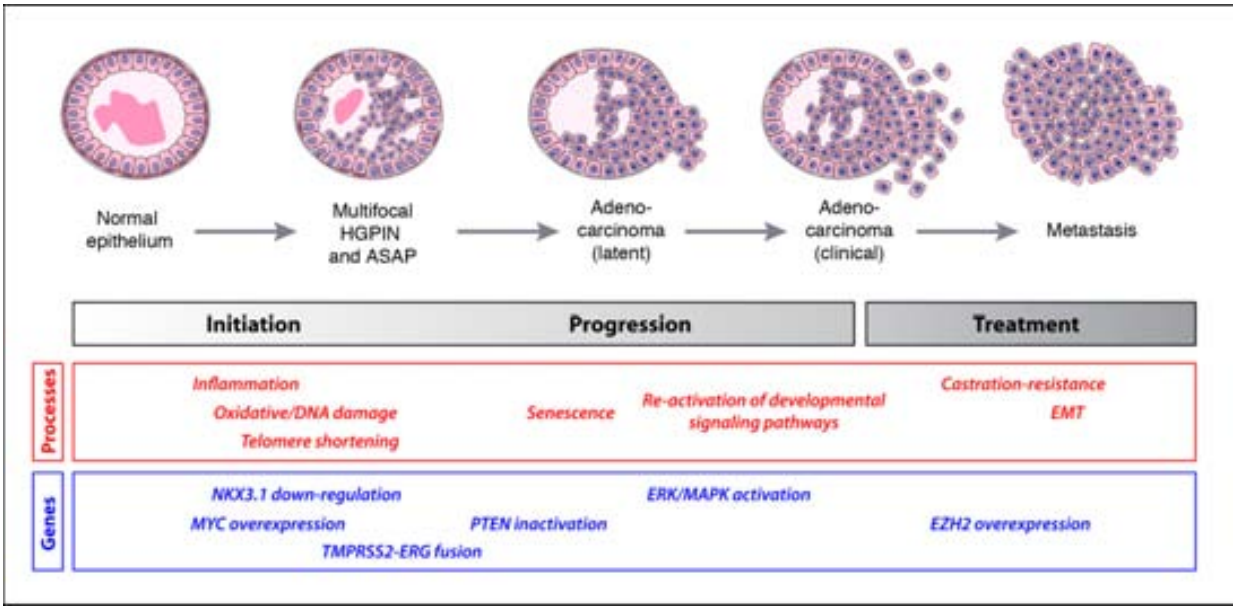



**Figure 8. Pathway for human PCa progression adapted from Abate-Shen *et al,* 2010.** Stages of progression are correlated with loss of specific chromosomal regions and candidate tumor suppressor genes [73].

## 2.4.1 PROSTATE STEM CELLS OR TUMOR INITIATING CELLS

Stem cells are defined by three important properties: the ability to self-renew, the ability to differentiate into a number of different cell types, and the ability to be easily grown from clonal populations [6]. A tissue stem cell can be defined as a progenitor that is multipotent, capable of giving rise to distinct cell types from the tissue of interest, and also able to self-renew by maintaining the stem cell phenotype in the progeny following cell division [83]. In the adult prostate, the existence of epithelial stem cells is suggested by this tissue's ability to undergo repeated cycles of extensive regression in response to androgen deprivation, followed by full regeneration after androgen restoration. In consequence, the prostate epithelium should contain a long-term resistant pool of stem cells that are castration resistant [84]. It is known that most luminal cells are androgen-dependent and undergo apoptosis

following castration, while most basal and neuroendocrine cells survive and are castration resistant [85, 86]. In theory, cancer could result from the transformation of a rare stem cell and/or the transformation of a more restricted cell type and its "dedifferentiation" in acquiring the self-renewal properties characteristic of stem cells. Given the luminal phenotype of human PCa, the cell origin should correspond to either a luminal cell or basal progenitor that can rapidly differentiate into luminal progeny after oncogenic transformation [73]. Several studies have been trying to find TICs from PCa. Despite promising findings in this area, it still remains unclear whether normal stem cells and cancer stem cells display conserved marker expression or whether the markers themselves display specificity for cancer stem cells. Nowadays, it seems that the authentic PCa stem cells have not yet been definitively identified [73].

### 2.4.2 PROCESS THAT PROMOTE PROSTATE CARCINOGENEIS

The hallmarks of cancer comprise eight biological capabilities acquired during the multistep development of human tumors. They include sustaining proliferative signaling, evading growth suppressors, resisting cell death, enabling replicative immortality, inducing angiogenesis, activating invasion and metastasis, reprogramming of energy metabolism and evading immune destruction [71, 72]. Underlying these hallmarks are genome instability, which generates the genetic diversity that expedites their acquisition, and inflammation, which fosters multiple hallmark functions.

PCa as well as other types of cancer, occurs when the rate of cell division is higher than the rate of cell death, leading to uncontrolled tumor growth. The causes of these events are still unknown, although some processes, which promote this carcinogenesis, are known. Androgens are regarded as major contributors to prostatic carcinogenesis (see Role of Androgens in PCa); nevertheless, only scant direct evidence demonstrates that androgens actually cause PCa.

As explained earlier, the single most significant risk factor for PCa is advanced age. However, PCa is not simply a product of aging, since the incidence among different populations varies considerably. In other words, the relationship of PCa to advanced age reflects the idea that the environment, as well as physiological and molecular influences play a role in PCa carcinogenesis and that, along with the normal consequences of aging, the effects of these influences accelerate the development and progression of this disease[73]. Moreover, various studies have described gene expression changes associated with aging and, in particular, changes involving the prostatic stroma, including genes involved in inflammation, oxidative stress and cellular senescence [87-89].

Chronic inflammation has been linked to PCa at the epidemiological, pathological and molecular levels [73]. The susceptibility of the prostate gland to infection is known from reported incidences of chronic bacterial prostatitis; thus, a potential role for bacterial infection has been suggested through the identification of multiple bacterial species in most prostatectomy specimens [90]. Moreover, some regions of inflammation present high proliferation. As we mentioned before, PIA has been proposed as a precursor to PCa [16, 91]. Regions that present PIA can be identified in men with advanced age and are frequently associated with inflammatory response; moreover, these regions are normally located in the proximity of HGPIN and adenocarcinoma lesions.

Oxidative stress results from the imbalance of reactive oxygen species (ROS) and leads to cumulative lipid, protein, and DNA damage. The prostate appears to be exceptionally vulnerable to oxidative stress, perhaps as a consequence of inflammation, hormonal deregulation, diet, and/or epigenetic modifications [73]. On the other hand, telomerase shortening, which is generally associated with DNA damage and may lead to chromosomal instability, has been implicated in PCa initiation [92]. Furthermore, cellular senescence, which corresponds to a form of cell cycle arrest where cells remains viable but are non-proliferative, has been identified as a potent mechanism of tumor suppression that prevents manifestation of the malignant phenotype after oncogenic insults [93]. Thus, oncogene-induced senescence may play an important role in preventing the progression of preneoplastic lesions to a fully malignant state. In the prostate, cellular senescence has been shown to occur during aging-related prostate enlargement and has been implicated as a tumor suppressor mechanism for prostate carcinogenesis. One possible interpretation of the temporal difference between the occurrence of latent PCa and the appearance of clinical state PCa is that cellular senescence may be involved in suppressing the progression to aggressive disease. It is possible that additional oncogenic events may be required to bypass the senescence mechanism in order to promote disease progression [73].

Recent analyses studies using massively parallel sequencing on tumoral and matched, normal genomic DNA have shown that genomic rearrangements could comprise a major mechanism that drives PCa carcinogenesis [94]. Moreover, prostate tumors display global changes in chromatin modifications coincident with cancer progression [95]. In addition, there are many somatic alterations, such as gains or losses in chromosomal regions. Importantly, several of these genetic alterations have also been identified in HGPIN, as well as in PIA lesions, supporting the idea of a precursor relationship between these two lesions and PCa. Epigenetic alterations are also believed to represent important contributing factors in prostate carcinogenesis and may provide useful biomarkers for disease progression [96, 97].

## 2.4.3 MOLECULAR GENETICS FOR PROSTATE CANCER

Subsequent to the initial transformation event, the further mutation of a multitude of genes, including p53 and retinoblastoma, can result in tumor progression and metastasis. The most commonly altered genes involved in PCa initiation, progression, and arrival to an advanced stage are shown in Table 2.

**Table 2. Altered genes in the initiation, progression and advanced Prostate Cancer**

| GENE | DESCRIPTION | FUNCTION | EXPRESSION | REFERENCES |
|---|---|---|---|---|
| **PROMOTING PROSTATE CARCINOGENESIS** | | | | |
| GSTP1 | Glutation s-transferase | Responsible for detoxification of reactive species | Epigenetically silenced in a majority of PCa by DNA methylation. Associated with inflammatory response in PIA lesions. | Nakayama *et al*, [98] |
| H3 | Histone 3 | Compact DNA into chromatin | Associated with prostate carcinogenesis. Epigenetically silenced by methylation. | Varambally *et al*, [99] |
| NKX3 | Homeodomine transcription factor | Transcription factor and Tumor suppressor controlling PCa | Early stages of PCa, where its inactivation may contribute to the vulnerability of oxidative stress. | Markowski *et al*, [100]; Abate-Shen *et al*, [101]; Chen *et al*, [102] |
| RNASEL | Endoribonuclease for ssRNA | Might play a central role in the regulation of mRNA | Muation - decreased activity. Associated with increased risk of sporadic cancer. | Casey *et al*, [103] |
| APE/Ref1 | Multifunctional enzyme | Involved in redox control of key enzyme and base excision repair | Up-regulated in PCa. APE polymorfisms are associated with increased PCa risk. | Kelley *et al*, [104] |
| PTEN | Lipid phosphatase | Tumor supressor gene | Copy number loss as an early event in PCa and is correlated with preogression to agressive, castration-resistant disease. | Z Chen *et al*, [105]; Berger *et al*, [94]; Abate-Shen *et al*, [4] |
| **INITIATION AND PROGRESSION** | | | | |
| MXI1 | Max-interacting protein 1 | Transcription repressor | Infrequently mutated. | Abate-Shen *et al*, [4] |
| Rb | Retinoblastoma | Key regulator of entry into cell division that acts as a tumor suppressor | Functional studies suggest critical role in PCa development. | Abate-Shen *et al,* [4] |
| p27 | Cyclin-dependent kinase inhibitor 1B | Important regulator of cell cycle progression | Loss expression in tumours correlated with tumour grade. | Abate-Shen *et al,* [4] |
| p16 | Cyclin-dependent kinase inhibitor 2A | Important regulator of cell cycle progression | Protein expression up-regulated in PCa. | Abate-Shen *et al,* [4] |
| Telomerase | Telomerase reverse transcriptase | Ribonucleoprotein enzyme; replication of chromosome termini | Reduced telomer length and increased telomerase activity found in HGPIN and PCa. | Sommerfeld *et al*, [106]; Zhang *et al,* [107] |
| Myc | Myc proto-oncogene protein | Transcription factor | Overexpression in PCa initiation (HGPIN) and abscence of amplification in the majority of advanced carcinomas. | Gurel *et al,* [108] |
| FGFs | Fibroblast Growth Factors | Stimulate growth of cells | Up-regulation of FGF; mechanism for activation of Erk MAPK pathway activity during cancer progression. | Foster *et al,* [109]; Djakiew *et al,* [110] |
| E-cadherin | E-cadherin | Cell adhesion | Reduced expression in HGPIN and PCa; loss may be associated with poor prognosis | Umbas *et al,* [111]; Morton *et al,* [112] |
| c-CAM | Carcinoembryonic - Cell adhesion molecule | Cell adhesion | Expression is reduced in HGPIN and lost in PCas | Kleinerman *et al,* [113] |
| c-MET | Met proto-oncogene | Tyrosine-kinase receptor | Overexpressed in HGPIN, PCa and metastasis | Pisters *et al,* [114] |
| Integrins | | Cell interaction | Reduced expression of specific family members during cancer progression | Cress *et al,* [115] |

| | | | | |
|---|---|---|---|---|
| FOXP3 | Forkhead box protein P3 | Transcription factor; Role in the control of immune response | Mutated in PCa. May repress Myc in PCa. | L Wang *et al,* [116] |
| TMPRSS2-ERG | Fussion gene | Chromosomal rearrangements linked with ETS family of transcriptor factors | Present in 15% of HGPIN and 50% of PCa. These rearrangement may occur after cancer initiation, or alternatively corresponds to an early event thar predisposes to clinical progression. | Abate-Shen *et al,* [73] and Berger *et al,* [94] |
| MAGI2 | Membrane guanylate kinase | PTEN-interacting protein | Rearrangement, loss of function. | Berger *et al,* [94] |
| Her2/Neu | Receptor tyrosine-protein kinase erbB-2 | Essential component of a neuregulin-receptor complex. | Activation through stimulation of AR signaling. Implicated in aggessive disease, progression to metastasis, and castration resistance. | Mellinghoff *et al,* [117] |
| EZH2 | Histone lysisne methyltransferase | Polycomb group (PcG) protein | Gene amplification. Up-regulated in advanced PCa and associated with aggressive tumors | Varambally *et al,* [99]; Saramaki *et al,* [118]; Bachmann *et al,* [119] |
| CHD1, CHD5, HDAC9 | | Chromatin modifiers | Regulation of embryonic stem cell plurypotency, gene regulation and tumor supression | Berger *et al,* [94]. Gaspar-Maia A *et al,* [120] |
| PRKC1 and DICER | | Micro RNA regulaors | Prostate tumorogenesis | Berger *et al,* [94] |
| MAP2K4 | Dual specificity mitogen-activated protein kinase kinase 4 | Kinase that activates the JUN kinases MAPK8 and MAPK9 as well as MAPK14 | Break-points rearrangement. Shown that induce anchorage-independent growth via muations | Berger *et al,* [94] |
| **ADVANCED CARCINOMA AND METASTASIS** | | | | |
| Androgen Receptor | Nuclear hormone receptor | Supress proliferation of basal cells, supports survival of luminal cells, and promotes metastasis | Expression maintained even in androgen-independent tumours, by gene amplification, gain-of-function mutations, alternative splice isoforms and finally enodgenous expression of androgen synthetic enzymes by tumor tissue. | Bentel and Tilley [121]; Niu *et al,* [122] |
| p53 | cellular tumor antigen 53 | Cell survival and proliferation, genome stability. Transcription/apoptotic regulator | Low mutation rate in primary cancer; frequently mutated in metastasis; p53 over-expression is correlated with poor prognosis. Frequently mutated in metastasis | Bookstein *et al,* [123]; Abate-Shen *et al,* [4] and Berger *et al,* [94] |
| IGF1 | Insulin-like growth factor I | Growth-promoting activity | Promotes growth of prostate epithelium; elevated serum levels is associated with cancer risk | Chan *et al,* [124] |
| TGFb1 | Transforming growth factor beta1 | Multifunctional protein that controls proliferation, differentiation and other functions in many cell types. | Negative regulator of prostate growth; shift to autocrine regulation associated with metastasis | Djakiew *et al,* [110] |
| EGF/TGFa | Epithellial growtrh factor /Transforming growth factor | Growth factors | Stimulates prostatic epithelial cell growth and invasiveness; may provude a mechanism for overcoming androgen-dependence | Djakiew *et al,* [110] |
| Ka1 | Putative integral membrane protein | | Shown to supress metastases, protein expression is down-regulated but is not mutated | Dong *et al,* [125] |
| Bcl2 | Apoptosis regulator Bcl-2 | Cell survival.Suppresses apoptosis in a variety of cell systems | Over-expression confers resistance to apoptosis in androgen-independent disease; key target for clinical intervention | Colombel *et al,* [126] |

## 2.4.4 CASTRATION-RESISTANT PROSTATE CANCER

Androgens regulate the prostate gland as the major stimulus for cell division in the prostate epithelium. Circulating androgens are essential to normal prostate development, as well as to the onset of PCa through their interactions with the AR (see Box III).

In the 1940's, Huggins *et al,* demonstrated that the removal of testicular androgens by surgical or chemical castration led to a reduction of the prostate tumor [127]. While the initial growth of a prostate tumor is dependent on androgens, the transition to metastatic disease is generally followed by androgen-independence, which is often evoked by androgen deprivation therapy (ADT). Following ADT, the androgen-dependence of prostate tissue is manifested by rapid cellular apoptosis and an involution to the regressed state. However, ADT is usually associated with PCa recurrence, thereby making continued ADT ineffective [128]. This recurrent disease has been called, *castration-resistant*. Unfortunately, castration-resistant PCa has remained essentially untreatable. Even when PCa progresses to castration-resistant PCa, AR activation and signaling remains sustained though a variety of mechanisms. Notably, castration-resistant tumors express AR, as well as AR target gens, such as PSA, indicating that the pathway activity has remained intact [129]. Nowadays, it is unclear when castration-resistant PCa normally appears within prostate tumors. The conventional adaptation model proposes that castration-resistant cells arise through the genetic/epigenetic conversion of previously androgen-dependent cells under conditions of androgen deprivation. The alternative, clonal selection model suggests that the emergence of castration resistance reflects the proliferation of a previously quiescent population of rare castration-resistant cells within an otherwise androgen-dependent tumor [130].

BOX III: THE ROLE OF ANDROGENS IN PROSTATE GROWTH

Growth of the prostate gland depends, like any cellular biology, on the balance between cell proliferation and cell death (apoptosis). If these two components are equivalent, as usually happens in normal prostate tissue, there is no increase in prostate growth. However, when the index of cell proliferation is greater than that of apoptosis, there is continuous growth of the prostate gland, and the number of cells increases. In the prostate, this balance between cell proliferation and apoptosis is regulated by androgens.

Testosterone, which is the major circulating androgen in peripheral blood, is produced mainly at the testicular level. At the prostate level, testosterone is converted into 5-alpha-dihydrotestosterone (DHT) through the action of 5-alpha-reductasa isoenzymes. Although both androgens are able to join the AR, DHT has a much higher affinity to the AR than testosterone [131-133]. The direct effect of testosterone on prostate epithelial cells is that it induces differentiation, while the indirect effect, proliferation, is mediated by the production of growth factors by the prostatic stroma. Androgens also directly stimulate the production of vascular endothelial growth factor (thus, inducing angiogenesis) in both normal prostate tissue and in neoplastic prostate tissue.

## 2.4.5 PROSTATE CANCER METASTASIS

Although common sites for secondary PCa metastasis are the lung, liver, and pleura, when PCa metastasizes, it goes first into the bone marrow stroma of the axial skeleton. This is the principal cause of PCa morbidity and mortality. Furthermore, PCa displays characteristic osteoblastic, rather than osteolytic, lesions [134]. Local invasion is a fundamental, initial step in the metastatic process, as without it, tumor spread could not occur. To develop invasive potential, malignant cells must down-regulate their cell–cell and cell–matrix adhesive characteristics, become motile and acquire the ability to break down extracellular matrices [135]. Once the malignant cells have escaped the tumor capsule, they must enter vascular or lymphatic circulation. Then, these cells must migrate via the circulation system to find a new place, where they can attach and proliferate and/or coalesce with other metastasized cells to form micro-metastases [136]. This can only happen when the environment at the secondary site is favorable. The epithelial to mesenchymal transition (EMT) process plays an important role in this progression (see Box IV).

A recent study showed that circulating tumor cells (CTCs) were able to be detected in the bone marrow of a significant number of patients with localized disease, suggesting that the disseminated tumor cells had not attained their full metastatic capability. This theory proves that CTCs from patients with metastatic disease show multiple chromosomal rearrangements typical of advanced PCa and consistent with the genomic instability acquired during cancer progression [137]. Generally, increased levels of CTCs predict a worse outcome in patients with metastatic, castration-resistant and clinically localized PCa [138, 139].

### BOX IV: EPITHELIAL TO MESENCHYMAL TRANSITION IN PCA

The formation of the prostate requires reciprocal interactions between the urogenital sinus mesenchyme and the epithelium and is dependent on testicular androgen synthesis [140]. It is known that an AR-dependent signal from the urogenital mesenchyme is required for prostate formation. While AR is not initially required in the urogenital epithelium for prostate organogenesis, it is subsequently necessary for epithelial differentiation and secretory protein expression. EMT is a transient, but biologically significant phenomenon that occurs during cancer progression and increases the innate invasive and metastatic capabilities of cancer cells. EMT is associated with histological, molecular and transcriptional changes. An interplay of several extracellular signals, growth factors, their effectors and transcription factors induce EMT and have the propensity of serving as EMT markers, consistent with tumor aggressiveness [141]. Recently, EMT has been mechanistically linked to stem cell signatures in PCa cells [141]. The acquisition of stem cell properties can lead to increased resistance to apoptosis, diminished senescence and escape from immune surveillance.

## 2.5 CLASSIFICATION OF PROSTATE CANCER

### 2.5.1. GLEASON GRADING

The Gleason grade was first described in 1966 by Donald F. Gleason [142]. More than 40 years after its introduction, the Gleason grading system still remains one of the most powerful prognostic predictors in PCa. However, the original descriptions of each pattern have undergone significant revisions over the years, first by Gleason *et al,* [143] and, most recently, at the 2005 International Society of Urological Pathology Consensus Conference [144]. To account for this heterogeneity, Gleason proposed a grading system that is now predominantly used by pathologists, since it is an excellent prognostic indicator. The Gleason score can be assessed using biopsy material or prostatectomies. It is the sum of the two most common patterns (grades 1 to 5) of tumor growth found in a sample. The Gleason score ranges between 2 and 10, with 2 being the least aggressive and 10 the most. In 2005, Amin et al recommended that the worst grade should be included, even if it is present in less than 5% of all cases [145] (Figure 9).

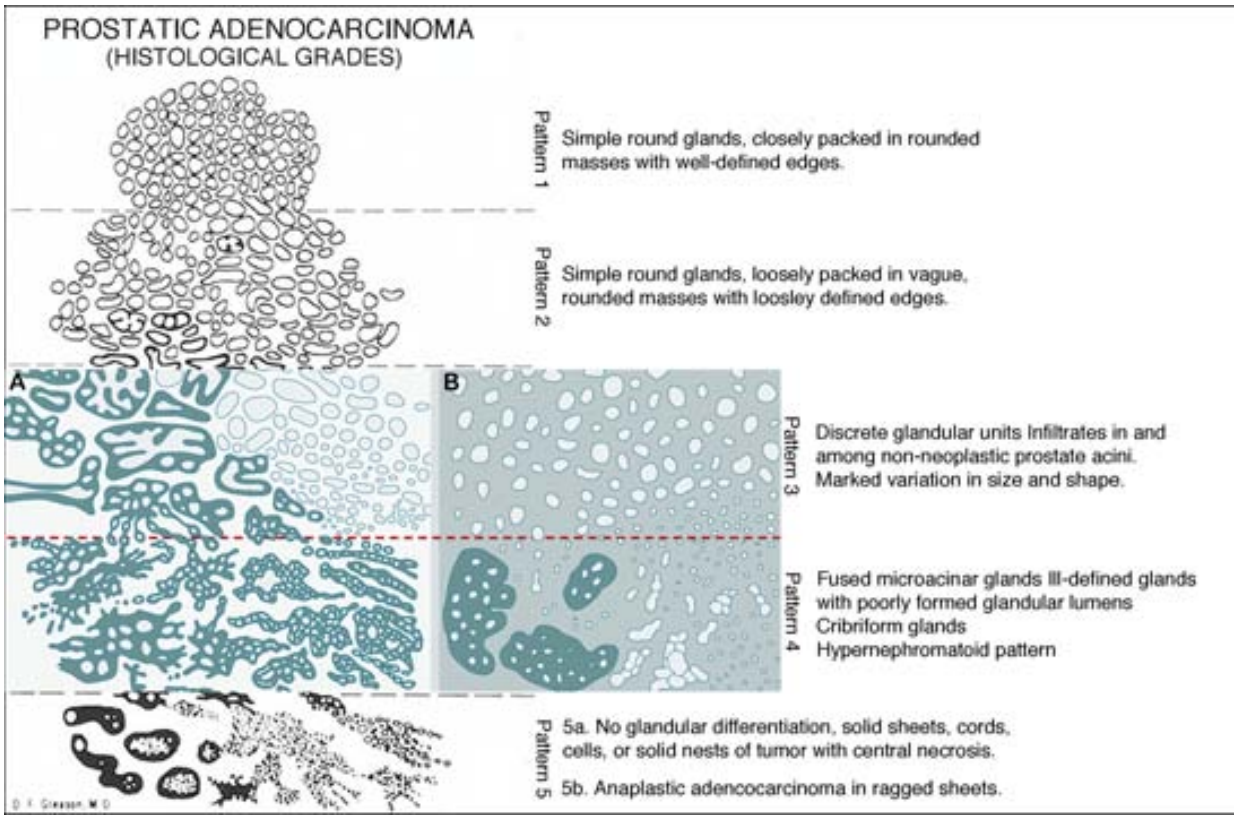



**Figure 9. Gleason grade.** This illustration shows Dr. Gleason's own simplified drawing of the five Gleason grades of prostate cancer. Grade 1 appears on the top and grade 5 on the bottom of the drawing (Adapted from Gleason *et al*, [142]). Color image adapted from Lotan *et al,* [146]; Comparison of the original Gleason and International Society of Urological Pathology (ISUP) 2005 modified systems for pattern 3 and pattern 4 carcinoma. (A) Gleason's original grading system and (B) ISUP 2005 modified system.

### 2.5.2 TUMOR NODE METASTASIS CLASSIFICATION

The tumoral node metastasis (TNM) classification is based on the status of the primary tumor, ranging from organ-confined to fully invasive (T1 to T4), with or without lymph node involvement (N0 or 1) and the presence and degree of distant metastasis (M0 and 1a-c) [147]. The TNM staging system provides a basis for survival prediction, initial treatment selection, patient stratification in clinical trials, accurate communication among healthcare providers, and a uniform method for reporting the end result of cancer management [148]. The 2009 TNM classification for PCa [149] is shown in Figure 10.

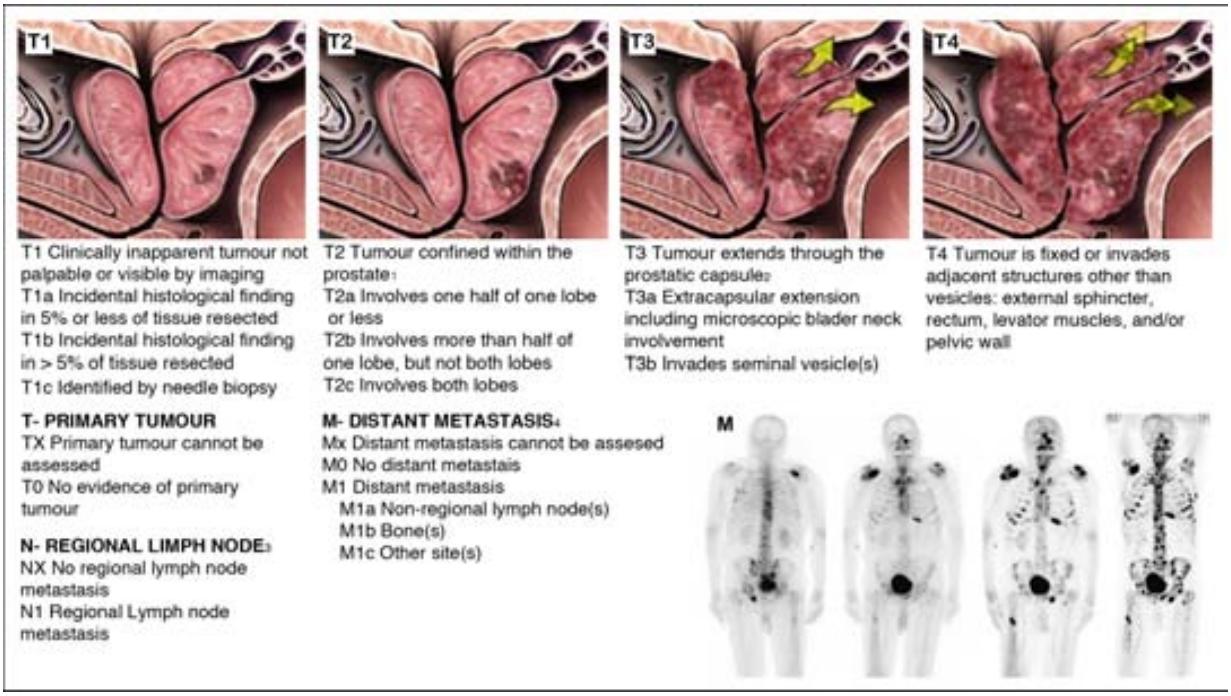



**Figure 10. Tumor Node Metastasis (TNM) classification of PCa**. Adapted from EAU guidelines 2010 edition [150]. (1) Tumor found in one or both lobes by needle biopsy, but not palpable or visible by imaging, is classified as T1c. (2) Invasion into the prostatic apex, or into (but not beyond) the prostate capsule, is not classified as pT3, but as pT2. (3) Metastasis no larger than 0.2 cm can be designated pN1. (4) When more than one site of metastasis is present; the most advanced category should be used.

## 2.6. PROSTATE CANCER TREATMENT

The therapeutic management of PCa has become increasingly complex, due to the various therapeutic options available, even in cases of clinically localized disease, which have equal oncological efficacy but with different, treatment-related side effects. Treatment recommendations vary by disease severity and life expectancy, since the side effects of treatment may outweigh the potential benefits for men whose cancers are unlikely to progress in their lifetime (Table 3). Additionally, a multidisciplinary approach may be advisable from the beginning in patients with high risk PCa, because it is very likely that adjuvant

treatment will be necessary for locally advanced disease. The main treatments for PCa from the Clinical Practice Guidelines in Oncology 2009 [151] are summarized below:

-*Watchful waiting (WW)*: This term, which was coined in the pre-PSA screening era, refers to the conservative management of PCa until the development of local or systemic progression, at which point the patient is afforded palliative treatment. The rationale behind WW is the observation that PCa often progresses slowly and is often diagnosed in older men for whom there is a high incidence of death from other disease.

-*Active surveillance (AS)*: AS is now an accepted management strategy for men with low-risk PCa who previously faced radical whole gland treatment (surgery, external beam radiotheraoy or brachytherapy (EBRT) [152]. AS involve monitoring the course of the disease with the expectation of intervening if and when the cancer progresses. It is often offered to men who have a limited life expectancy. Monitoring under AS involves PSA testing every 3 to 6 months, DREs every 6 to 12 months and possible, additional PBs.

-*Radical Prostatectomy (RP)*: This treatment involves the removal of the entire prostate gland between the urethra and the bladder and the resection of both seminal vesicles, along with sufficient surrounding tissue to obtain a negative margin. Regional lymph nodes may also be removed for examination to determine whether lymph node metastases are present.

-*Radiation therapy:* Radiation therapy normally consists of EBRT or brachytherapy for localized PCa. In EBRT the patient receives radiation treatment from an external source over an 8 to 9 week period. Brachytherapy involves placing small radioactive pellets, sometimes referred to as seeds, into the prostate tissue.

-*Hormonal therapy: ADT* alters the effects of male hormones on the prostate through medical or surgical castration (the elimination of the testicular function) and/or the administration of anti-androgen medications.

**Table 3. PCa treatment recommendations, by disease characteristics and life expectancy**

| Risk of progression and recurrence | Clinical characteristics of PCa | Life expectancy | Recommended initial treatment options |
|---|---|---|---|
| Low | T1-T2a and Gleason score 2-6, serum PSA levels <10 ng/mL | < 10 years | Active surveillance |
| | | > 10 years | Active surveillance or radical prostatectomy or radiation therapy |
| Intermediate | T2b-T2c, or Gleason score 7 or serum PSA level 10-20 ng/mL | < 10 years | Active surveillance or radical prostatectomy or radiation therapy (EBRT+/brachytherapy-) +/- ADT |
| | | > 10 years | Radical prostatectomy or radiation therapy (EBRT+/brachytherapy-) +/-ADT |
| High | T3a, or Gleason score 8-10 or serum PSA level >20 ng/mL | All | Radical prostatectomy (selected patients) or radiation therapy (EBRT) + long-term ADT |

Adapted from Prostate cancer NCCN Clinical Practice guidelines in Oncology 2009 [151].

# 3. PROSTATE CANCER DIAGNOSIS

## 3.1 PROSTATE CANCER SCREENING AND DIAGNOSIS

In order to cure patients with PCa successfully, it is important to detect the disease at an early stage, as well as monitor its progress accurately. PCa is typically a slow growing tumor that affects older men. Despite its slow growth, PCa is still a lethal disease. Early PCa usually has no symptoms. With more advanced disease, patients may present symptoms related to urethral obstruction (urinary frequency, hematuria, difficulty in initiating urination, or dysuria). Nevertheless, these symptoms can occur as the result of non-cancerous conditions, such as BPH. Normally, most cases of PCa are diagnosed before these symptoms develop. Due to the long latency period of this cancer and its potential curability, this disease is an excellent candidate for screening strategies that attempt to identify the disease in its early, curative state [153]. The diagnostic tools for detecting PCa can be separated into those that screen for the disease, such as PSA and DRE, and the decisive diagnosis set of transrectal ultrasound guided prostate biopsies (TRUS) [154].

PCa screening is widely utilized and considered to be an effective detection method for PCa. One of the limitations of serum PSA as a tumor marker is its lack of specificity (around 30%), which results in a high rate of negative biopsies. Elevated PSA levels can also be attributed to other factors such as BPH, prostatitis, prostate irritation, and recent ejaculation [155, 156]. As a consequence of the current screening parameters, around 2/3 of the approximately 1,300,000 biopsies made yearly in the United States and 390,000 in Europe are unnecessary [157, 158]. In contrast, the false positive rate of a biopsy is about zero, although the false negative rate in the first biopsy may oscillate between 12% and 32% [159]. As a result of their persistently elevated PSA levels, but negative biopsy results, these men undergo repeated biopsies to rule out PCa. This situation is called the "PSA dilemma" [155, 156] (See section 3.2).

### 3.1.1 PROSTATE SERUM ANTIGEN

Human kallikrein-related peptidase 3 (KLK3) is commonly referred to as PSA. The first report on the detection of PSA in serum was made by Papsidero *et al*, [160]. PSA is used as a serum biomarker to monitor and screen for PCa since 1986 and 1994, respectively [161]. Use of PSA-based screening has become widespread as a cancer marker. PSA has led to increased PCa detection and has served as an alert to stage migration with a decreased number of metastatic or locally advanced cases of cancer at diagnosis [162] (Figure 4a). The elevation of serum PSA signals an abnormality in the prostate, whether it is a benign enlargement, an inflammation or a PCa. This ability to find early-stage cases of PCa has made PSA an

interesting biomarker for use in PCa screening. However, as many men die with, but not as a result of PCa, a controversy exists regarding the ability of PSA to save lives [153].

There is continuing disagreement over the threshold levels of PSA, which would indicate the need for PB. A recommendation for biopsy has been set at an arbitrary serum PSA level of 4 ng/mL. At this threshold, PSA displays 93% sensitivity and a poor specificity of 24% for the detection of PCa [163]. In addition to PCa, PSA is expressed in normal prostate tissue, prostatitis, and BPH [164]. Furthermore, 27% of the men with borderline serum PSA levels (3.1-4ng/mL) have detectable PCa by biopsy [165]. The PSA levels are continuous parameters: the higher the value, the more likely the existence of PCa (Table 4). PSA testing cannot be expected to resolve all ambiguity with respect to PCa; instead, it may be best considered as an indicator of risk to be weighed in combination with other factors. Normally, annual PSA screening is recommended for all men over 50 years old, however, in patients with familiar PCa a screening should start at age of 40.

BOX V: THE ROLE OF PSA IN THE PROSTATE

PSA (or KLK3) is a glycoprotein belonging to the family of kallikrein-related serine proteases that are produced in normal prostate secretion. Its physiological role is believed to be the liquefaction of seminal fluid. The transcription of PSA is governed by androgens, which restrict its high-level production to the prostate epithelium. PSA is synthesized in healthy prostate tissue, in BPH and in PCa at all grades and stages [166]. For this reason, PSA is defined as being organ-specific, but not cancer-specific. The normal prostate architecture keeps PSA confined to the organ, so that only a minimal amount leaks into the circulatory system. Increased blood levels of PSA in men with cancer or with other prostate disease conditions cannot be explained by increased PSA expression. In fact, during the development and progression of PCa, PSA expression may actually decrease slightly [167]. So, the increased blood PSA levels must be caused, instead, by an increased release of PSA into the blood. Even so, there are no experimental data or mechanisms of increased release that are believed to result from the disruption of the prostate architecture seen in prostate tumors, such as the disordering of the basement membrane and a loss of the basal cell layer, ductal lumen architecture and epithelial cell polarity [161] (Figure 11).

PSA strongly discriminates between the different cancer stages. It is higher in men with localized disease than in cancer-free controls, is associated with stage and grade in localized disease and is higher in patients with metastatic compared to localized disease (Figure 11). Moreover, men with high PSA levels at the initial time of therapy have an increased risk of recurrence. PSA is a sensitive indicator of recurrence after RP, but it is far less sensitive as an indicator of recurrence after radiation therapy [161].

**Table 4: Risk of PCa in relation with PSA values** [150]

| PSA level (ng/mL) | Risk of PCa |
|---|---|
| 0 – 0.5 | 6.6 % |
| 0.6 – 1 | 10.1 % |
| 1.1 – 2 | 17.0 % |
| 2.1 – 3 | 23.9 % |
| 3.1 – 4 | 26.9 % |
| 4 – 10 | 20 – 35 % |
| > 10 | 50 – 80 % |
| > 20 | 95% |

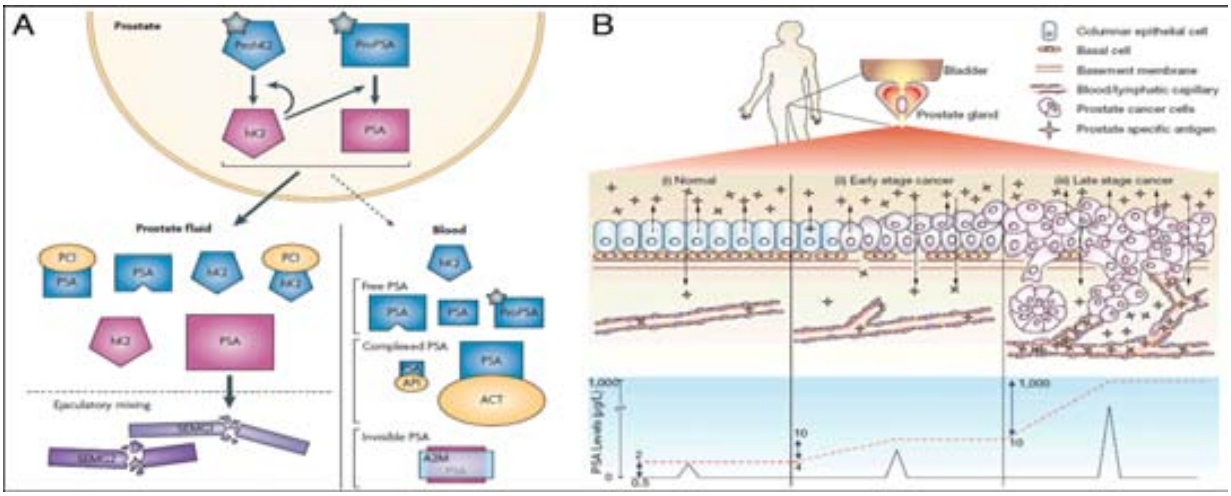



**Figure 11. PSA isoforms and PSA behaviour.** (A) PSA subforms and interaction, adapted from Lilja et al 2008 [161]. (B) Adapted from Kulasingam V and Diamandis EP 2008 [168]; Destruction of prostate tissue architecture as a measure of PSA elevation. (i) Healthy men (serum PSA 0.5-2 ng/mL) low levels of PSA enter the circulation via diffusion. (ii) Early-stage PCa (serum PSA 4-10 ng/mL) result of destruction of tissue architecture. (iii) Late-stage PCa (serum PSA 10 to 1,000 ng/mL), invasion of tumor cells, considerable amounts of PSA leak into the bloodstream.

Various attempts have been made to overcome the limitations of PSA screening, including the use of PSA variants such as, ratio PSA, age-adjusted PSA, fPSA, PSA density (PSAD), PSAV, PSA- doubling time (PSADT) and PSA isoforms (proPSA, cPSA, BPSA and intact PSA). All have been proposed as a means of improving serum PSA specificity in the detection of PCa. Nevertheless, there has been no evidence to suggest that any of these testing strategies improves health outcomes [169].

PSAD is the ratio of serum PSA to prostate volume (measured by TURS). PCa causes the release of more PSA into circulation per unit than does BPH [170]. Hence, PSAD might help to distinguish between BPH and PCa. In addition, a low PSAD has also been associated with tumor aggressiveness and unfavorable pathological features. Therefore, PSAD requires a

transrectal ultrasound, which is expensive, time-consuming and causes significant patient discomfort.

FPSA is the measure of non-complex PSA in serum, which is unreactive with plasma protease inhibitors. In proportion to total PSA, fPSA is lower in men with PCa than in men with BPH [171]. The free/total PSA ratio (expressed as the percentage of fPSA, %fPSA) is the concept most extensively investigated and most widely used in clinical practice to discriminate between BPH and PCa. This ratio is used to stratify the risk of PCa for men who have total PSA levels between 4 and 10ng/mL and negative DREs [150].

In general, PSA dynamics, such as PSAV and PSADT, add little to the predictive power of total PSA alone. PSAV is defined as the change in PSA levels over a specific period of time (ng/mL/year). PSAV is strongly associated with the diagnosis of PCa [172] and with a risk of recurrence and/or cancer-specific death after treatment [173] [174]. Finally, PSADT, which measures the exponential increases of serum PSA over time, reflects a relative change [175]. PSADT is mostly used to monitor disease progression for patients after surgery and radiotherapy and for patients who choose surveillance rather than definitive treatment [176]. However, prospective studies have shown that these measurements do not provide additional information compared to PSA alone [177-180].

The following is a list of PSA isoforms: ProPSA, intact PSA, nicked PSA and BPSA. PSA is initially produced as a 261 amino-acid (aa) prepotein. The cotranslational removal of an aa-terminal leader generates a non-catalytic zymogen (ProPSA). The subsequent removal of the 7-residue propeptide generates the catalytically active mature form (237-residue single-chain enzyme containing 5 intra-chain disulphide bonds). FPSA in the blood is a mixture of mature PSA and proPSA. Several studies have suggested that proPSA could help in discriminating PCa from benign disease [181]. FPSA in serum can be divided into intact forms (mature and proPSA) and nicked forms. The level of intact PSA and the ratio of nicked PSA to total PSA have shown potential for improving the ability to discriminate between PCa and BPH [182]. A second, distinct, cleaved form BPSA has been associated with prostate volume and, therefore, could also be helpful in distinguishing PCa from BPH [183]. See Figure 11.

Apart from screening, PSA is also used as an indicator of disease recurrence. Monitoring PSA after definitive local therapy for PCa provides a means of detecting recurrent PCa long before the tumor is detectable by any other means. Increasing PSA levels after treatment generally indicate the presence of tumor cells. However, PCa can also recur in the absence of a PSA increase or detectable PSA levels [161]. In contrast, detectable PSA levels after prostatectomy provide evidence of treatment failure, though its presence mostly defines the biochemical recurrence with PSA levels > 0.5ng/mL. Even so, some patients develop detectable PSA levels

after RP, but have no further increases in PSA [184]. PSA levels are also useful after radiation therapy, even though those levels slowly decrease. The time it takes to reach PSA nadir can range from months to years after treatment. Moreover, this time has been associated with radiation dosage [185], size of the prostate and pretreatment PSA levels [186]. Finally, PSA has been used to help control the standard ADT-based treatment for PCa. After initiation of ADT, PSA in the blood almost always decreases and then stabilizes for varying intervals [187]. ADT causes an initial decrease in PSA, not only because of tumor regression, but also because ADT suppresses the transcription of the PSA gene, which is androgen-dependent [188]. The progression to androgen-independent disease is generally defined by two consecutive increases of PSA after the post-ADT PSA nadir [189]. These increases predict tumor progression. The androgen-independent PSA increase not only indicates renewed tumor growth, but also the reactivation of androgen despite the castrated levels of testosterone. This reactivation (by mutation, gene duplication, etc.) is a common feature of androgen-independent PCa [190].

PSA can also help in monitoring advanced-stage cancer. In general, PSA is only modestly associated with survival in men with such advanced stages of disease. In addition, post-treatment changes in PSA may not accurately indicate the presence or absence of a response to treatment. The value of PSA as a surrogate measure in clinical trials has been questioned. For this reason, there is a great need to discover more accurate markers of response [161].

### 3.1.2 DIGITAL RECTAL EXAMINATION

The DRE, Latin *palpatio per anum* or PPA, is a relatively simple procedure. DRE is performed on all patients in whom PCa is suspected (by elevated levels of serum PSA). The patient is placed in a position where the anus is accessible. The physician usually examines the external area (anus and perineum) for any abnormalities, such as hemorrhoids, lumps, or rashes. Then, as the patient strains down, the physician slips a gloved and lubricated finger into the rectum through the anus and palpates the inside for approximately sixty seconds.

Most cases of PCa are located in the peripheral zone of the prostate and may be detected by DRE when the volume of this cancer is ≤ 0.2mL. A suspicious DRE is an absolute indicator for PB. However, the probability that an abnormal DRE is highly suggestive of PCa depends significantly on the PSA values [191]. Often, a palpable cancer is already advanced in both grade and stage and is potentially no longer organ confined. In experienced hands, DRE has a specificity of 83.6% and a sensitivity of 53.2% and an abnormality in either PSA or DRE alone confers a 20-25% chance of PCa [192]. The overall positive predicitive value (PPV) for cancer

detection increases to 50% when DRE and PSA are used in combination [193]. In approximately 18% of the patients that presented PCa, the disease was detected thanks to DRE alone, irrespective of the PSA levels [194].

### 3.1.3 PROSTATE BIOPSY

The need for PBs should be determined on the basis of the PSA levels and/or a suspicious DRE. TRUS is currently considered the normal standard of care. Although a transrectal approach is used for most PBs, some urologists prefer to use a perianal approach. Sextant biopsy is no longer considered adequate, as different studies have shown that increasing the number of biopsied cores increases PCa detection rates [195]. The Vienna nomogram offers an easy tool for selecting the optimal number of biopsy cores, based on age and total prostate volume in the PSA range of 2-10ng/mL. This system improved the cancer detection rate 66.4% compared to the old, systematic sextant biopsies [196]. Indications for repeat biopsy are rising and/or persistent PSA levels, a suspicious DRE and ASAP [150]. HGPIN, as an isolated finding, is no longer considered an indication for repeat biopsy. However, in the event of extensive HGPIN (multiple biopsy sites), repeat biopsy could be prescribed, since the risk of subsequent PCa is slightly increased [197].

Thanks to PBs, clinicians are able to obtain a representation of the prostate tissue. Following this procedure, the histopathological grading of PCa can be performed using the Gleason scoring guidelines (Figure 9). TRUS-guided biopsy have an overall sensitivity and specificity for PCa detection of 32% and 51 [198]. The histopathology of prostate tissue can be definitively identified in most cases of PCa. This method is the most commonly used prognostic indicator for PCa and results in a Gleason score. The lower the Gleason grade, the better the prognostic outcome [144].

However, there are limitations to this diagnostic method. First of all, a biopsy must be performed in order to obtain the cancer tissue for testing, and this procedure is considered to be invasive procedure that usually requires sedation of the patient. Second, needle biopsies only provide a small amount of tissue that is later available for histological examination, and problems arise because these biopsies often identify only a few malignant glands among many benign glands [44]. Finally, there is also disagreement regarding the thresholds of scoring afforded by different pathologists, since the Gleason grading scale that is used by them is semi-quantitative [199]. For these reasons, the Gleason scores themselves have limited quantitative value. The use of biomarkers will overcome the problem of quantification, and thus, they can provide a more accurate method for early diagnosis [200].

## 3.2 PROSTATE CANCER AND THE DIAGNOSTIC DILEMMA

As explained above, current screening techniques are based on the measurement of serum PSA levels and the DRE. However, the use of serum PSA as a cancer-specific diagnostic test has well-recognized limitations. It not only has a low PPV, which results in aprox. 60% negative biopsies, it also often leads to repeat PSA measurements and biopsies (the "over-diagnosis" problem). It also can lead to "over-treatment" by detecting non-life-threatening tumors [201], especially in the so-called gray zone (serum PSA between 4-10ng/mL), which represents a dilemma for discriminating between patients with PCa and BPH patients, or discriminating between prostatitis and urethral manipulations, which can also increase PSA levels [202]. Conversely, a significant number of diagnosed carcinomas present PSA levels below the threshold of 4ng/mL, resulting in undiagnosed cases of the disease[203]. See figure 12.

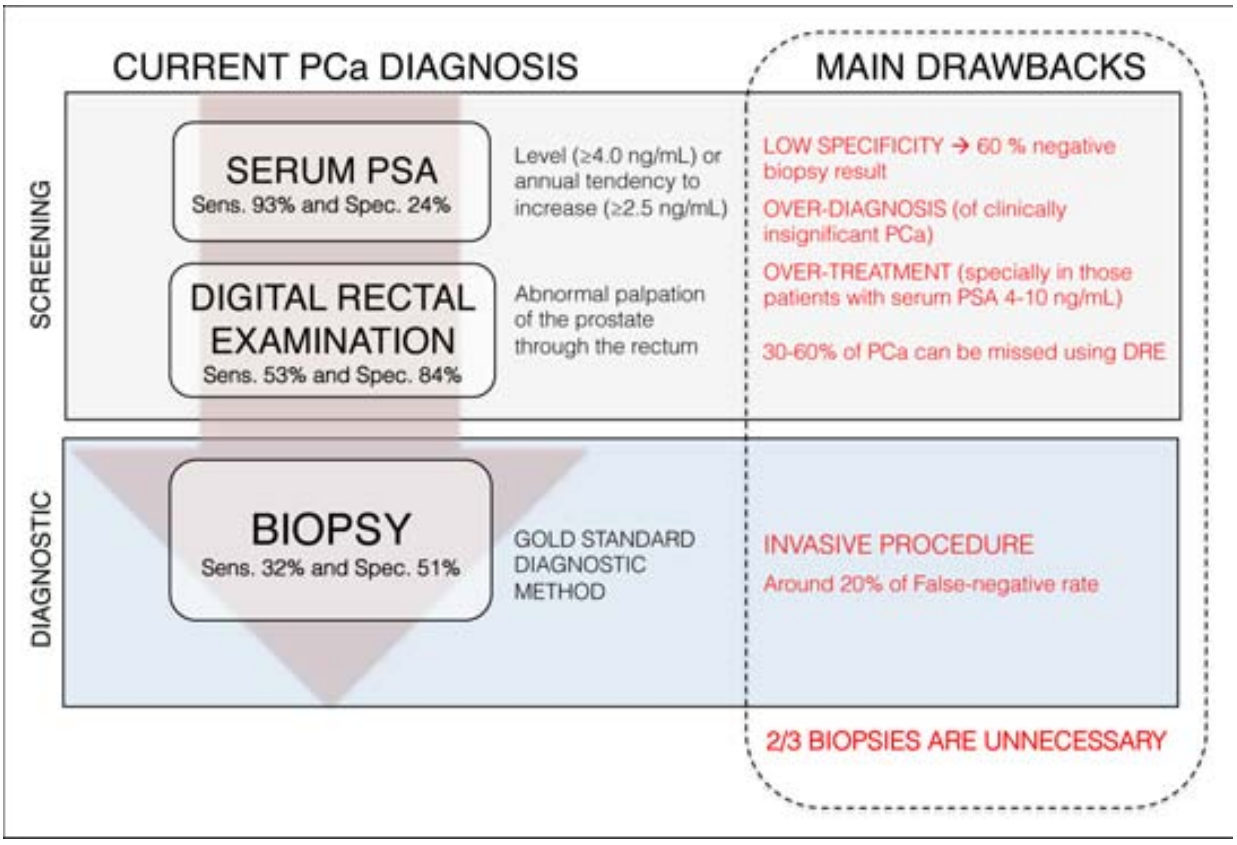



**Figure 12. Current squeme of diagnostic for PCa.**

Recent recommendations now suggest later and less frequent PSA screenings. These changes present a major clinical challenge for PCa diagnosis and treatment [204]. These new recommendations were proposed, because the widespread use of PSA testing has led to a vast increase in the diagnosis of patients having clinically localized, low Gleason grade carcinomas that may not require treatment, since their tumors are relatively indolent. Patients with a score of Gleason 3 or less almost never relapse after local therapy and can

very likely be managed conservatively with WW. Nonetheless, a small fraction of these tumors will progress rapidly and require immediate treatment [205]. This problem was first defined by Telesca et al. as the "over-diagnosis" problem. They defined over-diagnosed cases as those that would not have been detected by clinical means within the man's lifetime [206]. The over-diagnosis of clinically insignificant cases of PCa is considered one of the major potential drawbacks of PSA screening, especially in older patients. Approximately 70% of the men who died from unrelated causes also had demonstrated cases of PCa [30]. Currently, there is no existing method that allows clinicians to distinguish between insignificant cases and potentially lethal cancers, and, therefore, to effectively determine which men require treatment and which require WW. This problem could be addressed by the identification of new biomarkers able to distinguish between PCa and other prostate pathologies. In an attempt to predict the presence of these types of non-significant PCa cases before definitive treatment and to try to solve this "over-diagnosis" problem, the Epstein Criteria were developed [207] (Table 5).

Subsequently, another clinical challenge is posed by our current inability to readily distinguish between indolent and aggressive tumors in PCa patients, who present with low Gleason grade tumors upon biopsy [208]. The absence of this prognostic information has led to a significant "over-treatment" of patients, who would otherwise require only conservative management. D'Amico and coll. describe the *D'Amico Criteria,* which combine the examination of tissue from the prostate after radical prostatectomy, in combination with serum PSA values, in order to group patients into different risk groups of low, intermediate and high risk. The reasoning for this grouping is that only cases of clinical PCa require immediate treatment. Active surveillance can be an alternative option for men with low risk PCa [209] (Table 5). This prognostic challenge could effectively be addressed by a better understanding of the molecular basis of cancer initiation, which would ultimately lead to the identification of biomarkers that can distinguish between indolent and aggressive forms of PCa.

**Table 5: D'Amico and Epstein stratification critarias**

| Criteria | Groups | Description |
|---|---|---|
| D'AMICO | Low Risk | cT1-2a and Gleason score 2-6 and PSA < 10 ng / mL |
| | Intermidiate Risk | Intermediate risk-co cT2b Gleason score 7 or PSA 10-20 ng / mL |
| | High Risk | cT3a or more, or Gleason score 8-10 or PSA > 20 ng / mL |
| EPSTEIN | Insignificant PCa | 0.5cm$^3$ or less and confined to the prostate, Gleason pattern 4 to 5, PSAD: 0.1 to 0.15, low-grade cancer or intermidiate and core involvement of only 1 needle core biospy < 3mm. |
| | Unfavorable PCa | Gleason sum. 7-10 and/or non-organ-confined disease |

Recently, two large, randomized trials of PCa screening with PSA testing have been completed. The US-based Prostate, Lung, Colorectal and Ovarian Cancer (PLCO) screening Trial did not observe a mortality benefit from screening [210], while the European Randomized Study of Screening for PCa (ERSPC) concluded that PSA-based screening reduced the rate of death from PCa by 20%. However, this type of screening was also associated with a high risk of over-diagnosis [211]. Moreover, the ERSPC study reported that after two rounds of screening, 43% of the men who underwent RP had cancers that were considered minimal [211]. The major Urological Societies have concluded that, at present, widespread mass screening for PCa is not appropriate. Rather, early diagnosis should be offered to the well-informed man. Nevertheless, two key questions remain open and empirical: at what age should early detection start, and what is the interval for PSA and DRE? A baseline PSA determination at age 40 has been suggested, upon which subsequent screenings may then be based. A screening interval of 8 years might be enough in men with PSA levels ≤ 1ng/mL. Further PSA testing is not necessary in men older than 75 years of age with a baseline PSA ≤ 3ng/mL, because of their very low risk of dying from PCa [150].

As mentioned earlier, PCa screening is not the only diagnostic technique to show several limitations; the PB, which is currently the preferred diagnostic technique, also has its limitations. In summary, in order to cure patients with PCa successfully, it is important to detect the disease at an early stage, as well as monitor its progression accurately. In order to successfully accomplish this, it is necessary to find new biomarkers that can distinguish between patients with malignant prostate pathologies and those with benign conditions (early diagnostic markers). Moreover, it would also be very helpful to find new biomarkers that can distinguish between cases of low Gleason grade PCa and those of high Gleason grade (progression markers).

For all of the reasons listed above, the search for new biomarkers, which can help to improve PCa screening, represents a major step in successfully curing patients suffering from this disease. Moreover, it is not only important to find new diagnostic markers, but also to look for prognostic markers that can help stratify patients and to treat them effectively.

# 4. BIOMARKERS FOR PROSTATE CANCER

## 4.1 BIOMARKER DEFINITION

The National Cancer Institute defines a biomarker as "a biological molecule found in blood, or other body fluids, or tissues that is a sign of normal or abnormal process or of a condition or disease." A biomarker should be objectively measured and evaluated as an indicator of a stage or process, disease progression or pharmacological responses to a particular therapeutic intervention [212].

The discovery of biomarkers is based on the following research principle: the comparison of physiological states, phenotypes or changes across control and case (disease) patient groups [213]. A key approach to biomarker discovery research is to compare cases versus control samples in order to detect statistical differences, which could lead to the identification and prioritization of potential biomarkers. Theoretically, every disease may be detected and characterized by its unique biomarker. However, viewing this biomarker as a single molecule is just one option. A more complete view of a biomarker is a panel of up- and down- regulated molecules and/or proteins with altered post-translational modifications (PTMs) that differ in their diseased and normal states [214, 215]. Biomarker classification can be accomplished using a different set of parameters; however, for our study we needed to focus our biomarker classification system on the basis of their potential applications for screening, diagnosis, or prognosis (see Box VI).

BOX VI: TYPES OF BIOMARKERS ON THE BASIS OF THEIR APPLICATIONS

*Screening biomarkers* are used to predict the potential occurrence of a disease in asymptomatic patients.

*Diagnostic biomarkers* are used to make predictions on patients suspected of having a disease. An ideal diagnostic biomarker should enable unbiased determination, particularly in patients without specific symptoms. It should fulfill several criteria: (1) high specificity for a given disease (low rate of false positives), (2) high sensitivity (low rate of false negatives), (3) ease of use (rapid procedure), (4) standardization (consistent reproducibility), (5) clearly readable results for clinicians [215], (6) economical, and (7) quantifiable in an accessible biological fluid or sample.

*Prognostic biomarkers* are applied to predict the outcome of a patient suffering from a disease.

One of the key issues in biomarker research is the type and accessibility of a high quality biological matrix. Blood is most commonly used, though other biofluids, such urine, may provide valuable information. However, these biofluids are not routinely used, and the

availability of samples is often restricted. Preanalytical factors such as biospecimen collection, processing, storage and distribution also play important roles in biomarker research.

The identification of a new candidate biomarker is followed by a thorough operational evaluation, in order to validate its applications within clinical setting (see Box VII). The conventional biomarker development pipeline involves a discovery stage, followed by a validation stage on a larger cohort of samples prior to clinical implementation. Despite the intensified interest and investment, only a few novel biomarkers have been successfully introduced into clinical practice [216]. The reasons for this failure are the long and difficult path from candidate discovery to clinical assay and the lack of coherent and comprehensive pipelines for biomarker development [217].

The new biomarker development pipeline, which is an improvement over the previous one in terms of its success, is defined by different essential process components [217] (Figure 13). It starts with the discovery stage, which is the initial step for semi-quantitatively analyzing the molecular content in common biospecimens and selecting and prioritizing disease-related biomarkers [214]. Discovery can employ different model systems or a variety of human samples, such as body fluids, and usually comprises a binary comparison between the diseased and controlled states, avoiding contamination by other disease conditions [217]. The product of the discovery phase is a list of candidate biomarkers with a high false discovery rate of differentially expressed proteins, at least for lower abundant proteins [217]. This "candidate biomarker list" can be supplemented by candidates drawn from the literature, from alternative discovery approaches, or from domain knowledge. The next step, "qualification," is used to confirm that the differential candidate can be seen using an alternative target method and to confirm differential expression in simplified comparisons between the diseases and normal states [217]. In the verification stage, the analysis is extended to a larger number of samples. Verification is defined as the bridge between discovery and quantification and is the process of credentialing prioritized biomarker candidates using analytically robust, reproducible and quantitative assays on -a statistically- powered number of samples having clinical relevance. Credentialed biomarkers that successfully pass this stage are considered verified biomarkers, which are of high value for translation into large-scale, clinical validation studies [214] (see Box VII).

BOX VII: VALIDATION OF BIOMARKERS

Biomarkers that have been scientifically scrutinized must pass through several proposed practical tests prior to being accepted into clinical practice. Several phases for biomarker development have been suggested [218] [219]: (1) preclinical exploratory, (2) clinical assay and validation, (3) retrospective longitudinal, (4) prospective screening, and (5) cancer control.

The first phase involves identifying the biomarker and evaluating it for a specific clinical indication. Ideally, biomarkers should be validated analogously in prospective, well-controlled clinical studies on diverse patients across multiple institutions with well-established standards for all of the steps in the process [148]. However, those ideal conditions are rarely applied.

Prior to submitting a biomarker to the US Food and Drug Administration (FDA) or EU European Medicines Agency (EMA) for approval, analytical and clinical confirmation needs to be performed. After approval by the FDA, the Center for Medicaid and Medicare service (CMS) have to conclude that this biomarker is essential for the patient care (Figure 13).

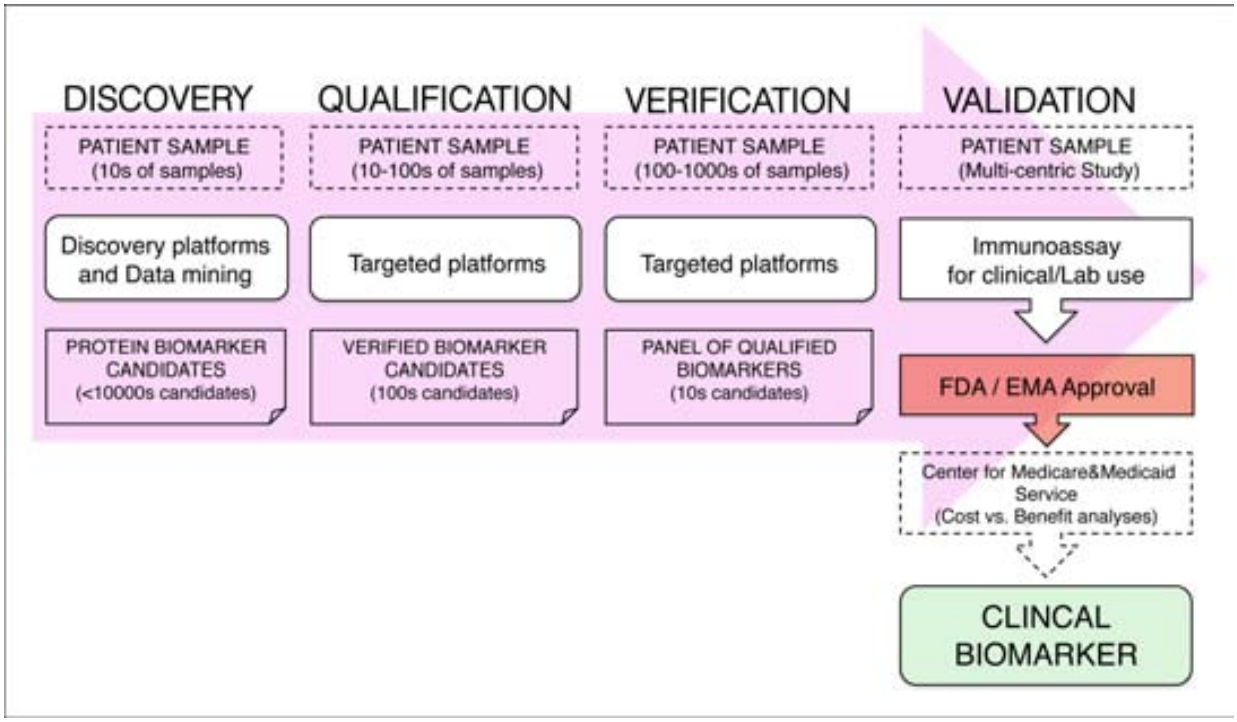



**Figure 13. New Biomarker Pipeline and Steps involved in the validation of a biomarker.** Adapted from Rifai *et al,* [217] and Madu CO *et al,* 2010 [220].

## 4.2 URINE AS A SOURCE OF BIOMARKERS

Human urine plays an important role in clinical diagnostics. For centuries, physicians have attempted to use urine for the non-invasive assessment of disease. The philosopher Hermogenes (5th century before the common era) long ago described the color and other attributes of urine as indicators of certain diseases [221].

Urine is produced by the kidneys and allows the human body to eliminate waste products from blood. The kidneys also maintain homeostasis and produce hormones [222]. The human kidney is composed of 1 million functional units, called nephrons, which filter the plasma and yield what is termed, "primitive" urine. Also found in the kidney is the renal tubule, which reabsorbs most of the primitive urine (Figure 14). In 24h, about 900 liters of plasma flow through the kidneys, where 150-180 liters are filtered. Low molecular weight proteins (< 40 kDa) readily pass through the glomerular filtration barrier and are reabsorbed (more than 99% of this primitive urine is reabsorbed). Because of their low plasma concentrations, only small amounts are seen in the urine. Normal urine contains up to 150 mg / 24h of protein [223]. The remainder, or "final" urine, exits the kidney via the ureter into the bladder. Therefore, urine may contain information not only from the kidney and the urinary tract, but also from distant organs via the plasma and obtained through glomerular filtration (See Table 6). In healthy individuals, 70% of the soluble urinary proteome originates from the kidney and the urinary tract, whereas the remaining 30% originates from the ultrafiltration of plasma by the glomerulus [224]. Diseases, which adversely affect kidney function, also cause the excessive loss of proteins in the urine (see Box VIII). The analysis of the urinary proteome and genome could, therefore, allow the identification of biomarkers for both urogenital and systemic diseases. To define the disease-specific urinary biomarker(s) it is crucial to include appropriate and sufficient controls [225]. Urine samples obtained from patients with other diseases or disorders possessing clinical, biochemical and metabolic profiles similar to those of the disease of interest, must be included as controls [226]. This inclusion should also consider variations that may occur throughout the differential stages of the disease.

Until now, most of the studies looking for specific disease biomarkers have been performed using tissues obtained directly from the affected organs. Although using tissue samples has the main advantage of providing information, which can only be obtained directly from the affected organs, the translation of this technique to clinical practice is much more complicated [225]. In general, working with body fluids, such as urine, has several advantages compared to working with tissue samples.

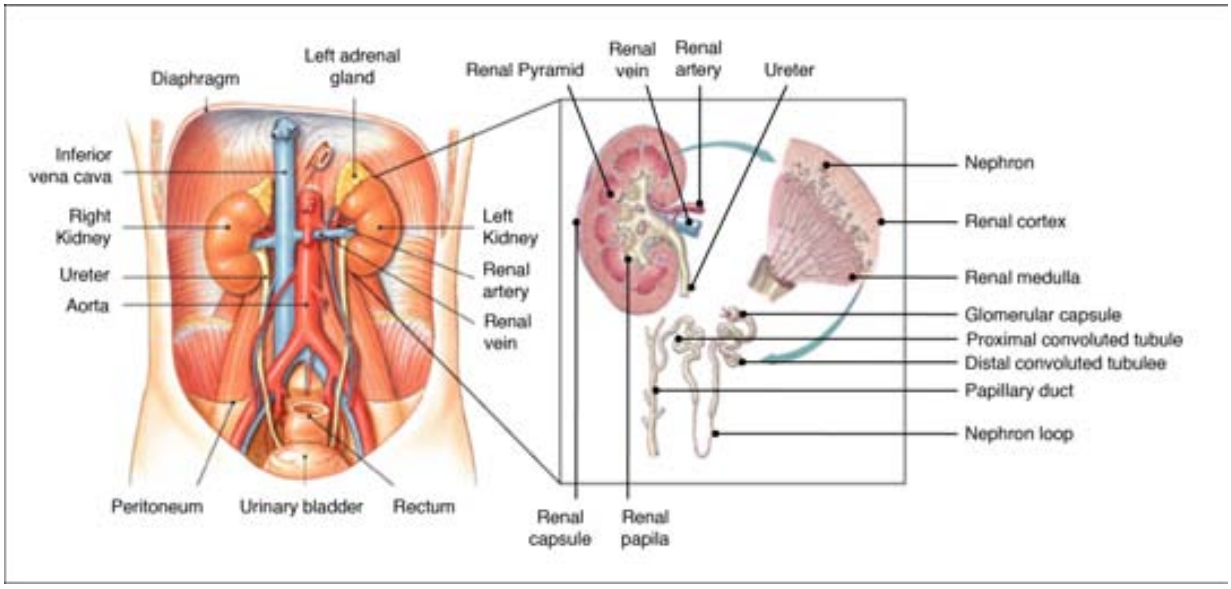



**Figure 14. Urinary System squeme**. Adapted from Human Anatomy book [227].

**Table 6. Urinary protein composition.** Adapted from Barratt J et al., 2007 [228]

| Protein type | % Proteins excreted | Source | Abnormalities |
|---|---|---|---|
| Soluble proteins | 49 | (1) Glomerular filtration of plasma proteins (free passage of proteins < 40 kDa). (2) Some soluble proteins are excreted into urine by epithelial cells. | (1) Defects in glomerular filtration; increase excretion of high-molecular weight proteins. (2) Defects in reabsortion of glomerual filtrate; increase excretion of low-molecular weight proteins. |
| Urinary sediments* | 48 | (1) Mainly sloughed epithelial cells (from podocytes to urethral epithelial) and casts#. (2) Shedding of microvilli or paopotosis of epithelial cells may generate small fragments of cell membranas. | Numbers of whole cells and casts increased in many renal diseases. |
| Urinary exosomes | 3 | All epithelia lining the urianry tract (from podocytes to urtehral epithelia). | Currently unclear whether excretion of urinary exosomes is altere in diseases of the kidney and urinary tract. |

* Urinary sediments can be separated by moderate centrifugation speeds.
# Urinary casts are cylindrical structures produced by the kidney and present in the urine in certain disease states.

## THE MAIN ADVANTAGES OF USING URINE:

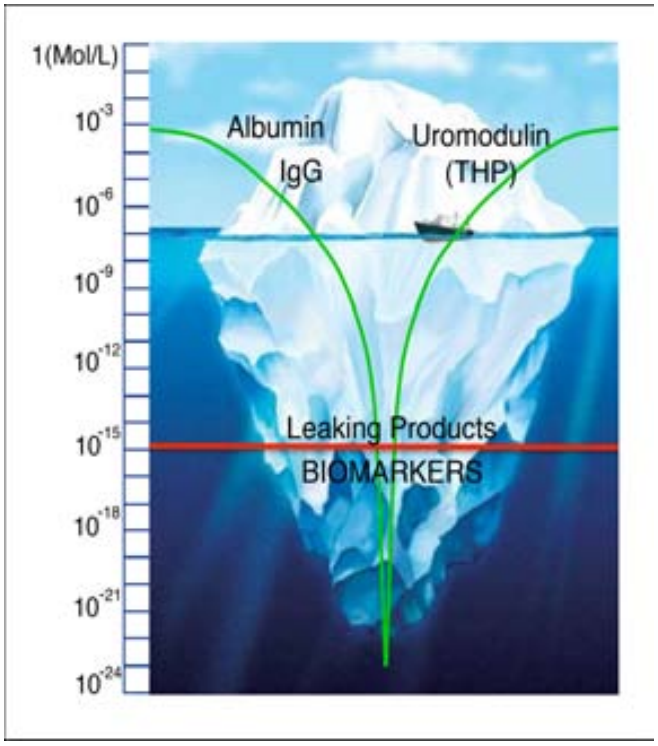
(1) Urine can be obtained in large quantities using non-invasive procedures. This allows repeated sampling of the same individual for disease surveillance. The availability of urine also allows an easy assessment of reproducibility and/or the improvement in sample preparation protocols. (2) Urinary peptides (fragments of proteins) and lower molecular mass proteins are in generally soluble. Therefore, solubilization of these low molecular weight proteins and peptides is generally not an issue. Further analysis of these lower molecular mass compounds (< 40 kDa) can be performed easily without additional manipulation. (3) The urinary protein

content is relatively stable, probably due to the fact that urine "stagnates" for hours in the bladder; hence, proteolytic degradation by endogenous proteases may essentially be completed at the time of voiding [229] [230]. In addition, urine can be stored for several years at -80°C without significant alterations of its proteome. (4) Since urine contains serum proteins and peptides that pass through the ultrafiltration (< 40 KDa) process in the kidneys (including prostatic proteins), changes in the kidney and genitourinary tract are reflected in the urinary proteome, as well as changes from more distant sites. Urine can be used to detect exfoliated normal cells and cancer cells or their secreted protein products, such as prostate secretions. (5) As urine contains sediments, including exfoliated normal or cancer epithelial cells and urinary casts, we can encounter entire cells in urine samples that contain a wealth of genetic material, such as DNA, RNA and microRNA.

THE MAIN DISADVANTAGES OF USING URINE:

(1) Urine widely varies in cell, protein and peptide concentrations, mostly due to differences in a patient's daily intake of fluid or diet. However, this shortcoming can be countered by standardization, based on creatinine levels or the total amount of protein content. (2) The definition of disease-specific biomarker(s) in urine, and most likely in other compartments, is also complicated by significant changes that are likely caused by variations in the diet, metabolic and/or catabolic processes, circadian rhythms, and exercise, as well as the circulatory levels of various hormones [231]. (3) The reproducibility of any analysis is reduced by physiological changes, even if the analytical method shows high reproducibility. (4) The high complexity of the sample may present a problem. (5) High levels of salts and/or other interfering compounds may hinder the method of analysis. (6) Urine samples possess a high degree of variation (intra-individual and inter-individual). (7) These samples also possess a high dynamic range (Figure 15). The dynamic range is defined as the ratio between the largest and smallest possible values of a changeable quantity, in this case, protein quantity. Plasma samples exhibit tremendous variations in individual protein abundances, typically of the order of $10^{10}$ or more, with the result that in any two-dimensional (2D) map, only the high-abundance proteins are displayed [232]. In the case of urine, the problem is further aggravated by its very low protein content, requiring a concentration step of 100- to 1000- fold [233]. Exploring the hidden or low abundance urine proteome through the equalization of the most abundant proteins allows us to reveal the "deep proteome." It is believed that, within this deep proteome, a few, potential novel biomarkers for the disease may be represented [234]. (8) Urine contains active enzymes, such as proteases, DNAses and RNAses that can degrade both DNA, RNA, and proteins. The addition of Proteases Inhibitors is recommended, in order to avoid protein degradation [235]. The protection of RNA in a urine sample is more

complicated, as the addition of compounds usually involves cell lysis. This process could produce a loss of RNA material during subsequent sample preparation steps. To avoid RNA degradation, urine sediments have to be kept on ice (4°C). (9) Screening for biomarkers in urine may have limitations when the protein biomarkers derived from the affected organs, tissues, or cells are present at very low levels. Moreover, the protein components in each biofluid can be affected by several organ systems and various cell types [225].




BOX VIII: ABNORMAL CLINICAL AND LABORATORY FINDINGS FOR URINE

(1) Hematuria is the presence of red blood cells (erythrocytes) in the urine. (2) Eosinophiluria is the abnormal presence of white blood cells (eosinophils) in the urine. (3) Proteinuria is defined as an excessive rate of plasma protein excretion in the urine. (4) Glycosuria, Ketonuria, Bilirubinuria, Hyperuricosuria / Hypouricosuria, and Aminoaciduria are different urination disorders that all include the abnormal presence of small molecules in the urine. (5) Bacteriuria denotes the presence of bacteria in urine that is not the result of sample contamination during collection.

**Figure 15. Dynamic range representation squeme; Iceberg.** (Adapted from www.hispalive.es)

## 4.2.1 URINE AS A SOURCE OF BIOMARKERS FOR PROSTATE CANCER

Body fluids are liquids originating from inside the body of living people. They include fluids that are excreted or secreted from the body. A proximal fluid is a body fluid closer to, or in direct contact with, the site of disease. Those fluids may have several properties that make them attractive for biomarker discovery, since they normally contain secreted proteins, which are shed or leaked, as well as cells or cell fragments from diseased tissue [217].

Given the location of the prostate and its involvement with the urethra, the possibility of prostate epithelial cells being released into the urine is a verified fact. The main function of the prostate gland is the secretion of proteins into the prostatic fluid, which is then combined with seminal vesicle derived fluid that is released into the prostatic urethra, in order to promote sperm activation and function [236]. Malignant cells may also secrete specific products that will end up in the same place. Expressed-prostatic secretion (EPS) fluid is the fluid

secreted by the prostate gland after a prostate massage (PM). The gentle massage of each side of the prostate gland during PM stimulates the release and movement of prostatic fluids and detached epithelial cells into the urethra [236]. EPS can be collected in the first voided urine after the DRE examination. Although in 1958 Papanicolaou described the presence of PCa cells in voided urine [237], the frequency of those cells was low [238]. There is much evidence supporting the idea of prostate manipulation by PM, in order to increase the presence of PCa cells in voided urine. All of these studies have generally reported that the cells are low in number and are also fragile; therefore, urine requires careful collection, manipulation and storage. Koss *et al,* hypothesized that PCa cells in voided urine without DRE manipulation could represent a sign of advanced cancer, compared to patients presenting cancer cells in their urine after PM [239]. This could be explained by the premise that larger, more invasive tumors are more likely to shed cells into the prostatic urethra.

In summary, post-PM urine can easily serve as a mirror of what is happening within the prostate. Furthermore, urine collection can be accomplished without disruption of clinical standard practice. It can also be repeated several times throughout the course of the prostatic disease. For all of these reasons, post-PM urine can serve as a potential source of prostate disease biomarkers. Nevertheless, using urine for biomarker discovery represents an important technical challenge, both in genomic and proteomic approaches. Although there are many studies that focus on genomic approaches and biomarker discovery and identification, there still exists some controversy regarding the standardization of collection procedures, sample processing, storage and normalization. In addition, with regards to proteomics, there is still quite a lot of work to do (see section 5; Clinical Proteomics).

## 4.3 PROSTATE CANCER "URINE" BIOMARKERS

The existing clinical biomarkers for PCa are not ideal, since they cannot specifically differentiate between those patients who should be treated for aggressive forms of the disease and those who should avoid over-treatment of the indolent form [220]. For PCa diagnosis, PSA is the only conventional biomarker accepted by the FDA. Even if it is not an ideal biomarker, it has been employed worldwide to diagnose and monitor men with PCa. As explained earlier, PSA is notoriously weak on specificity. Therefore, there is a need for PCa biomarkers possessing better specificity, especially within the serum PSA "gray zone," for use in the clinics.

The ideal biomarkers for use in the early detection of PCa should be prostate-specific and able to differentiate between normal, BPH, HGPIN and PCa [220]. It is for this reason that the selection of the study population should be included at all steps of the research: from

biomarker discovery to biomarker verification, every stage needs to adhere to strict, thorough protocols. However, diagnostic markers are not the only biomarkers needed; there is also a lack of prognostic markers, which can identify those patients with early, localized PCa, who would benefit from curative therapy, and those patients, who present advanced disease needing specific treatment to prevent large, metastatic lesions. Most of the currently used biomarkers were discovered unexpectedly, while others, such as PSA, are often used without knowing their specific molecular functions within the given pathologic condition. Many pilot and discovery research has been published, but large validation studies using body fluids have failed to replicate promising results. However, in the last 5 years biomarker discovery programs using urine samples have emerged for detecting and predicting the aggressiveness of PCa, and some of these programs have reported promising results [240].

Molecular biomarkers can take different forms, and as a consequence of this fact, a variety of strategies have been adopted for their discovery. The search for effective biomarkers has principally included transcriptional profiling, DNA methylation, metabolomics, fluxonomics, and more recently, proteomics [217]. These emerging biomarkers will be beneficial and critical in the development of new and clinically reliable indicators, which will have a high specificity for the diagnosis and prognosis of PCa. Ideally, acquisition of these biomarkers will be less invasive, will be useful for screening PCa patients (especially older patients), and will be able to guide patient management, in order to provide maximum benefit while minimizing treatment-related side-effects and risks [200].

On the basis of their description, biomarkers can be divided into three groups: (i) DNA-based, (ii) RNA-based, and (iii) Protein-based. DNA-based biomarkers include single nucleotide polymorphisms (SNPs), chromosomal aberrations, changes in DNA copy number, microsatellite instability, and differential promoter-region methylation [148]. In urine, CpG hypermethylation of GSTP1 [98, 241, 242] is detected using methylation-specific PCR on urinary cells. GSTP1 is the main urine DNA-based biomarker investigated that, alone or in combination with other methylation markers, has shown promising results for PCa detection. In this project, we first focused on urine RNA-based biomarkers that have already been described in the literature, and then, we tried to define new, urine protein-based biomarkers for PCa. RNA-based and protein-based biomarkers are further explained below. Recently, another novel approach to finding biomarkers for PCa was described. This approach analyzed urine exosomes or proteasomes, which are secreted vesicles that contain proteins and functional RNA molecules. In fact, several studies have found a more elevated rate of exosome secretion in men with PCa than in those with other diseases [243]. The most studied PCa biomarkers, which have been found in urine, are summarized in Table 7.

**Table 7. Summary of most relevant PCa biomarkers in the literature**

| GENE | DESCRIP-TION | PROTEIN | EXPRESSION | TYPE OF BIOMARKER | SAMPLE | REFERENCE |
|---|---|---|---|---|---|---|
| PCA3 | Prostate Cancer Gene 3 | Non conding mRNA (unknown function) | Prostate specific and highly up-regulated in PCa. | Diagnostic (in gray zone) | Tissue and urine | Hessels *et al,* [244]; Grok Skopf *et al,* [245] |
| PSGR | Prostate Specific G-coupled Receptor | Receptors coupled to heterotrimeric GTP-binding proteins | Over-expressed in PCa. | Diagnostic | Tissue and urine | Xu LL *et al,* [246, 247]; Rigau M *et al,* [248] |
| PSMA | Prostate Specific Membrane Antigen | type II membrane protein | Over-expressed in Pca compared to BPH and normal. | Diagnostic | Tissue, blood and urine | Horoszewicz JS *et al,* [249]; Beckett ML *et al,* [250] |
| AMACR (P504) | Aplha-Methylacyl-CoA Racemase | Enzyme involved in branched chain fatty acid oxidation | Over-expressed in PCa, in HGPIN and in some other carcinomas, both at RNA and protein level. | Diagnostic (in grey zone) and Prognostic | Tissue, blood and urine | Sreekumar A *et al,* [251]; Zehentner BK *et al,* [252] |
| TMPRSS 2: ERG | Gene fussion | Androgen drives the expression of ETS and causes tumor proliferation | The most common gene fusion in PCa. Over-expressed PCa and related to PCa aggressiveness. | Prognostic biomarker for aggressive PCa and detection of PCa | Tissue and urine | Tomlins, SA *et al,* [253]; Hessels, D *et al,* [254] |
| GSTP1 | Glutathione S-transferase P1 | Enzyme involved in protecting DNA from free radicals | Loss of GSTP1 expression due to the promoter hypermethylation (< 90% of PCa). | Diagnostic (for predicting repeating PB) | Tissue and urine | Nakayama M *et al,* [97;] Hoque MO *et al,* [244] |
| ANXA3 | Annexin A3 | Calcium and phospholipid binding proteína; Implicated in cell differenciation and migration | Presence in urinary exosomes and proteasomes. Lower production in PCa than in BPH, HGPIN and benign. | Under-expressed in PCa. Prognosis | Tissue and urine | Wozny W *et al,* [255] Gerke V *et al,* [256]; Pisitkun T *et al,* [257]; Schostal M *et al,* [258] |
| TERT | Telomerase reverse transcriptase | Mantains the telomeric ends of chromosomes | Over-activity in 90% PCa (association with Gleason score). | Prognosis | Urine | Meid FH *et al,* [259]; Vicentini C *et al,* [260] |
| Sarcosine | Sarcosine | N-methyl derivative of the amino-acid glycine | Differentially exressed metabolite elevated during PCa progression to metastasis. | Prognosis | Urine | Sreekumar A *et al,* [261] |
| GOLM1 | Golgi membrane protein 1 | cis-Golgi membrane protein (unknown function) | Overexpressed in PCa. | Diagnostic | Urine | Laxman B *et al,* [262]; Varambally S *et al,* [263] |
| KLK2 | Human Kallikrein 2 | Serine protease | Over-expressed in PCa. | Diagnostic and Prognostic | Tissue and blood | Darson MF *et al,* [264]; Haese A *et al,* [265] |
| PAP or ACPP | Human Prostatic acid phosphatase | Enzyme phosphatase | Over-expressed in PCa and in bone metastasis. | Diagnostic and Prognostic of PCa bone metastasis | Blood | Gutman AB *et al,* [266] |
| PSCA | Prostate Stem Cell Antigen | Membrane glycoprotein | Specific production in the prostate and possible target for therapy. | Prognostic | Tissue and blood | Gu Z *et al,* [267]; Gu Z *et al,* [268] |
| IL-6 | Interleukin-6 | Secreted cytokine are involved in the immune and acute-phase response | Increased concentrations of IL-6 and IL-6R in metastatic and androgen-independent PCa. | Diagnosis and Prognosis | Blood | Shariat SF *et al,* [269]; Nakashima J *et al,* [270]; Shariat SF *et al,* [271] |
| uPA and UPAR | Urokinase Plasminogen Activator and Receptor | Urokinase plasminogen-activation cascade | Over-expressed in BPH and PCa vs benign. | Prognostic (associated with bone metastasis) | Tissue and blood | McCabe NP *et al,* [272]; Shariat SF *et al,* [273] |
| EPCA / EPCA-2 | Early Prostate Cancer Antigen / 2 | Nuclear matrix protein | Over-expressed in PCa. | Diagnostic and Prognostic | Tissue and blood | Paul B *et al,* [274]; Leman ES *et al,* [275] |

### 4.3.1 RNA-BASED BIOMARKERS FOR PROSTATE CANCER

The term "omics" informally refers to a field of study in biology ending in *-omics*, such as genomics or proteomics. The related suffix *-ome*, as used in molecular biology, refers to a *totality* of some sort; it is used to address the objects of study by field, such as the genome or proteome, respectively. The genome is the entirety of an organism's hereditary information. In humans it is encoded in DNA. The genome includes both the genes and the non-coding sequences of the DNA. The transcriptome is the set of all RNA molecules in one or a population of cells. Messenger RNA (mRNA) is the molecule that belongs to the DNA sequence, which is transcribed to RNA, in order to finally be translated into a protein. As the transcriptome includes the study of mRNA transcripts, it somehow reflects the genes that are actively expressed at any given time and that are somehow related to the proteins to be translated. The RNA-based biomarkers mainly include over- or under-expressed transcripts and regulatory RNAs, such as microRNAs [148].

Improvements in RNA microarray platforms, quantitative Real Time Polymerase Chain Reaction (RTqPCR), and the development of new high throughput technologies, such as Next-generation Sequencing, allow us to better understand the expression profile of these single cells, populations of cells, specific tissues, while also allowing comparison between different pathological conditions. RTqPCR is generally considered the "gold standard" against which other methods are validated. It can now be performed at relatively high-throughput [148]. However, there are considerable limitations to those technologies, mainly due to the lack of standardization protocols. The absence of standardized methods makes it difficult to integrate findings, as well as treat the RNA from clinical samples, since it is often very unstable [276].

In recent years, within the field of PCa, there have been several studies that focused on the study of expression profiles in different types of samples, such as tissue, blood, and urine. For our purposes, we were interested in those studies that focused on the use of body fluids, such as urine, as a non-invasive technique for sample acquisition. A wide range of promising PCa biomarkers, which are not only prostate-specific, but are also differentially expressed in prostate tumors, have been identified, such as the TMPRSS2:ERG gene fusions [254, 277-280] and prostate cancer antigen 3 (PCA3) [278, 281-283]. Although urine RNA-based testing for PCA3 expression has already been documented in large screening programs, there are only a few studies that take into account the heterogeneity of cancer development based on a diagnostic profile. A combination of various biomarkers would clearly improve performance over the use of a single biomarker [284], since a single marker may not necessarily reflect the multifactorial nature of PCa. Using multi-gene panels or "fingerprints" can normally overcome

single biomarkers in the ability to differentiate between the diseased and benign states. Nevertheless, choosing which genes to include in the pattern adds an additional layer of statistical complexity and also prompts new developments in biostatistics and bioinformatics [148].

#### 4.3.1.1 Prostate Cancer Gene 3 (PCA3)

PCA3 ($PCA3^{DD3}$ or $DD3^{PCA3}$) is located on chromosome 9q21-22 and was originally described as consisting of 4 exons with alternative polyadenylation at 3 different positions on exon 4. Exon 2 was found to be absent in 60% of the cDNA clones analyzed [285]. Recently, Clarker and coll. described additional complexities of the PCA3 gene with 4 new transcription start sites, 2 new differentially spliced exons and 4 new polyadenylation sites; however, both of these new forms only represented a minority of the PCA3 transcripts [286]. One important observation within the PCA3 cDNA sequence was the presence of a large number of stop codons. The nuclear localization of transcripts, together with the high frequency of stop codons, suggested that PCA3 was a noncoding RNA (ncRNA) [285] (see Box VIII for ncRNA).

BOX VIII: NON-CODING RNA

The "central dogma" of molecular biology is that genetic information flows in one direction with proteins as the end product. However, growing evidence has emerged to describe the role of the RNAs that are not translated into proteins. NcRNAs comprise microRNAs, anti-sense transcripts and other Transcriptional Units containing a high density of stop codons and lacking any extensive "Open Reading Frame" [288]. Several types of ncRNAs have been implicated in gene regulation via modification of the chromatin structure, alterations to DNA methylation, RNA silencing, RNA editing, transcriptional gene silencing, post-transcriptional gene silencing, and enhancement of gene expression [289-291]. It is becoming clear that these novel RNAs perform critical functions during development and cell differentiation [288]. Small ncRNAs, such as microRNAs (miRNAs) and small interfering RNAs (siRNAs), are well-studied in the roles they play in gene silencing and are reported to be aberrantly expressed in many cancers [289]. Together, all this information indicates that ncRNAs are emerging as a new class of functional transcripts in eukaryotes, formally described as junk DNA.

PCA3 was first identified in 1999 using a differential display analysis that compared mRNA expression in normal and tumor-bearing prostate tissues in RP specimens [285]. In 95% of the 56 human prostatectomy specimens examined, PCA3 was highly over-expressed in primary PCa tumors compared to adjacent non-cancerous tissue [285]. By Northern blot analysis, the normal prostate and BPH tissue, taken from the same subjects, expressed low levels to no PCA3 at all [285]. By RTqPCR analysis, PCA3 was demonstrated to present low levels of expression in

normal prostate and BPH tissue, whereas the median up-regulation of PCA3 in PCa cells, relative to non-malignant tissue, was about 60-fold [244, 287]. In addition, no PCA3 was detected in other tissues or tumors [285, 287].

The measurement of PCA3 mRNA versus PSA mRNA in urine was first proposed as a marker for PCa by Hessels *et al*, [244]. Using time-resolved fluorescence RTqPCR, they detected PCA3 in post-DRE urine sediments. The levels of PSA mRNA were used to normalize the expression of PCA3 mRNA. PSA mRNA was expressed in benign cells and PCa cells, as well, at approximately the same levels with a modest (1.5-fold) decrease in PSA expression in PCa cells [292]. They reported an Area Under the Curve (AUC) value of 0.72 and a diagnostic accuracy of 67% and 83% (values of sensitivity and specificity, respectively) [244]. Later on, this study was verified in a large, European multi-center study, which concluded that PCA3 possessed potential as an aid in PCa diagnosis [293].

Currently, there exists a commercially available, highly sensitive novel transcription-mediated amplification based test (TMA) method (PCA3, Gen-Probe Incorporated) for urinary assay, which is already in use in the clinics [245]. This method measures PCA3 mRNA and PSA mRNA in the first urine catch collected after DRE. Instead of using urine sediments, this assay has been optimized to measure PCA3 in whole urine samples mixed with a detergent stabilization buffer. The PCA3 score is the ratio of PCA3/PSA mRNAs multiplied by 1,000. The assay demonstrated a sensitivity of 69%, a specificity of 79% and an AUC of 0.75 [245]. Several studies have confirmed that PCA3 performance was independent of serum PSA levels and, moreover, that PCA3 was unaffected by prostate size [294-296]. Furthermore, the PCA3 score correlated with the results of a subsequent PB, as a high score increased the likelihood of a PCa positive result. It had been demonstrated that men with a PCA3 score < 5 had a PCa positive biopsy rate of 14%, while 69% of the men with a PCA3 score > 100 had PCa upon biopsy [297]. A cutoff PCA3 score of 35 was adopted, and this score combined the greatest cancer sensitivity and specificity (54% and 74%, respectively) [295]. In 2006, the Gen-Probe PCA3 assay was accepted for use in Europe, under the name, PROGENSA ® PCA3.

A larger, follow-up validation study in North American men and two European studies were then developed. The largest international, multicentric, double-blind study reported to date was published by Aubin and coll. for the evaluation of the clinical performance of PCA3 in men with elevated PSA levels and previous negative biopsy [298]. The Reduction by DUtasteride of Prostate Cancer Events (REDUCE) trial was designed to evaluate the use of a drug for the chemoprevention of PCa [299]. This trial was a prostate cancer risk reduction study that evaluated men with moderately increased serum PSA and negative biopsies at baseline [298].

Urine PCA3 scores were determined for patients in the placebo arm of the REDUCE trial before their year 2 and year 4 biopsies. The informative rate of this study was 94%, corresponding to the total number of subjects analyzed (1072 patients). The PCA3 score was associated with PCa ($p<0.0001$) and also correlated with Gleason grade 6 or less than 6 versus Gleason more than 6 ($p=0.0017$). The performance of PCA3 was better than that of serum PSA (AUC = 0.69 vs. AUC = 0.61, respectively). A multivariate model using PCA3, serum PSA, %fPSA, prostate volume, age and family history (AUC = 0.75) outperformed the PCA3 and serum PSA as single markers. PCA3 at year 2 was a significant predictor of the 4 years biopsy outcome (AUC 0.63, $p=0.0002$), whereas serum PSA and % fPSA were not predictive. These results confirmed that PCA3 correlated with the probability of biopsy results and also displayed an associative significance for biopsy results. Finally, PCA3 was demonstrated to be insensitive to prostate volume. In conclusion, increased PCA3 scores indicated an increased risk of PCa and predicted future biopsy outcomes [298].

In addition some authors have correlated PCA3 scores with tumor aggressiveness at different levels. PCA3 scores have also been positively correlated with tumor volume [300, 301]. Moreover, PCA3 scores were found to be significantly associated with prostatectomy Gleason scores (6 versus ≥ 7) and "significant" cancers [300]. Finally, PCA3 scores were found to be independent predictors for the extracapsular extensions of tumors [301]. These data, while requiring further confirmation, suggest that PCA3 can be useful in the identification of aggressive PCa, which requires more aggressive treatments [302].

Some studies have also correlated the urine PCA3 score with the presence of HGPIN at biopsy, yielding conflicting results. Dereas *et al*, reported no difference between the PCA3 scores for HGPIN and for normal samples [295], whereas Haese *et al*, found increased PCA3 scores in men with HGPIN [294].

After the development of the PCA3 test assay different studies are published using PCA3 as a biomarker for detecting PCa. As these studies used different methodologies as, target capture with magnetic particles, TMA and a hybridization protection assay, to detect PCA3 transcripts in patient urine and used different disease prevalences, directly comparing of the results would have been inappropriate however the determination of the cut-off score, PCA3 transcripts/PSA transcripts in urine, is identical. The main drawbacks of this test are the high costs around 250€ per test and the fact that it is a closed system, which not permits the user to test in parallel additional biomarkers.

In summary, PCA3 testing may be of valuable help in several PSA quandary situations (Table 8). Some potential applications of the PCA3 assay include testing prior to first biopsy and deciding whether or not to repeat a biopsy in men with elevated serum PSA levels and previous negative biopsies [302] [297]. Other potential applications that still need further validation include detecting local recurrence following radical prostatectomy or radiation therapy and monitoring patients receiving drug therapies that affect serum PSA levels [297].

**Table 8. The possibilities of the PCA3 test in clinical practice.** Adapted from Hessels and Shaken, 2009 [297]

| PCA3 score + prostate biopsy result | Course of action to consider | Established prognostic factor |
|---|---|---|
| Low + negative | Conservative follow-up | PSA-kinetics |
| High + negative | Advanced imaging | PSA-kinetics |
| Low + positive | "Active surveillance" | Clinical stage and grade; PSA-kinetics |
| High + positive | Intervention | Clinical stage and grade |

### 4.3.1.2 Prostate Specific G-coupled Receptor (PSGR):

Prostate specific G-protein coupled receptor (PSGR), also known as the OR51E2 gene is located on chromosome 11p15 and is transcribed into a 1474-bp cDNA sequence [246]. The PSGR gene has one short non-coding exon 1 and one long coding exon 2 (encodes the full protein sequence), separated by a 14.9Kb intron [303]. Two distinct promoters have been described that control the transcriptional regulation of PSGR in human prostate cells [304]. The presence of 2 distinct promoters and clusters of transcription factors binding motifs upstream of position -311 suggest that the expression of PSGR may have two mRNA transcripts; those transcripts have been confirmed by Northern Blot analysis [304]. Both promoters show prostate cell-specific characteristics and can be activated by Interleukine-6 (IL-6) in human prostate cancer cells [304]. Moreover, it has been hypothesized that binding IL-6 to its receptor could activate 3 downstream signaling pathways, including PI3-k, Stat3 and MAPK, which could potentially normalize the transcriptional regulation of a large number of genes, such as PSGR [304]. Also, steroid hormones, such as androstenone and androstadienone have been identified as odorant receptor (OR) ligands. These hormones elicit rapid $Ca^{2+}$ responses in LNCaP PCa cell line and in primary human prostate epithelial cells. The activation of PSGR has been found to causes phosphorylation of p38 and SAPK/JNK MAPKs, resulting in reduced proliferation rates in PCa [305].

PSGR expression has been found to be restricted to human prostate tissues in prostatic epithelial cells [246], although mRNA from PSGR has also been found to be expressed in the olfactory zone and the medula oblongata of human brain [303]. The PGSR protein sequence (320

aa, 35,4 kDa protein) is based on 7 transmembrane-spanning domains and a conserved DRY motif in the second intracellular domain [246]. The PSGR protein is homologous to the G-protein-coupled OR, and it also corresponds to the sub-family Gα12 receptor [246].

Receptors coupled to heterotrimeric GTP-binding proteins (GPCRs) are integral membrane proteins involved in the transmission of signals from the extracellular environment to the cytoplasm. GPCRs, which are activated by factors ranging from small amines to hormones and chemokines, initiate signaling cascades that result in cell shape changes, formation of a migrating front, and altered adhesion [306]. Heterotrimeric G-protein family is composed of α, β, and γ subunits. The Gα proteins comprise 4 subfamilies (Gs, Gi/o, Gq and G12/13). Gα12 receptor, which is a member of the subfamily of PSGR, has been implicated in different cellular functions, such as Rho-dependent cytoskeletal shape change, activation of the c-Jun N-terminal kinase and stimulation of the Na+/H+ exchange. Rho proteins are central regulators of the dynamic reorganization of the actin-based cytoeskeleton and are key mediators of several cellular processes, including cell migration and polarity [306].

PSGR has been described as having a high prostate tissue-specific expression and tumor over-expression, which is restricted to prostate epithelial cells [246, 247, 303, 307]. Comparing HGPIN to PCa, PSGR was found to be over-expressed, especially in early-stage prostate tumors, and was considered to be involved in the presumably early stages of prostatic carcinogenesis [308, 309]. Weng and coll. defined an AUC from Receiver Operating Characteristics (ROC) curve of 0.90 for PSGR, indicating an excellent power of discrimination for PSGR as a PCa diagnostic marker, especially for distinguishing between benign prostate tissue and malignant prostate tissue. However, no significant difference was detected for the expression of PSGR between normal prostate and BPH ($p > 0.05$) [309]. Later on, PSGR was found to be over-expressed in non-organ confined (pT3) prostate tumors compared to organ-confined tumors (pT2), suggesting that this marker could be involved in cancer progression [247]. More recently, PSGR was found to be over-expressed in the peripheral blood of PCa patients. It was suggested that it could serve as a useful diagnostic or prognostic marker; however, the authors stated that these findings were low on specificity and, therefore, recommended further analysis [310]. Finally, PSGR was investigated as a possible diagnostic tool, able to distinguish between normal tissue and PCa in urine samples, since as a single marker [248], or in combination with other biomarkers, it outperforms PSA as a diagnostic marker [311].

#### 4.3.1.3 Prostate Specific Membrane Antigen (PSMA):

Prostate Specific Membrane Antigen (PSMA) was first described in 1987 by Horoszewicz *et al*, as a glycoprotein expressed in the plasma membranes of prostate epithelial cells [249]. PSMA was discovered in the LNCaP human prostatic cell line as the target of mAb 7E11 [249]. The PSMA gene consists of 19 exons that span 60 Kb of genomic DNA and encodes a type II transmembrane protein of 750AA [312]. PSMA is a non-shed integral membrane protein with folate hydrolase activity produced by prostatic epithelium. The extracellular domain mediates homodimer formation. In the normal prostate, PSMA exists as a splice variant, known as PSM', that lacks the transmembrane domain [313].

Expression analysis of PSMA has revealed that it is highly restricted to secretory cells within the prostatic epithelium [249], although PSMA has also been detected in other tissues, such as the kidney, the proximal small intestine, and the salivary glands [314]. Furthermore, the presence of PSMA has been listed as absent to moderate in most BPH and benign tissues. However, PCa tissues have been shown to possess the greatest intensity [249]. It has been suggested that the expression of this molecule could be linked to the degree of tumor differentiation [315]. Subsequent studies have confirmed that PSMA expression is highest in high-grade tumors, metastatic lesions and androgen-independent disease [316]. These findings have spurred the development of PSMA-targeted therapies for PCa, including immunotherapy, radioimmunotherapy, chemotherapy and gene therapy. They have also initiated clinical trials of first-generation products [316]. However, the general clinical application of these therapies still requires extensive clinical studies, in order to test their clinical safety, stability and efficacy.

Several studies have demonstrated that PSMA is abundantly expressed in the new blood vessels that supply most non-prostatic solid tumors, though it is not found within the normal vasculature [316]. PSMA was first proposed as serum prognostic marker for PCa in 1999; however, its use is controversial [250]. A Dual-Monoclonal Sandwich Assay for PSMA was developed to be used on tissues, seminal fluid and urine [317]. Later on, levels of serum PSMA were suggested as a marker for distinguishing between BPH and PCa [318], and the same results were found in urine samples [319]. The transcript levels of the PSMA gene in blood samples show a poor prognosis value, especially in those patients who present more aggressive, metastatic disease [320]. Recent clinical trials have provided significant proof of principle for PSMA as a target for PCa and neovascular therapy [316]. A commercial imaging test for PSMA (ProstaScint) has been developed, that uses a radiolabeled 7E11 antibody to PSMA in a radioimmunoscintigraphy assay [321]. ProstaScint is a monoclonal antibody scanning technique that has implications in the staging of patients with newly diagnosed PCa, as well

as use in evaluating patients believed to have recurrent disease. Finally, PSMA has also been found to be present in the urinary exosomes of PCa patients after therapy [322]. However, the use of PSMA has not yet been adopted in clinical practice.

#### 4.3.1.4 Pannels of urine RNA-based biomarkers

The introduction of PSA testing has revolutionized how PCa is diagnosed and managed; however, controversy still exists regarding both the utility of PSA screening for reducing PCa mortality and the risks associated with PCa over diagnosis. For this reason, novel markers are required to improve the specificity of PSA testing. Currently, biopsy strategies may miss heterogeneous tumor foci, and a urine-based assay would possess a great advantage, since the cells of multiple foci from the entire prostate could be released into and collected in the urine [280]. Moreover, as a single marker may not necessarily reflect the multifactorial, multifocal and heterogeneous nature of PCa, a combination of various biomarkers would clearly improve performance over a single biomarker [284]. The use of multiple markers, in combination with clinical and demographic data, will aid in predicting patients who are at risk for developing PCa and for assessing their prognoses.

Nowadays, only a few studies have taken into account the heterogeneity of PCa development based on a diagnostic profile. Those studies followed the approach of combining various biomarkers for the detection of PCa in urine. The principle that underlies the combined biomarker approach is consistent with tests offered for the detection of PCa in tissue specimens [323, 324] and takes into consideration the heterogeneity of cancer development based on a diagnostic profile. The combined model that resulted from these studies provided overall increased sensitivity without decreasing the specificity (Table 9).

Hessels *et al.* performed a study on 108 patients using urine sediments, where the authors combined PCA3 with the TMPRSS2:ERG fusion status. The combination of both markers remarkably increased the sensitivity for the detection of PCa [254]. Another important study came from Laxman *et al.* who combined PCA3 with GOLPH2, SPINK1, and TMPRSS2:ERG fusion status [262]. Ouyang *et al.* developed a duplex RTqPCR assay for the detection of PCa, based on the quantification of AMACR and PCA3 in urine sediments. Finally, Talesa and coll. analyzed the mRNA of PSMA, Hepsin, PCA3, GalNAC-T3 and PSA biomarkers using RTqPCR. They concluded that the best combination of those biomarkers included urinary PSA and that PSMA transcripts were the best predictors of PCa [319].

**Table 9. Summary of the most important studies combining pannels of urine biomarkers for PCa**

| Study | Marker | PCa / Study | Sens. | Spec. | AUC |
|---|---|---|---|---|---|
| Hessels et al. 2007 | mRNA (PCA3 and TMPRSS2:ERG) | 78 / 108 | 73% | ND | ND |
| Laxman et al. 2008 | mRNA (PCA3, GOLPH2, SPINK1 and TMPRSS2:ERG) | 152 / 257 | 66% | 76% | 0.76 |
| Ouyang et al. 2009 | mRNA (AMACR and PCA3) | 43 / 92 | 72% | 53% | ND |
| Rigau et al. 2010 | mRNA (PCA3 and PSGR) | 73 / 215 | 96% | 34% | 0.73 |
| Cao et al. 2010 | mRNA, protein and metabolite (PCA3, TMPRSS2: ERG, Annexin A3, Sarcosine, and urine PSA) | 86 / 131 | 95% | 50% | 0.86 |
| Rigau et al. 2011 | mRNA (PSMA, PSGR, and PCA3) | 57 / 154 | 96% | 50% | 0.82 |

## 4.3.2 PROTEIN-BASED BIOMARKERS FOR PROSTATE CANCER

Proteomics is regarded as a sister technology to genomics. Even though the pattern of gene activity may be abnormal in a tissue with a pathological lesion, there can still be a poor correlation between the level of the transcription of different genes and the relative abundance within the tissue of the corresponding proteins. Consequently, the information regarding a specific pathological process, which can be derived at the level of gene transcription, remains incomplete [325]. The contribution of proteomics to the understanding of the pathogenesis and diagnosis of genitourinary tract diseases has been considerable. Proteomic approaches allow the discovery of disease-specific targets and biomarkers, providing comprehensive diagnostic and prognostic information (see section 5; Clinical Proteomics).

Protein-based biomarkers include cell-surface receptors, tumor antigens (such as PSA), phosphorylation states, carbohydrate determinants, and peptides released by tumors into serum, urine, sputum, nipple aspirates or other body fluids [148]. Most of the FDA-approved cancer biomarkers in clinical use are single proteins, and most of them are serum-derived proteins. However, for every article demonstrating one new clinically relevant biomarker, there are hundreds of articles containing negative or positive results, which are either non-reproducible or of no clinical value [326].

The mapping of the human genome represents a true milestone in medicine and has led to an explosion in discoveries and translate research in the life science [327]. The expansion of biological knowledge through the Human Genome Project (HGP) has also been accompanied by the development of new high throughput techniques, which provide extensive capabilities for the analysis of a large number of genes or the whole genome. The completion of the

human genome has presented a new and even more challenging task for scientists; the characterization of the human proteome. Unlike the genome project, there are major challenges in defining a comprehensive Human Proteome Project (HPP) due to (1) potentially very large number of proteins with PTMs; (2) the diversity of technology platforms involved, (3) the variety of overlapping biological "units" into which the proteome might be divided for organized conquest and (4) sensitivity limitations in detecting proteins present in low abundances. The idea is to provide a human proteome map, where at least one protein product for each coding gene will be characterized at three different levels: abundance, partners in interactions and expression localization [328]. Another goal is to characterize the human proteome with respect to the main diseases affecting humans. When complete, these results will generate a map of the protein-based molecular architecture of the human body, and this will become a resource able to help elucidate biological and molecular functions, as well as advance the diagnosis and treatment of diseases [328].

#### 4.3.2.1 Prostate Cancer Secretome

The term, "secretome," refers to the rich, complex set of molecules secreted from living cells, including proteins that are secreted by the cell via the classical, non-classical or exosomal pathways [329], as well as proteins shed from the cell surface. Proteins of "secretomes" play a key role in cell signaling, communication and migration [330]. Understanging this language could largely increase our knowledge regarding the molecular mechanisms of neoplasia. In addition, extracellular matrix components and other molecules secreted by tumor cells are a rich source of potential markers and drug targets for cancer treatment [331]. Proteins secreted by cancer cells are also essential in the processes of differentiation, invasion and metastasis. These cancer secreted proteins or their fragments always enter body fluids, such as blood or urine, and can be measured via non-invasive assays. For this reason, the cancer secretome may reflect a broad variety of pathological conditions and represent a more reliable source of biomarkers. To date, only a few studies have analyzed cancer secretomes; however, the results regarding biomarker discovery are rather exciting [332].

The need for developing more effective cancer biomarkers and therapeutic methods has led to the study of the cancer cell "secretome" as a means of identifying and characterizing diagnostic and prognostic markers and potential drug and therapeutic targets [329]. Proteomics profiling is a relevant approach to biomarkers discovery and is extensively utilized in secretome analysis. Advances in liquid chromatography (LC) and two dimensional polyacrylamide gel electrophoresis (2DE), in combination with mass spectrometry (MS), have significantly facilitated the challenging detection of these secreted proteins [168].

The main challenges in "secretome" proteomic analysis include the fact that these proteins are frequently found in low concentrations, due to their high dilution in body fluids or cell culture media (CM). Also, their masking and contamination by cytoplasmic or other normally non-secreted proteins, which are released following cell lysis and death, can present problems in their analysis; and, finally, "secreted" proteins are normally hidden by serum proteins.

As a result of these technical problems, only some articles have been published on this subject. In particular, Sardana *et al*. analyzed the CM of 3 human PCa cell lines of differing origins (PC3, LNCaP and 22vR1) using shotgun proteomics. Well-known PCa biomarkers, such as KLK3 and KLK2, were identified, among others. Follistatin, chemokine ligand 16, pentraxin 3 and spondin 2 serum levels were tested by Enzyme-Linked Immuno Sorbent Assay (ELISA) and found to be increased in PCa patients when compared to healthy controls [333]. The same group also performed a qualitative proteomic analysis on conditioned medium from the PCa cell line, PC3(AR)6. 262 proteins were identified from the CM. Mac-2BP and KLK11 levels were found to be elevated, whereas KLK5 and KLK6 decreased in the serum samples of PCa as compared to controls [334]. Martin *et al*. analyzed the prostatic cell line LNCaP grown under androgen stimulation or starvation. 600 proteins were identified, 27% of which were classified as secreted or membrane bound. Knitz-type protease inhibitor 1, suppressor of tumorigenicity 14, vascular endothelial growth factor, insulin-like factor binding protein 2 precursor and amyloid-like protein 2 precursor were confirmed by Western blot [335]. Finally, Goo *et al*. identified 116 secreted glycoproteins from cultured normal prostate stromal mesenchyme cells using the glycopeptide-capture method followed by MS [336]. Stromal proteins found to be upregulated in the prostate included cathepepsine-L, follistatin-related protein, neuroendocrine convertase and tumor necrosis factor.

# 5. CLINICAL PROTEOMICS

The amino acid sequence of proteins provides a link between proteins and their coding genes via the genetic code, and, in principle, a link between cell physiology and genetics [337]. Proteins and peptides "leaked" from tissues into clinically accessible body fluids, such as blood or urine, have led to the possibility of disease diagnosis at earlier stages and the ability to monitor responses to therapy simply by testing for the presence of disease-related biomarkers in these fluids [338].

Ever since the field of proteomics demonstrated the ability to characterize a large number of proteins and their PTMs, including identification and quantification, it has been applied to various areas of science, such as in the discovery of potential diagnostic markers [214]. The field of clinical proteomics shows great promise in medicine, thanks to the new types of proteomic technologies being developed and its combination with advanced bioinformatics, which are currently being used to identify the molecular signatures of diseases based on protein pathways and signaling cascades [214]. However, without proper study design and the correct implementation of these robust analytical techniques, efforts and expectations may be easily hindered.

Proteomic biomarker discovery shares a lot of characteristics with genomic or transcriptomic profiling, including the analysis of biological samples within a complex matrix, and sophisticated statistical analysis, including large sets of variables. However, proteomic biomarker discovery is inherently more complex, mainly due to the vast range of analyte concentrations that must be detected and identified, as well as the fact that protein products cannot be amplified, since an equivalent PCR for proteins does not exist [217]. Innovative protein-based approaches to identify and quantify proteins in a high throughput manner have furthered our understanding of the molecular mechanisms involved in diseases [214]. Because of the inherent complexity of the proteome, all approaches to its examination are generally based on a separation step (gel-based or non-gel based) followed by ionization and a subsequent analysis by MS [339] (Table 10). Furthermore, the depletion of high abundance proteins by affinity columns, affinity enrichment of targeted protein analytes and/or multi-dimensional chromatographic fractionation coupled with MS have all expanded the dynamic range of detection for low abundance proteins by several orders of magnitude in serum and plasma [340], as well as in urine [234].

As explained in section 4, the new biomarker development pipeline involves a discovery stage followed by qualification, verification and, finally, prior to clinical implementation, validation on larger cohorts of samples. Traditionally, the discovery stage is performed on an

MS-based platform for global, unbiased sampling of the proteome, while biomarker qualification and clinical implementation will generally involve the development of antibody-based protocols, such ELISA. The recent explosion in proteomic technologies centering on MS and protein microarrays has provided great opportunities for researchers to use these "bridging technologies" for the clinical, proteomic investigation of disease-relevant changes in tissues and biofluids [214].

Common proteomic methodologies are categorized into two classes: those for differential proteomics in the discovery phase and those for quantitative proteomics.

*Differential proteomics* is defined as the scientific principle that compares normal and diseased states for biomarker discovery without providing specific protein concentrations in the biological matrix.

*Quantitative proteomics* is defined as the absolute quantification of proteins used in targeted biomarker verification and quantification studies. Selected Reaction Monitoring (SRM) is introduced as one of the main multiplex quantitative methodologies in the biomarker pipeline for verification [214].

**Table 10. Advantages and disadvantages of each mass spectrometry-based proteomics technique for use in clinical applications.** Adapted from Decramer et al., 2008 [338]

| TECHNOLOGY | ADVANTAGES | DISADVANTAGES |
|---|---|---|
| 2DE-MS | Detection of large molecules, enables estimation of their molecular weight. Sequencing of biomarkers easy to perform from 2D spots. | Not applicable to molecules < 10 Kda, no automation, time-consuming, quantification difficult, medium througput, moderate comparability. |
| LC-MS | Automation, high sensitivity, used for detection of large molecules (> 20 Kda) after tryptic digest, sequence determination of biomarkers provided by MS/MS. | Time-consuming, relatively sensitive toward interefering compounds, restricted mass range, medium throughput. |
| SELDI-TOF* | Easy to use, high throughput, automation, low sample volumen. | Restricted to selected proteins, low resolution MS, lack of comparability, sensitive toward interefring compounds, low information compound. |
| CE-MS# | Automation, high sensitivity, fast, low sample volume, multidimensional, low cost. | Generally not suited for larger molecules (> 20 kDa). |

*Surface-enhanced Laser Desorption Ionization-Time-of-Flight (SELDI-TOF)

#Capillary electrophoresis coupled to mass spectometry (CE-MS)

## 5.1. DIFFERENTIAL PROTEOMICS FOR BIOMARKER DISCOVERY

The principal technology for protein biomarker discovery is MS. All mass spectrometers produce mass spectra, which plot the mass-to-charge ratio (m/z) of the ions (fragments of proteins) observed (x-axis) versus detected abundances (y-axis) [217]. These instruments are used either to measure the molecular mass of a polypeptide or to determine additional structural features, such as the aa sequence or the site of the attachment and type of PTMs [337]. Single-stage mass spectrometers are used to essentially act as balances to weigh molecules. Double-stage mass spectrometers (MS/MS) are used to act as two individual balances. After the initial mass determination, specific ions are selected and subjected to fragmentation through collision. In a second step, the mass of these fragment ions is determined. In such experiments, referred to as tandem mass spectrometry (MS/MS), detailed structural features of the peptides can be inferred from the analysis of the resulting fragment masses [337].

### 5.1.1. GEL ELECTROPHORESIS

The term "proteomics" originated in the context of 2DE [341]. 2DE represented a fundamental evolution in the field of separation technologies. 2DE has proved to be a reliable and efficient method for the separation of proteins based on mass and charge. It can achieve the separation of thousands of different proteins on one gel [342]. In 2DE experiments, the staining pattern of proteins from two samples is compared, and the "up-regulated" and "down-regulated" proteins are identified. The main problem of 2DE is its lack of reproducibility that difficult the comparison of different experiments.

To overcome 2DE reproducibility, the simultaneous staining of two samples has been developed in recent years to allow a one-step analysis and a direct comparison of different mixtures of proteins [343]. This methodology is known as Differential In-Gel Electrophoresis (DIGE). Briefly, the samples are differentially labeled with fluorescence dyes (eg. Cy3 and Cy5), and then, the samples are resolved simultaneously within the same 2DE. This technique allows the introduction of a third sample "an internal standard" (a pool of all the samples included in the study), which is labeled with a third dye (Cy2), thereby allowing a quantitative analysis of the different gels. The clear advantage of DIGE is that the different samples are subjected to the same procedures, reducing experimental variability and producing separate 2D images of the same gel (Figure 16).

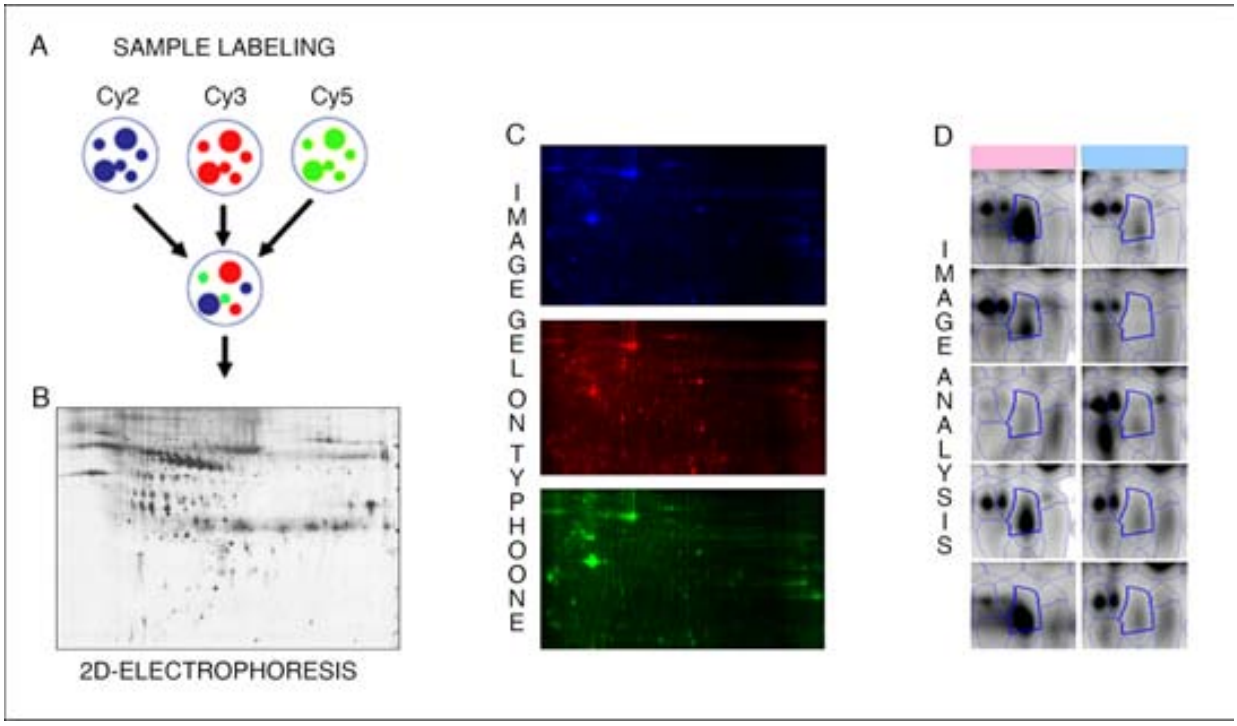



**Figure 16: Differential In-Gel Electrophoresis (DIGE).** (A) Sample labeling with different dyes. (B) 2DE of the different labeled samples in the same gel. (C) Typhoone scanner using different wavelength to obtain the different images from the same gel. (D) Image analysis using specific software (Progenesis Samespots v2.0 software (NonLinear Dynamics, Newcastle, U.K.)).

Although DIGE overcomes the problems associated with 2DE, it is still considered a low throughput, time consuming and expensive research oriented technology. On average, it takes days to complete a study on one biological sample. It can be used, however, in certain cases for biomarker discovery, where the high throughput processing of samples is not required. Moreover, the standardization of protocols for gel electrophoresis, in order to generate reproducible results, has been undertaken. These efforts to standardize protocols include the creation of a normal urine protein map to ensure the reproducibility of 2DE results across laboratories [344].

### 5.1.2. LIQUID CHROMATOGRAPHY COUPLED TO MASS SPECTROMETRY

LC coupled to MS (LC-MS) based proteomics has emerged as the most effective method for studying complex proteomes. LC provides a powerful fractionation method that is compatible with any mass spectrometer. This method separates large quantities of analytes on an LC column and offers high sensitivity [339]. In an MS experiment, the intensity of the signal, as a peptide (ion) elutes from the chromatographic column, can be plotted over time. The area under the curve (spectrum) represents the extracted ion current (XIC) and, for the same peptide and under the same experimental conditions, is linearly related to its amount. The XICs of different peptides from the same protein are also very different [345].

Not all peptides are able to be analyzed by MS. Large proteins (above 10-20 kDa) cannot generally be analyzed; instead, they must first be cleaved by a protease (for example, digested with trypsin) and then, the subsequent peptides may be analyzed [339]. The proteins representing a proteome, or a subset thereof, are enzymatically digested to generate peptides, which in turn may be analyzed by LC-MS [337, 346]. LC-MS is a powerful technique that may be used for many applications. It has very high sensitivity and selectivity. Generally, its application is oriented towards the specific detection and potential identification of chemicals in the presence of other chemicals (in complex mixtures, such as body fluids). The main limitation of LC-MS includes difficulties with the comparative analysis, in part because of the variability of the multidimensional separation and the substantial time required for analysis of a single sample [339].

### 5.1.3. PROTEIN IDENTIFICATION USING MASS SPECTROMETRY

The proteolytic digestion of separated proteins into peptides and the mass analysis of these peptides provide a peptide mass fingerprint (PMF), which can be researched against the theoretical fingerprints of sequences in protein databases [214]. The data can be used to search for protein sequences and nucleotides in the databases and analyze them by the application of algorithms, comparing them with theoretical production spectra of proteins in the database [347]. Protein identification is nearly effortless, as only two unique peptides are usually required to recognize a protein. However, the failure to identify or detect a peptide does not necessarily mean that the protein is absent, since the peptides may simply be below the threshold of detection [345]. A protein's score is a sum of the identification scores of its peptides, and one might surmise that a high protein-identification score would correlate with the higher abundance of that protein [348].

*Time-of-Flight Mass Spectrometry (TOF-MS):* TOF-MS is MS in which ions are accelerated by an electric field of known strength. In TOF analyzers, the m/z of an analyte ion is deduced from its flight time through a tube of a specified length, which is under a vacuum. The m/z ratio of an ion is proportional to the square of its drift time [337]. Sample preparation is critical to this approach. The protocols must be optimized, in order to reduce the variability and gain representative data by acquiring and averaging many single-shot spectra from several positions within a given sample spot [214].

*Matrix Assisted Laser Desorption Ionization-Time-of-Flight (MALDI-TOF):* Ionization by MALDI [349] involves a protein suspended or dissolved in a crystalline structure of small, organic and UV-absorbing molecules. The crystal absorbs energy at the same wavelength as the laser, which is used to ionize the protein or peptide. The laser energy strikes the matrix causing the

rapid excitation of the matrix and causing the matrix and ions to pass into a gaseous phase. MALDI is able to analyze proteins down to mole quantities and can tolerate small amounts of contaminant. The information obtained by MALDI analysis can be automatically submitted to a database search.

*Surface-enhanced Laser Desorption Ionization-Time-of-Flight (SELDI):* SELDI [350] technology reduces the complexity of the sample by the selective adsorption of proteins to different surfaces. Proteins bind to a specific surface with varying degrees of selectivity, while the unbound sample is washed away. A matrix that absorbs energy and allows the vaporization and ionization of the sample by laser is added. The sample is analyzed by MS, which provides a "low resolution" fingerprint [339]. Differentially expressed proteins may be determined from these protein profiles by comparing peak intensities. This technique can be used for protein purification, expression profiling, or protein interaction profiling with a small number of samples. The power of this technology is the integration of on-chip selective capture, relative quantification, and partial characterization of protein peptides [214]. An algorithm can then be used to identify trends in the m/z ratio peaks across a spectrum of proteins. These differential expression profiles may then be used as biomarker candidates for specific disease types, as has been performed for PCa [351].

Both, MALDI and SELDI techniques are attractive for disease biomarker identification. The differential expression of proteins analyzed by MALDI- and SELDI-TOF-MS can be used to determine the approximate size of a putative biomarker. However, the major weaknesses of these methods include a lack of fragmentation capability for peptide sequencing (protein identification), a lack of specificity due to a matrix effect in the low mass region (m/z 300-600), and a lack of accurate quantitation that requires careful optimization of the experimental parameters [214]. Nowadays SELDI technology is rearely used.

## 5.2 QUANTITATIVE PROTEOMICS FOR BIOMARKER VERIFICATION AND VALIDATION

Along with protein identification, protein quantification is another main component of the biomarker discovery process for determining changes in protein expression between the disease and control states. Quantitative proteomics have been developed to explore the dynamics of whole proteomes and generally provide a relative comparison of data between a few samples[352].

Traditionally, the *ELISA* [353] has been the "gold standard" used for targeted quantitation of a protein, providing good sensitivity and throughput. The ELISA technique allows the easy

detection of single proteins. Another powerful tool for studying the biochemical activities of proteins and potentially identifying disease biomarkers and therapeutic targets are *protein microarrays*. This technique can be considered a modern version of the old method for immune detection of multiple proteins. The samples are hybridized to a surface, where specific antibodies or antigens are imprinted, normally including several hundred targets[354]. Finally, *immunohistochemistry (IHC)*, which was first described in the 1940s by Coons *et al*, [355], is used to determine the location and distribution of a specific protein in the tissue. IHC can be used in different situations: (1) the histologic diagnosis of morphologically non-differentiated neoplasias, (2) the subtyping of neoplasias, (3) the characterization of the primary sites of malignant neoplasias, (4) the research for prognosis factors and therapeutic indications for some diseases, and (5) the discrimination between the benign and malignant natures of certain cell proliferations, as well as the identification of structures, organisms and materials secreted by cells.

BOX IX: PITFALLS OF ELISA, PROTEIN MICROARRAYS AND IMMUNOHISTOCHEMISTRY

The drawbacks of the ELISA assay are the following: (1) it usually measures a single biomarker per assay; (2) while the ELISA can recognize changes in protein concentration and/or its presence/absence, it is unable to detect PTMs; (3) 5-7h are required for the results; (4) its dynamic range is only reliable across 3 orders of magnitude; and, (5) it requires specific antibodies, which limit the speed and affordability of the development of an assay for emerging biomarkers. Clinically, the ELISA assay can give an FP or an FN result, resulting in an incorrect diagnosis.

Protein microarrays are often limited by the difficulty and complexity of protein purification, as well as by high throughput protein expression systems. There is a need for a specific probe for every protein to be analyzed, and the generally low density only allows for the detection of a few proteins. Finally, PTMs are usually not detected.

Limitations of IHC include: (1) the usefulness and contribution of IHC to solving pathological anatomy problems is directly proportionate to the experience of the hands that perform the reactions and the eyes that interpret the results; (2) the acquisition, handling, fixation and specimen delivery to the laboratory, as well as the antigen retrieval, are all critical variable factors that need to be considered; and (3) IHC is semi-quantitative.

Until very recently MS has been used almost exclusively in the discovery phase of the biomarker pipeline, whereas the qualification, verification, and validation of candidate biomarkers have traditionally been carried out by classic affinity methods [356]. Emerging target MS assays, such as SRM, which possess increased selectivity and sensitivity, have the potential to speed-up the extensive and time consuming biomarker rationale [357].

### 5.2.1 SELECTED REACTION MONITORING

The concept of targeted quantitative proteomics using MS-based quantitation strategies have been actively developed and implemented for preclinical proteomics applications, in order to help circumvent the throughput limitations and the complexity of proteomic analysis [214]. SRM is emerging as a technology that ideally complements the discovery capabilities of shotgun strategies through its unique potential for the reliable qualification and quantification of low abundance analytes in complex mixtures [357]. SRM has been the referenced quantitative technique for small molecule analysis for several decades. It is now emerging in proteomics as the ideal tool to complement shotgun qualitative studies; targeted SRM quantitative analysis offers high selectivity, sensitivity and a wide dynamic range [358].

The SRM technique was introduced in the late 1970's, along with the development of the first triple quadrupole mass spectrometers (QQQ-MS) by Enke and Yost [359] . When a QQQ-MS is operated in SRM mode, the first and the third quadrupole serve as mass filters to specifically selected predefined m/z values, corresponding to the precursor ion and a specific fragment ion of that precursor. The second quadrupole is used as a collision cell (Figure 17a). Initially this mode of operation was called SRM or Multiple Reaction Monitoring (MRM), depending on whether one precursor/fragment ion pair (transition) was monitored or a series of transitions were measured iteratively. Finally, the term of SRM was established, in order to avoid ambiguity [358].

The "peptide signature" of each protein is defined by those peptides (ions) that are unique to the proteome of an organism (proteotypic peptides or PTPs). In an SRM experiment, a series of transitions (precursor ion and fragment ions), in combination with the retention time of the targeted peptides, are monitored during a single LC-MS experiment [357] (see Box X and Figure 17b). Normally, from 2 to 3 PTPs and 3 to 4 fragment ions are monitored throughout the LC-MS, in order to get high specificity. SRM offers multiplexing detection capabilities with an excellent dynamic range. This approach potentially provides high selectivity and specificity, and it avoids most of the problems associated with the optimization of multiple assays in a single measurement.

Generally, the quantitative proteomic data is represented in two forms: the relative change in protein amounts between two states and the absolute amount of the proteins in a sample. In *relative quantification*, the amount of a substance is defined with relation to another measure of the same substance [345], normally between different stages, such as the benign and tumoral stages. *Absolute quantification* is the determination of the amount of the substance in question. In principle, absolute quantification encompasses relative comparisons; if the absolute amounts of the proteins are known in two samples, their relative ratios can be

calculated easily. Absolute quantification can be performed in the MS mode by comparing the extracted ion signal (peak height or peak area) of the internal standard peptide, which is normally and isotope labelled peptide (ILP) and the native forms of a given peptide [352].

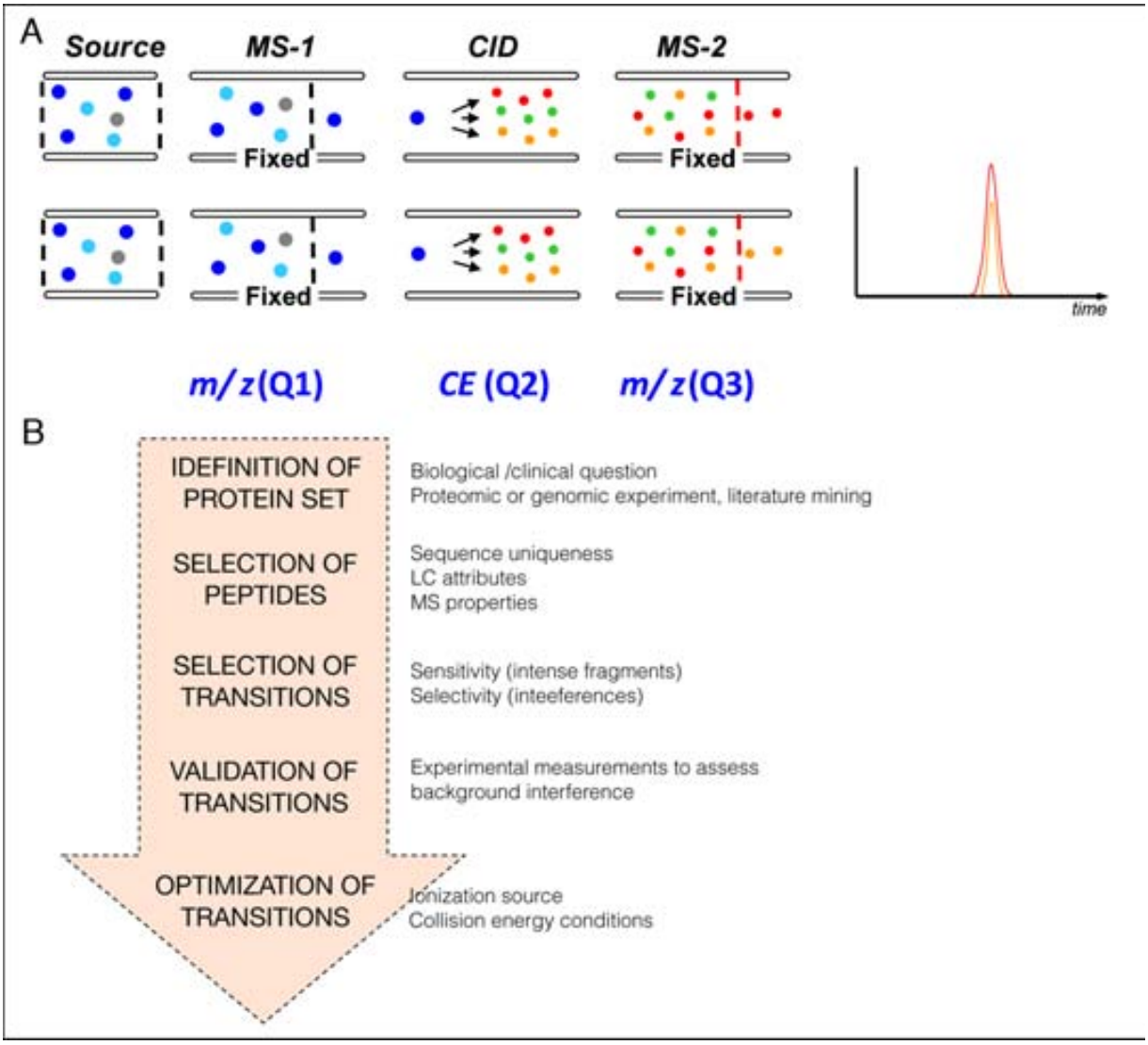



**Figure 17: SRM Workflow.** (A) Principle of the selected reaction monitoring performed on a triple quadruple mass spectrometer. The precursor ion selected by the first mass filter (Q1) enters the collision cell (Q2) where it undergoes collision-induced dissociation. One fragment ion, is then selected by the second mass filter (Q3). Multiple precursor/fragment ion pairs can be monitored sequentially within a measurement cycle. Adapted from (Domon and Aebersold 2006) [337]. (B): Workflow of an SRM-based proteomic experiment: Adapted from Galien *et al.* 2011 [358].

When SRM is coupled with the ILP, which is always present in the sample, it has several advantages for quantitative analysis. (1) It provides a useful confirmation that the instrument is functioning properly, as well as calculating more accurate ratios (endogenous peptide/ILP). One can be certain that the level of a peptide in the sample is below the detection limit if no signal is detected in the sample when the ILP is detected. (2) Ion suppression and matrix effects can be controlled, because the chemically identical ILP and the native forms of the

peptides will be suppressed to the same extent. (3) SRM analysis with ILP is capable of the sensitive and absolute determination of peptide concentrations across a wide dynamic range. (4) The amount of ILP added is different for each different protein and is a function of each protein's individual, relative abundance within a sample. The amount can be optimized to obtain a low coefficient of variancy (CV) of 5-10%, along with high sensitivity and precision, making it attractive for translational and clinical research.

BOX X: WORKFLOW OF SRM-BASED PROTEOMIC EXPERIMENTS;

(Adapted from Gallien *et al*, [358])

(1) After considering a specific biological or clinical question, a set of proteins of interest must be defined.

(2) For each protein a set of peptides representing each protein need to be determined. Normally, tryptic peptides result from the enzymatic digestion of the samples selected. A few representative peptides for each protein are targeted to infer its presence in a sample for quantification. Those selected peptides must be *PTPs*. The aa sequence uniquely associated with the proteins of interest are consistently observed in LC-MS analyses and correlate with good ionization efficiency and m/z within the practical mass range of the instrument. Since a single peptide only defines a small portion of a protein, ideally, several peptides should be selected for any given protein for distribution across the full sequence.

(3) It is critical to select transition ions that maximize the sensitivity and specificity of the SRM experiment. While sensitivity is related to the signal intensity of a transition, specificity is associated with interference from the co-eluting species that fall within the mass selection Q1 and Q3 analyzer windows. Current practice involves selecting the two to three most intense transitions for construction of the SRM assay.

(4) The experimental validation of the transitions should be evaluated in the context of the biological matrix, in order to account for the unspecific contributions of the fragment ions derived from co-eluting species with similar properties. When isotopically labelled reference peptides are available, checking the co-elution of the traces of the native peptides against those of their labelled counterparts represents the ultimate validation method. This method is also used to determine the Limit of Detection (LOD) and Limit of Quantification (LOQ) and is limited to experiments aiming at the quantification of a limited number of analytes.

(5) Quantification and high sensitivity are required, as well as an additional optimization of transitions. Signal intensities are determined by a combination of peptide ionization efficiency, peptide transfer into the analyzer and peptide disassociation into, ideally, a few intense fragments.

A calibration curve for each ILP made in the biological sample matrix and included with each batch of samples would permit the signal of each peptide in each sample to reflect the real concentration in a multiplexed fashion [358]. Incomplete digestion, due to a high concentration of matrix proteins, likely plays a large role in the reduced yield of peptides, along with other

sample preparation losses. It is for this reason that it is crucial to spike in an exogenous protein, preferably the stable ILP version of the native protein of interest, into each sample in order to gauge and normalize digestion efficiency [358]. However, the high cost and extensive efforts associated with the development of high quality reagents, including an ILP for each target protein and also internal standard controls, are limiting factors.

A current trend attempts to detect very large sets of peptides and estimate their abundance. SRM is being used as a directed discovery tool to screen for putative screening biomarkers and to investigate protein network pathways. The critical parameters driving such large-scale SRM experiments are the number of peptides to be analyzed, the number of transitions measured for each peptide, and the dwell time of each transition [358]. Recent developments in data acquisition techniques have enabled the analysis of a larger number of peptides by using the LC elution time as an additional constraint for monitoring the transitions of specific peptides in the corresponding time windows. This acquisition technique is called Time-scheduled SRM [360]. This method is divided into segments, during which only subsets of peptides are targeted. This acquisition mode increases the number of peptides monitored in one LC-MS analysis, while keeping the same sampling rate and the same degree of sensitivity and without compromising the dwell time.

Nevertheless, while SRM shows considerable promise for protein quantification, this technique is still in its early stages for clinical application.

(1) Accuracy and reproducibility continue to be issues. It is conceivable that analytical variability will be further reduced with the automation of sample preparation, while specificity will be improved by monitoring multiple transitions for a single peptide and multiple peptides from the same proteins.

(2) SRM has to overcome its technological obstacles if it is to establish itself as the method of choice for high molecular weight analytes, as it already has for low molecular weight species.

(3) The level of complexity of a tryptic digest of unfractionated plasma is several orders of magnitude greater than the equivalent samples, which are typically analyzed for metabolite quantitation [358].

Nonetheless, the prospects for the adoption of SRM as an essential element in the quantitative proteomics toolbox for the high throughput screening of patient samples in the qualification and verification stages are promising.

## 5.3. URINE IN CLINICAL PROTEOMICS

The contribution of proteomics to the understanding of the pathogenesis and diagnosis of diseases from the genitourinary tract has been considerable [361]. Urine has been defined as a fluid biopsy of the kidney and urogenital tract and provides much more information about these organs (including the prostate). A number of studies have clearly shown that urinary proteomics can be used not only for diagnosis, but also for kidney and non-kidney related disease prognosis[225]. Furthermore, as a filtrate, urine contains protein components that are similar to the ones that can be found in blood. Therefore, pathological changes in human organs, which may be found in blood, may also be reflected in the urinary proteome [223]. Thus, urinary proteomics becomes one of the most interesting sub-disciplines in the clinical proteomics area for biomarker discovery [338, 362, 363].

Urinary proteomics can be analyzed directly (whole urine) or separated by centrifugation into different fractions. Supernatants, from low-speed centrifugation, contain proteins that are derived from filtered plasma proteins and secreted by tubular epithelial cells. These supernatants can be further centrifuged (ultracentrifugation), yielding pellets with exosomes, small vesicles with cell membranes and cytosolic proteins [339].

Urine from a healthy individual contains a significant amount of peptides and proteins. One of the first attempts to define the urinary proteome was published in 2001 [364]. Using LC-MS, tryptic peptides of pooled urine samples were analyzed, and 124 proteins were identified. In 2004, this number increased to 1400 distinct spots on a two-dimensional electrophoresis gel, on which 420 identified spots yielded 150 unique protein annotations [365]. This number of identified urinary proteins increased significantly to around 1500 in 2006 by combining 2DGE and reverse-phase LC coupled with mass spectrometry (Orbitrap) [366]. In 2008, it was determined that the human urinary proteome apparently contains over 100,000 different peptides with at least 5000 high frequency peptides [367]. It is, therefore, safe to state that urine is indeed a rich, non-invasive source of potential biomarkers for disease that awaits further exploration.

Analysis of the urinary proteome could, therefore, permit the identification of biomarkers for both urogenital and systemic diseases. The ultimate goal of urinary proteomics is the identification and quantification of biomarkers related to specific diseases, mainly focusing on proximal tissues, such as the prostate. However, the obstacles for establishing true biomarkers are daunting, and most of the candidate biomarkers to date are not likely to be specific upon further scrutiny [368]. Clearly, proteomic technology needs to be improved in terms of sensitivity, throughput ability and quantitative accuracy [369].

The main practical challenges in the detection and quantification of proteins in urine by proteomic techniques are the high levels of salt and the wide dynamic range of protein concentrations spanning more than ten orders of magnitude [370]. It is known that the dynamic range of urine is important, as it is in plasma. The highly abundant proteins and their degradation products hamper the reliable detection of low-abundance components [371], which are frequent targets in biomarker development. Furthermore, since urine contains different high abundance proteins and the dynamic range is not as high as that of plasma, commercial plasma pretreatment depletion kits are not useful. Also, the protein content in urine is about a factor 1,000 less compared to that of plasma [366]. In consequence, it requires approximately 100-fold concentration.

Another important question that needs to be addressed in urinary proteomics is the variability of the proteome. Some studies have shown that, using low-resolution shotgun proteomics to identify proteins in urine collected from a number of individuals over the course of several days, only a few proteins were identified in common between those individuals [372]. Recently, Mann *et al,* studied the variability of the normal urinary proteome by using label-free quantitative proteomic techniques [369]. They found that the differences between technical replicates contributed to below 8% of the total variability. This was less than the contribution of either intra-individual (45%) or inter-individual (47%) variation to total variance. They also investigated the protein content of all of the samples analyzed and found that a large set of proteins were common in most of them ("the core urinary proteome"). They were able to identify 808 proteins and a full 587 of those proteins were found in each of the urine samples analyzed from each day. Moreover, they found that the added peptide intensity of each protein correlated well with the total amount of protein. Finally, they described a fraction of 20 proteins to be the most abundant proteins that taken all together contributed 2/3 of the total urinary content. The most dominant protein was human serum albumin (HAS), which represented 25% of the total amount. The dynamic range of this urinary core proteome spanned 5 orders of magnitude.

Last, but not least, is the task of resolving the lack of appropriate bioinformatic and statistical standardization procedures for data evaluation. Biomarker discovery projects generate large data sets. So far, no standards have been developed for data evaluation, resulting in a set of different solutions that may work well for only one particular problem [339]. The definition of common standards for data representation and the use of common formats and transparent tools for data analysis are crucial. Finally, the creation of data repositories is essential for the comparison, exchange, and sharing of data within the scientific community [373].

## 5.4. PROSTATE CANCER URINE PROTEOMICS

New technology platforms, which are currently being used in the discovery of novel PCa markers, will aid in the search for new markers; however, the use of appropriate study design and clinical-data analyses are key factors for obtaining results that are unbiased and reproducible. The high-throughput proteomic analysis of biological fluids, such urine, has recently become a popular approach for the identification of novel biomarkers, due to the non-invasive nature of samples collection and the reduced complexity of the serum [338]. However, only a limited number of studies have focused on PCa.

On several recent occasions, the detection of under-expressed PSA protein levels in urine has been reported [374-377]. Bolduc *et al*. compared a small cohort of urine samples (collected without previous PM) with benign, BPH and PCa and suggested that the ratio of serum PSA to urine PSA could have diagnostic utility [374]. The same idea was also suggested in another independent study [378]. In that study, PSA levels were also determined in urine [379], where no differences between urinary PSA pre- and post-PM were found. It was suggested that PSA may be a useful tool for the detection of local recurrence. Drake *et al*. [236] performed a study in which they focused on the characterization of PSA and PAP by using an ELISA assay on EPS samples. They found a clear trend towards lower levels of expression for both proteins in non-cancer samples.

One of the first proteomic urine profiling experiments for the detection of PCa were performed by Rehman *et al*. using a gel-based proteomic strategy that compared PCa samples with BPH samples [380]. They identified calgranulin B (MRP-14). However, this data was not able to be verified in an independent study. More recently, several studies have focused on the characterization of urine samples in a high-throughput manner. Teodorescu *et al*. performed a pilot study for PCa using CE-MS and were able to define a potential urinary polypeptide pattern with 92% sensitivity and 96% specificity [381]. However, again, these data could not be verified on a larger cohort. Later, the same group described a refinement of the PCa specific biomarker pattern using samples from 51 PCa and 35 BPH patients [382]. The model, containing 12 potential biomarkers, resulted in the correct classification of 89% of the PCa cases and 51% of the BPH cases in a second blind cohort of 213 samples. The inclusion of age and fPSA increased the sensitivity and specificity to 91% and 69%, respectively. M'Koma and collaborators performed a large-scale urinary proteomic analysis among samples of BPH, HGPIN and PCa [383]. Using MALDI-TOF, the group reported 71.2% specificity and 67.4% sensitivity for discriminating between PCa and BPH, while they also reported a specificity of 73.6% and a sensitivity of 69.2% for discriminating between BPH and HGPIN. Finally, Okamoto *and coll*. used a SELDI-TOF MS approach to analyze post-PM urine samples,

obtaining a heat map of 72 peaks, which was able to distinguish PCa from benign lesions with a sensitivity of 91.7% and a specificity of 83.3% [384]. The major study in the field of PCa metabolomics was performed by Seekumar *et al.* by mapping the differential metabolomic profile between tissues [261]. They identified a profile that was able to distinguish between benign tissue, clinically localized PCa and metastatic tissue. 1,126 metabolites were analyzed among 262 clinical samples, which included 110 urine samples. This study was conducted using liquid and gas chromatography MS. Sarcosine and the N-methyl derivative of the acid glycine were found at highly increased levels in PCa and were associated with disease progression to metastasis.

In conclusion, in the search for urinary biomarkers for PCa, a number of factors relating to methodology need to be taken into account. In particular, urine collection must be standardized. Although it has been demonstrated that post-PM urine sampling represents a rich source of biomarkers for PCa, there is also some evidence that this procedure can be avoided. Moreover, urine can be obtained in any urology clinic and does not represent any change in routine clinical practices. However, only a few studies have been conducted at the protein level, mainly due to the complexity of samples and the lack of standardization procedures for proteomic analysis. In addition, currently no potential internal standard has been identified at the protein level [236].

For all of the reasons mentioned above, any prospective proteomics studies with the goal of identifying and verifying new, specific and sensitive protein biomarkers for PCa offer an exciting challenge.

# HIPOTHESIS & OBJECTIVES

As small portion of normal prostate cells and their products continuously disseminate from the prostate epithelium and can be found in urine and;

As small portion of cancer prostate cells and their products continuously disseminate from the prostate epithelium and can be found in urine and;

As prostatic massage leads to an enrichment of prostatic fluid and prostatic cells in the first urine catch after prostate massage;

Thus the urine may reflect a broad variety of pathological conditions of the prostate and represents a more reliable source of biomarkers than serum or other more distal body fluids.

*We hypothesized that the utilization of targeted genomic and proteomic techniques on urine samples from patients suspected of having PCa can provide a pattern of biomarkers able to efficiently distinguish between the presence or absence of a prostate carcinoma and, further, can help to identify clinically significant prostate cancer patients*.

## GENERAL OBJECTIVES

*Short term objectives:*

- Diagnose asymptomatic prostate cancer by non/minimally-invasive means using RNA or Protein in urine after prostate massage
- Overcome the low specificity of PSA by the use of additional biomarkers to reduce the number of unnecessary biopsies (reduce financial costs for society, reduction in unwanted secondary effects)

*Long term objectives outside the duration of this thesis:*

- Identify clinical significant prostate cancers
- Increase the survival of patients (diagnosis of early-stage disease will be substantially better than that of patients who receive a diagnosis of late-stage disease)
- Reduce treatment costs (the detection of prostate cancer at an earlier stage will improve prostate cancer detection and staging, which will, in turn, help to reduce health care costs).

## SPECIFIC OBJECTIVES

*1. Objectives of the transcriptomic approach:*

1a. Characterization of new urine diagnostic marker for PCa; "PSGR and PCA3 as biomarkers for the detection of Prostate Cancer in urine" (Publication 1 and Annex 1)

1b. Improving the specificity of PSA for the detection of PCa using a urine multiplex biomarker model; "A three-gene panel on urine increases PSA specificity in the detection of Prostate Cancer" (Publication 2 and Annex 1)

1c. Characterization of the behaviour of PCA3 in HGPIN; "Behaviour of PCA3 gene in the urine of men with high grade prostatic intraepithelial neoplasia" (Annex 2)

*2. Objectives of the proteomic approach:*

2a. Discovery and qualification of new urine biomarkers for PCa; "The Discovery and Qualification of a Panel of Urine Biomarkers for Prostate Cancer Diagnosis" (Annex 3)

2b. Verification of urine PCa candidate biomarkers; "Qualification and Verification of Prostate Cancer candidate biomarkers in urine using Selected Reaction Monitoring approach"

(Annex 4)

# GLOBAL SUMMARY

# 1. TRANSCRIPTOMIC APPROACH

In recent years, the explosion of genomic and transcriptomic approaches to research have resulted in increased biomarker discovery. However, their translation into clinical utility has been limited, possibly due to a lack of samples analyzed in validation studies. It may also be due to the fact that traditional approaches tend to be limited to a single biomarker when, indeed, panels of biomarkers would be more useful. The recent discovery of PCA3 in urine as a biomarker for the detection of PCa and studies to determine its applicability in routine diagnosis represent a significant success for the scientific community in this field. However, implementing the use of this biomarker in a screening method for PCa is still not available to everyone, principally due to the high cost of the existing test.

In this study, we aimed to follow the same strategy of using urines as a source of biomarkers and as a minimally invasive PCa diagnostic technique. First, we wanted to characterize a new urine candidate biomarker (PSGR) to be compared with PCA3, and second, we planned to use a panel of biomarkers, in order to improve diagnostic accuracy. Finally, we proposed to better characterize the well-known biomarker PCA3 as a tool for the early detection of pre-neoplastic PCa lesions, such as HGPIN.

## 1a. Characterization of New Urine Diagnostic Biomarker for PCa; “PSGR and PCA3 as biomarkers for the detection of Prostate Cancer in urine”

PSGR is a member of the G-protein coupled OR family. PSGR has previously been described to be highly prostate tissue-specific and over-expressed in PCa tissue. Our aim was to test whether PSGR could also be detected by RTqPCR in urine sediment obtained after PM. A total of 215 urine samples were collected from consecutive patients (34% with PCa), who presented for PB due to elevated serum PSA levels (> 4 ng/mL) and/or an abnormal DRE. These samples were analyzed by RTqPCR. First, we analyzed PSGR alone and compared these findings to those of PCA3. By univariate analysis we found that PSGR and PCA3 were significant predictors of PCa. A ROC curve was used to assess the outcome predictive values of the individual biomarkers. We obtained the following AUC values: PSGR (0.68) and PCA3 (0.66). Both markers individually overcame the AUC value for serum PSA (0.60). Finally, we combined those markers to test if a combination of both biomarkers could improve the sensitivity of PCA3 alone. By using a multivariate extension analysis, multivariate ROC (MultiROC), the outcome predictive values of the paired biomarkers were assessed. We obtained an AUC value of 0.73 for the combination of PSGR and PCA3 (PSGRvPCA3). Then, we tested whether a combination of PSGR and PCA3 could improve specificity by fixing the

sensitivity at 95%. We obtained specificities of 15% (PSGR) and 17% (PCA3) for each individual marker and 34% for PSGRvPCA3. In summary, a multiplexed model that included PSGR and PCA3 improved the specificity for the detection of PCa, especially in the area of high sensitivity. This could be clinically useful for determining which patients should undergo biopsy.

## 1b. Improving the Specificity of PSA for the Detection of PCa using a Urine Multiplex Biomarker Model; *"A three-gene panel on urine increases PSA specificity for the detection of Prostate Cancer"*

Much evidence points to the fact that a single marker may not necessarily reflect the multifactorial and heterogeneous nature of PCa. The principle that underlies the combined biomarker approach is consistent with tests offered for the detection of PCa in tissue specimens and takes into consideration the heterogeneity of cancer development based on a diagnostic profile. The combined model that results from these combinations provides overall increased sensitivity without decreasing the specificity.

Our previous work demonstrated that by fixing the sensitivity at 95% and multiplexing PSGR and PCA3, we were able to double the specificity over that of each single biomarker. Following the same approach here, we combined three biomarkers to maximize individual specificities. PSMA, another well-known PCa biomarker, was used to test whether a combination of PSGR, PCA3 and PSMA was able to improve the specificity of the current diagnostic technique. We analyzed post-PM urine samples from 154 consecutive patients (37% with PCa), who presented for PB due to elevated serum PSA levels (**>** 4 ng/mL) and/or an abnormal DRE. We tested whether the putative PCa biomarkers PSMA, PSGR, and PCA3 could be detected by RTqPCR in the post-PM urine sediment. By univariate analysis, we found that the PSMA, PSGR, and PCA3 scores were significant predictors of PCa. We then combined these findings to test if a combination of these biomarkers could improve the specificity of an actual diagnosis. Using a multiplex model (PSGRvPCA3vPSMA), the area under the MultiROC curve (AUCm) was 0.74, 0.77 with PSA and 0.80 with PSAD. Fixing the sensitivity at 96%, we obtained a specificity of 34%, 34% with PSA and 40% with PSAD. Afterwards, we specifically tested our model for clinical usefulness in the PSA diagnostic ''gray zone'' (4–10 ng/mL) on a target subset of 82 men with no prior biopsy (34% with PCa) and a target subset of 77 men with the PSAD information (35% with PCa). Using a multiplex model, the AUCm was 0.82, 0.89 with PSAD. Fixing the sensitivity at 96%, we obtained a specificity of 50% and 62% with PSAD in the gray zone. This model would allow 34% of the patients to avoid unnecessary biopsies in the gray zone (42% when using PSAD). Taken

together, these results provide a strategy for the development of a more accurate model for PCa diagnosis. In the future, a multiplexed, urine-based diagnostic test for PCa with a higher specificity, but the same sensitivity as the serum-PSA test, could be used to better determine which patients should undergo biopsy.

## 1c. Characterization of the Behavior of PCA3 in HGPIN; "Behavior of PCA3 gene in the urine of men with high grade prostatic intraepithelial neoplasia"

An ideal biomarker for the early detection of PCa should also differentiate between men with isolated HGPIN and those with PCa. PCA3 is a highly specific PCa gene, and its score in post-PM urine seems to be useful in ruling out PCa, especially after a negative PB. The biopsy finding of an HGPIN is a frequent indication that the PB should be repeated. Some studies have correlated the urine PCA3 score with the presence of HGPIN at biopsy, yielding conflicting results. Today we know that PCA3 is expressed in the HGPIN lesions that surround PCa and also that the PCA3 score seems to be higher in men with HGPIN.

The aim of this study was to determine the efficacy of post-PM urine PCA3 scores for ruling out PCa in men with previous HGPIN. The PCA3 score was assessed by RTqPCR in 244 post-PM urine samples collected from men subjected to PB: 64 men with an isolated HGPIN (no cancer detected after two or more repeated biopsies), 83 men with PCa and 97 men with benign pathology findings (BP: no PCa, HGPIN, or ASAP). The median PCA3 score was 1.56 in men with BP, 2.01 in men with isolated HGPIN and 9.06 in men with PCa. A significant difference was observed among the three scores ($p < 0.001$) and also between HGPIN and PCa ($p = 0.008$); however, no differences were observed between HGPIN and BP ($p = 0.128$). The AUC in the ROC analysis was 0.71 in the subset of men with BP and PCa, while it decreased to 0.63 when only men with isolated HGPIN and PCa were included in the analysis. Finally, the median of the PCA3 scores was assessed in men with previously diagnosed unifocal HGPIN (2.63) and in men with previously diagnosed multifocal HGPIN (1.59). No differences were observed between unifocal and multifocal HGPIN ($p = 0.56$). In conclusion, the efficacy of post-PM urine PCA3 scores in ruling out PCa in men with HGPIN is less than in men with BP. For this reason, when HGPIN is found at PB, these results should be taken into consideration, in order to establish the clinical usefulness of the PCA3 score as a tool for avoiding unnecessary repeated biopsies.

# 2. PROTEOMIC APPROACH

In order to improve screening techniques and cancer prognosis, urologists need biomarkers that can provide better information than is afforded them by the histology of prostate tissue and by PSA, which is currently obtained from the blood. Moreover, the idea of getting this information via a minimally/non-invasive way is very attractive.

Urine proteomics, and additionally, urine PCa proteomics can help us to improve these diagnostic and prognostic drawbacks. The high-throughput proteomic analysis of urine samples has recently become a popular approach for the identification of novel biomarkers. However, the use of appropriate study designs and clinical-data analyses are key factors for obtaining results that are unbiased and reproducible.

To date, only a few studies have used urine for the discovery of PCa, protein-based biomarkers; however, the results regarding biomarker discovery are very promising. To our knowledge, no protein-based urine biomarker has yet appeared on the market.

In our studies (data not published, but included in the thesis annexes 3 and 4), we attempted to follow the "new biomarker pipeline" described in section 4, so that at the end of the process, should we obtain a candidate biomarker or a panel of candidate biomarkers for PCa diagnosis, we would be closer to success than to failure (Figure 18).

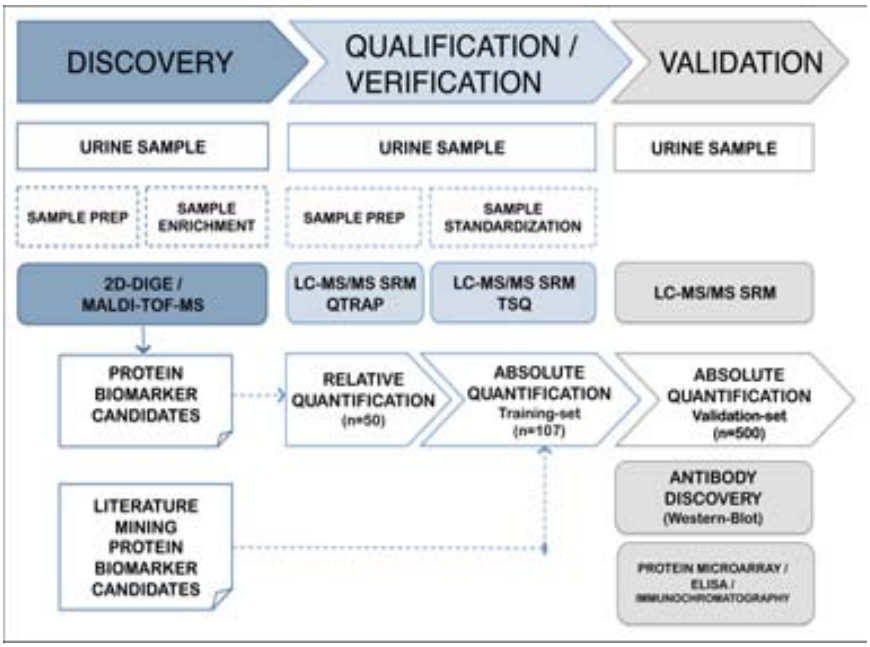



**Figure 18: Proteomic Analysis Workflow.** Discovery phase was performed with and without urine sample pre-treatment (light blue). The candidate biomarkers were by using SRM-based assay on a cohort of 50 urine samples (dark blue). Qualified proteins and other candidate were qualified and verified by using SRM-based assay in a total cohort of 107 urine samples (dark blue). Finally, credientialized proteins passing these steps will be good candidates to be introduced in a larger validation study (not contemplated within this work-gray part).

## 2a. Discovery of New Urine Biomarkers for Prostate Cancer; "The Discovery and Qualification of a Panel of Urine Biomarkers for Prostate Cancer Diagnosis"

The use of biomarkers is a main focus in the detection of presymptomatic PCa. Proteins secreted by cancer cells, also referred to as "cancer cell secretomes," are a promising source for biomarker discovery. A great advantage to these cancer-secreted proteins and/or their fragments is that in most cases, they enter body fluids, such as blood or urine, and therefore, can be measured via non-invasive assays. Since the protein products of PCa cells can be detected in urine, their use as a proximal body fluid in the detection of PCa is very attractive.

In this study we used DIGE proteomic analysis on 30 age-matched, post-PM urine supernatant specimens, in order to identify the differentially expressed proteins in patients with PCa. 24 potential biomarkers were identified, the majority of which were secreted proteins associated with several well-known, functional cancer pathways, such as NFκB, PDGFBβ, or β-catenin. Qualification of 15 of the 24 identified biomarker candidates was then undertaken by relative quantification using an SRM-based assay on 50 post-PM urine supernatant samples (38% with PCa). After statistical analysis, 7 peptides, corresponding to 5 different proteins, were selected. A multiplex ROC curve using those 7 peptides showed an AUC value of 0.93. Fixing the sensitivity at 95%, we achieved a specificity of 78%. This data demonstrated that proteomic analyses are able to reveal novel biomarkers/diagnostic profiles for PCa in urine. This constitutes an important step towards advancing the accurate diagnosis of PCa, which currently represents a major setback in our ability to cure patients who suffer from the disease.

## 2b. Qualification and Verification of Urine Prostate Cancer Candidate Biomarkers; "Qualification and Verification of Prostate Cancer candidate biomarkers in urine using Selected Reaction Monitoring approach"

Urine has been defined as a liquid biopsy of the urogenital tract, and it can provide much more information about these organs (including the prostate) than a tissue biopsy. The qualification and verification of candidate biomarkers is a critical stage in the great biomarker discovery pipeline. Credentialed biomarkers that have successfully passed through this stage are considered verified biomarkers, which are of high value for translation into large-scale, clinical validation studies. The evaluation of biomarkers in body fluids necessitates the development of robust methods to quantify proteins in body fluids, using large sets of samples. SRM is emerging as a technology that ideally complements the discovery capabilities of shotgun strategies through its unique potential for the reliable

qualification and quantification of low abundance analytes in complex biological samples, such as urine samples. However, approaches to biomarker discovery using urine have been hindered by concerns about reproducibility and an inadequate standardization of proteomics protocols. Reproducible procedures for the preparation of protein samples isolated from human urine are essential for meaningful proteomic analyses.

In the present study, we performed the qualification of a set of 42 candidate biomarkers for PCa diagnosis on a set of 107 post-PM urine supernatnat samples (36% with PCa) using SRM-based absolute quantification. Before that, urine sample preparation and analytical procedures were optimized for SRM methodology. We standardized preparation of the urine protein samples for SRM analysis by using 9 different protocols. We "spiked" our samples with exogenous yeast and bovine proteins to monitor and optimize protein extraction and digestion conditions. Each protocol was repeated for 5 separate days by 2 different individuals. Our goal was to obtain a panel of biomarkers that alone, or in combination with the existing PCa biomarker, would help us to better define patients with PCa. In addition, due to the large number of samples and their pathological conditions, we would also be able to define candidate prognostic markers. However, this study has yet to be completed.

# DISCUSSION

PCa is currently the second most common cause of cancer death in men [48]. The introduction of an effective blood test (PSA) has resulted in a more early-stage PCa diagnosis, when potentially curative treatment options should be more effective. Nevertheless, in spite of this early detection, PCa mortality has not significantly decreased in recent years. Despite significant progress in the investigation of PCa biomarkers, some men are still over-diagnosed with indolent PCa, while others die from aggressive disease that has been diagnosed too late. Consequently, in order to avoid over-treatment, the same stage of PCa may require different treatment strategies, based on an individual patient's life expectancy. Whether or not PSA screening is found to benefit the population at large, PSA will continue to be used in in caring for the individual PCa patient [161]. Nowadays, PCa is detected in only about one-third of all men biopsied [385]. As a result of their persistent serum PSA levels, but negative biopsy results, these men undergo repeated PB. This situation is referred to as the "diagnostic dilemma." Methods to enhance PSA specificity have assisted clinicians in deciding which patients should undergo biopsy; even so, there has been no evidence to suggest that any of these methods have improved diagnostic accuracy or facilitated optimal therapeutic decision-making. Due to the widespread use of the PSA test, it will be difficult to change the actual clinical routine; however, additional testing methods, in which have the ability to clarify the PSA gray zone and are taken in combination with PSA, could be useful in determining which patients should undergo biopsy. Actually, most of the current diagnosis and treatment algorithms are based on PSA levels [201]. Furthermore, there is an important need for biomarkers that are able to distinguish between aggressive and non-aggressive cancers and that can help to avoid actual over-treatment. A growing body of data suggests that individual DNA-, RNA-, and protein-based urine biomarkers hold promise in these settings. However, as a single biomarker does not necessarily fulfill the requirements needed to reliably detect a disease as early as possible, to unambiguously distinguish it from other pathological conditions, and to monitor the efficacy of therapy, an alternative strategy must be used that combines several markers. These markers may not present high specificity and sensitivity on their own, but taken together within a panel, these markers will be able to effectively work in concert [231].

Because the first portion of a voided urine sample contains the highest concentration of prostatic and urethral secretions [386], post-PM voided urine samples represent the best compromise between a minimally invasive technique and the possibility of obtaining enough material for a correct diagnosis. Since PM is part of the standard diagnostic tripod in PCa, urine specimens can be easily obtained after the routine examination process. Urine-based assays can monitor PCa with heterogeneous foci and provide a non-invasive alternative to

multiple biopsies [387]. Even in case that urine assays cannot detect cancers, which do no shed tumor cells into urine, they still deserve considerable attention. The combination of multiple urine biomarkers can be of special value to men who have persistently elevated serum PSA levels and a history of negative biopsies [387].

In order to define the disease-specific urinary biomarker(s), it is crucial to include the appropriate controls. Urine samples obtained from patients with other diseases or disorders, which have clinical, biochemical and metabolic profiles similar to those of the disease of interest, must be included as controls [388]. Here, we used urine samples from patients subjected to PB, due to their increased serum PSA levels and/or an abnormal diagnostic DRE. In our selection process, we excluded other types of cancer; thus, our study population was representative of the population at large. We used age-matched specimens as benign samples, due to the fact that these patients had the same clinical backgrounds found throughout this age group, such as BPH, prostatitis, atrophy, and HGPIN (with the only difference being the presence or absence of PCa).

A promising result in the field of improving PCa diagnosis has been the recent discovery of PCA3 as a useful marker in an RNA-based, non-invasive urine test. This is especially valuable for men with chronically elevated serum PSA levels, when considering whether there is a need for repeat biopsy [297]. In 2006, a urine-based, quantitative test was developed for European countries (EMA approval), the APTIMA® assay [245] (Gen-Probe, San Diego, CA, USA; PROGENSA-TM). However, a clinical trial is still being conducted in the United States to obtain US FDA approval. The clinical performance and the diagnostic accuracy of PCA3 have been evaluated in several studies to prove its diagnostic utility. PCA3 consistently outperforms PSA in diagnostic accuracy and also improves upon serum PSA due to its higher specificity [302]. Some potential applications of the PCA3 assay include testing prior to first biopsy and deciding whether or not to repeat a biopsy in men with elevated serum PSA levels and previous negative biopsies [302] [297]. However, due to the long-term, widespread use of serum PSA, PSA will likely still be used for years to come [302]. Another important point that needs to be addressed is the high cost of this new assay and the fact that this system is a closed system, which does not permit the user to test on parallel additional biomarkers. In contrast, the methodology that we used in our studies allowed the combination of several biomarkers; however, in terms of handling, our test is still a long way from becoming a routine part of clinical practice.

In order to compare the gold standard PCA3 urine test to other urine-based RNA candidates, it is necessary to note that since different studies use different methodologies for the detection of transcripts, drawing conclusions is not easy. Moreover, it would be

inappropriate to directly compare the results obtained from urine samples collected from different populations with different disease prevalence. In our first study (publication 1), we wanted to know if PSGR, a prostate-specific integral membrane protein that was described as over-expressed in PCa tissues, could serve as a urine-based RNA biomarker for PCa. We performed an RTqPCR urine-based assay on 215 patients who presented for PB (34% with PCa), in which the relative quantification of transcripts for PSGR was achieved by using a ratio $C_t$(PSGR)/$C_t$(PSA)*1000. By using the same experimental settings, PCA3 transcripts were also quantified. Urine contains not only cells from the prostate, but also cells from the kidney, bladder, transitional epithelium, blood, and also from the skin; for this reason, patient samples presented varying quantities of organ-derived cells, and therefore, the selection of a prostate-specific "housekeeping" gene was crucial. Since PSA is only expressed in prostate tissue, and the PSA mRNA levels in the prostate cells released in urine are generally unrelated to the PSA protein levels found in the blood, and they remain essentially unchanged in PCa [389], we used the mRNA of PSA as a "housekeeping" gene [244]. Thus, the quantification of PSA mRNA was required to normalize the total mRNA present in the urine samples. By ROC curve analysis we were able to show that PSGR (AUC 0.68) could represent an alternative or complement to PCA3 testing (AUC 0.66) (Table 11). In order to translate these findings to the clinics, we needed to find the way to create a test to pinpoint PCa patients. A false negative is much worse than a false positive, since the consequence of missing a cancer case may be fatal to the patient [390]. By fixing the sensitivity at 95%, we achieved specificities of 15% for PSGR and 17% for PCA3. Furthermore, when we combined both biomarkers using a multiROC approach, the results were better in terms of both biomarker performance (AUC 0.73) and specificity (34%) (Table 11). Finally, we analyzed PSGR and PCA3, in order to test their ability to predict clinical risk groups. However, we could not observe any relationship between the different clinical groups (annex 1). In conclusion, PSGR had a similar performance to the gold standard PCA3, however, future studies will need to be conducted to further improve the performance of this test. The improved performance of this test will fill important medical needs, such as helping to improve decision making for repeat biopsies in men who have elevated PSA levels and negative biopsies, resolving the so-called ''PSA dilemma'' and ruling out clinically significant PCa.

The second aim of this study was to better improve PSA specificity by multiplexing different urine-based RNA biomarkers (publication 2). By using an RTqPCR urine-based assay on 154 patients who presented for PB (37% with PCa), we demonstrated that PSMA, a marker also previously described to be over-expressed in PCa, has a similar performance (AUC 0.62) to the known urine markers PSGR (AUC 0.65) and PCA3 (AUC 0.60). In order to improve the

diagnostic efficacy of these biomarkers one by one, we combined PSMAvPSGRvPCA3 (3M) by using a multiROC approach. This analysis improved other combination methods, since it retained all of the simplicity of traditional ROC curve analysis interpretations, while additionally allowing comparisons between the performances of multivariate combinations. These comparisons were not restricted to the display of a single variable's performance or the comparison of individual tests [391]. Moreover, this kind of analysis has the advantage of not having the combinations of sensitivity and specificity fixed in advance. A major concern in creating a ROC curve to represent the performance of a combination of biomarkers is over-fitting. To control this bias, we used an approach based on the split-half (50%) method to validate the results [392].

Using the multi ROC approach, we found that the AUCm (0.74) was notably improved, compared to the individual AUC values. Moreover, if we combined 3M with serum PSA to a newly combined marker, the AUCm (0.77) was even better (Table 11). Because of the multifactorial nature of cancer, it is possible that marker A could be positive in one patient and marker B in another. Using our multiROC model, the combination of both markers was linear, assigning a single weight to each marker individually. The combined marker was declared positive if at least one of the scores was above its detection threshold. By fixing the sensitivity to the clinically interesting value range of 96%, the specificities for the individual markers dropped to almost 0%. However, the 3M model maintained a specificity of 34%, while no further increase was observed when 3MvPSA (34%) was combined (Table 11).

The next step was to evaluate the behavior of these three markers in the PSA diagnostic "gray zone" of 4-10 ng/mL when no prior biopsy information was available. This sub-population is of particular interest, due to the lack of previous biopsy information and also because serum PSA levels between 4-10 ng/mL do not provide a clear diagnosis of PCa and normally those patients undergone repeated biopsies to rule out PCa. A total of 82 patients were included in this subgroup of clinical interest (34% with PCa). We demonstrated that the prediction of PCa improved significantly for PSMA (AUC 0.74), while PSGR (AUC 0.66) and PCA3 (AUC 0.61) showed a similar performance. For the combined model (3M), we found an AUCm of 0.82, while there seemed to be no further advantage in using 3M in combination with PSA (Table 11). By fixing the sensitivity at 96%, the specificity of the individual markers was near 0%, while the combined model maintained a specificity of 50%. Moreover, in this clinically interesting sub-group of patients, the combination of the 3M and PSAD values (AUCm 0.89) significantly increased the specificity to 62%, when the sensitivity was fixed at 96% (Table 11) (annex 1). Although there was a significant increase, we did not take this result into consideration, since PSAD requires a transrectal ultrasound ecography, which is

expensive, time-consuming and causes significant patient discomfort. Therefore, this test could no longer be considered non-invasive.

In conclusion, our combined model results provided overall increased sensitivity without decreasing the specificity. Translated to the clinics, we achieved the same high sensitivity as that of the serum PSA-test alone (96%), but we increased the specificity considerably. By using this procedure, approximately 34% (Table 11) of the biopsies performed on patients in the "gray zone" could have been prevented. In addition, as we achieved a negative predictive value (NPV) of 96% (table 11), by using this test we would be able to be almost sure that the patient who will not been scheduled for a biopsy, is a true negative patient. Another interesting benefit would be the improved ability to decide whether or not to perform a repeat biopsy. Since the probability of finding a cancer patient in a repeat biopsy is higher, the sensitivity could be lower, in order to get a higher specificity, which would then translate to a lower rate of false positives. However, in order to obtain accurate results, future studies will be needed on larger populations of men who have undergone repeat biopsies.

A possible limitation of these studies is the fact that patients who present with a negative biopsy usually tend to have fewer cells of prostatic origin in their post-PM urine than do their malignant counterparts. Therefore, those patients who do not yield enough RNA material to being analyzed would need to be excluded. A possible explanation could be the loss of cell-cell contact in cancer [393]. Another possible limitation of our studies (RNA- and protein-based) is that the definition of patients negative for PCa is based on their having presented a recent, negative PB. However, this definition can be problematic, since approximately 20% of the patients who have presented a negative biopsy will actually be diagnosed with PCa at a later date [159]. An accurate follow-up for the patients included in these studies would, therefore, be necessary.

The finding of HGPIN lesions is a frequent indication for the repetition of a biopsy. In early studies, using limited sextant biopsy schemes, HGPIN was associated with high rates of PCa. It was suggested that its presence would indicate an immediate repetition of the biopsy [394]. However, if a more extensive biopsy scheme is initially used, the cancer detection rate on early repeat biopsy may be considerably lower. This is due to the fact that the number of cores sampled during the initial biopsy affects the likelihood of detecting PCa in any subsequent biopsies [395]. For this reason, some investigators believe that repeat biopsies may be unnecessary in the current era and that these men can be followed using serial DREs and PSA measurements [396]. In fact, HGPIN lesions do not contribute to the serum concentrations of PSA or modify the percent-free PSA [397]. Despite the current body of knowledge, the

urologic community still has many doubts regarding the attitude towards the presence of an isolated HGPIN in a PB.

Because the PCA3 gene has emerged as a tool for selecting those men with negative biopsies, who require a repeat procedure, it seems important to better know its behavior in men with isolated cases of HGPIN and in those men with PCa. In our opinion, it is important because PCA3 has been considered a highly specific marker for PCa. For this study (annex 2), we selected a subset of patients with a high probability of having an isolated HGPIN, since PCa had not been detected after one or more repeated biopsies. As control groups, we selected men with PCa and men with benign pathological (BP) findings (no PCa, no HGPIN). We observed that when compared to those men with BP the PCA3 score was only slightly higher in men with HGPIN, while the score was also higher in men with PCa. By ROC curve analysis we speculated whether PCA3 could be effective for ruling out PCa, when men with diagnosed HGPIN were included in the study (AUC 0.63) and whether the results were similar to those observed when only men with normal biopsies were included (AUC 0.71). Our results demonstrated that the ability of post-PM urine PCA3 to identify men with PCa was lower when we compared men with HGPIN to those without HGPIN. Although we know that our studies need further valifation, we suggest that the PCA3 gene is not only a specific marker for PCa, since it is also over-expressed in HGPIN. Therefore, we propose a prospective study that considers men with HGPIN in their first biopsy, incorporating this finding into a new nomogram based on the PCA3 score.

Proteomics is regarded as a sister technology to genomics; however, although the pattern of gene activity may be abnormal in a tissue with a pathological lesion, there can be a poor correlation between the level of the transcription of different genes and the relative abundance of the corresponding proteins within the tissue. Consequently, the information about a pathological process that can be derived at the level of gene transcription is incomplete [325]. Proteomic approaches allow the discovery of disease-specific targets and biomarkers, providing comprehensive diagnostic and prognostic information. In addition to linking genomics with proteomics, we know that, since the human genome project was concluded, we have arround than 20,000-25,000 genes that may be simultaneously suffering from alternative splicing, which can give rise to different proteins. Moreover, each protein can undergo different processes, which can change their biological functions, such as PTM, truncations, proteolysis or compartmentalization. The exact number of different proteins is unknown. It is estimated that the human proteome could contain up to a million of different proteins. Future proteomics efforts will need to continue technology development, optimization and standardization [398], in order to discover the vast, unknown human

proteome. Incorporation of the most up-to-date and efficient technologies is critical in successfully propelling the translation of proteomics findings into clinically relevant biomarkers. There should be the continued, rigorous assessment of biospecimens and data quality through Quality Assessment (QA)/Quality Control (QC) criteria at each step of the biomarker development pipeline, in order to make "go" or "no go" decisions [214].

Since RNA and DNA biomarkers generally rely on cellular shedding into the urine and require the exfoliation of cells and subsequent analysis of those cells, the analysis of secreted protein products in urine could easily detect changes throughout the course of cancer [240]. The "secretome" is referred to as a biological fluid that may be enriched with secreted and/or shed proteins from adjacent, disease-relevant cancer cells [399]. Proteins and peptides "leaked" from tissues into clinically accessible body fluids, such as blood or urine, have led to the possibility of diagnosing the disease at an early, presymptomatic stage, as well as being able to monitor its responses to therapy through testing these body fluids for the presence of disease-related biomarkers [214]. Blood sampling activates proteases, which induce massive protein degradation [382]. In contrast, urine stays in the urinary bladder for hours and subsequent sampling does not trigger any reaction. Besides, the PCa secretome can easily be found in proximal prostate fluids, such as urine. Since PM leads to an enrichment of prostatic fluid in the first urine catch after this procedure, urine is an ideal clinical sample for biomarker discovery. Moreover, urine is readily available in almost all patients, and its collection is very simple and non-invasive. Urinary proteomics has, thus, become one of the most interesting sub-disciplines in the clinical proteomics area. A number of studies have clearly shown that urinary proteomics can be used not only for diagnostics, but also for kidney and non-kidney disease prognosis [225].

The two main obstacles encountered in the discovery of biomarkers in urine, as well as in plasma, are the following: the high dynamic range, which results in only the high-abundance proteins being displayed on any typical 2DE map, and the low reproducibility. Urine, as any other body fluid, contains some highly abundant proteins, such as albumin, Immunoglobulin (IgG) and uromodulin. Therefore, it is necessary to decrease the level of the most abundant proteins, by enriching the rarest ones. This must be accomplished in order to get a better representation of the low-abundance proteins, which may be the most interesting for biomarker discovery. Nowadays, there is still great debate regarding the performance of pretreatments that have been used on plasma samples, as well as on urine samples, for biomarker discovery [400]. The main problem is that complex matrix samples, such as urines, are a vast expanse of unexplored "waters," which scientists are not exactly sure how to explore [401]. The other main obstacle mentioned is the low reproducibility of 2DE, which has

mostly been overcome by using the DIGE technique. Currently, LC-MS technologies are expanding their horizons to be introduced in the first steps of the biomarker pipeline. In our discovery phase, we combined data obtained from non-pretreated urine samples with data from pretreated urine samples. Pretreated samples were enriched/equalized using the ProteoMiner technique[234]. The ProteoMiner technique results in an important dilution of the most abundant species, resulting in a concurrent concentration of the diluted and rare ones. Although there is some discordance behind the use of this technology as a quantification technique, we opted to use it, since we obtained reproducible results in our trial experiments. In addition, this technique was only used in our discovery phase, while qualification and verification were performed directly on urine supernatants.

While proteomics holds great promise for biomarker development, until now it has been difficult to draw meaningful conclusions and translate them into clinical practice [214]. This is largely due to several factors: poor study designs, a lack of standardization in the preanalytical, analytical and post-analytical studies, a technological gap between discovery and clinical qualification in the biomarker development pipeline, and a lack of understanding among the stakeholders, who play different roles at every stage of the pipeline [214]. Thus, to efficiently translate proteomic technologies to the clinics, it is necessary to contemplate four stages in the biomarker development pipeline: discovery, qualification, verification, and finally, validation. Our aim was to follow this biomarker development pipeline in the discovery phase for identifying potential urine biomarkers for PCa (annex 3). By using 2D-DIGE coupled with MALDI-TOF-MS/MS, 24 potential biomarkers (15 down- and 9 up-regulated) were found differentially expressed in a significant manner in the urinary specimens of patients with PCa compared to the age-matched controls confirmed by PB (Discovery Phase). A critical factor that determines the selection of a candidate biomarker is the quality of the scientific and clinical results, such as linking the gene or protein function to the biology of the disease, relating the biomarker to the mentioned disease, variations in the different stages of the disease, response to therapy, and overall survival [220]. Thus, we performed a network analysis that showed that the majority of these identified proteins were secreted components of several well-known, functional cancer and inflammation networks, such as NFKB, PDGFBβ and β-Catenin. Moreover, due to the fact that we used urine supernatants, where we could find soluble secreted proteins, the majority of the identified proteins (62%) were localized in the extracellular space.

MS-based proteomics strategies have become an integral part of the biomarker development workflow, including a discovery phase and the subsequent qualification, verification, and validation of the candidates in body fluids [217]. Targeted proteomic approaches, such as the

SRM-based assay, are emerging in proteomics as ideal tools for complementing shotgun qualitative studies [402]. Our "candidate biomarkers" were first qualified/verified using an SRM-based approach. The reasoning behind this was the reduction of the candidate protein number for a second verification step using absolute quantification. SRM may be used to achieve the precise quantification of a specific group of proteins. Owing to its outstanding features, SRM arises as an alternative to antibody-based assays for the discovery and validation of clinically relevant biomarkers. In the present study (annex 3), 15 of the 24 biomarkers were qualified first by the SRM relative quantification based assay on an independent set of 50 samples (38% with PCa). 9 candidate biomarkers could not be analyzed using the SRM-based assay, because some good peptides that represented these proteins were missing, perhaps due to the small quantity of these proteins in the complex urine matrix, and also because some of these candidate biomarkers might not have been real candidates. After a logistic regression analysis of the data obtained through the relative quantification SRM assay, we obtained a panel of 7 peptides within 5 different proteins that were able to distinguish between PCa samples and benign control samples with a sensitivity of 95% and a specificity of 78% (Table 11). One of the proteins included in this panel of biomarkers was PSA. This protein showed an under-expression in PCa urine samples, in comparison with the benign counterparts. These results correlate with data from Boulduc *et al*. who found that PSA was down-regulated in urine from PCa patient [374]. They also suggested that the ratio of serum PSA to urine PSA could have diagnostic utility. Even though the number of samples used in our first step qualification/verification study was not very high, the results obtained were very promising. They outperformed our previous urine-based RNA testing panel. However, such an approach must be confirmed with a blind verification set.

While SRM shows considerable promise for protein quantification, the technique is still in its early stages for clinical application. The first important points are accuracy and reproducibility [214]. To obtain statistically significant data, expanding the number of analyzed components requires an increase in the number of samples analyzed and, consequently, greater computing power. Clinical proteome analysis can be seen as a sophisticated comparative analysis of a large number of variables in a limited number data set. Every sample manipulation increases the possibility of introducing artifacts, reduces reproducibility, and may further increase the complexity of samples [361]. Thus, reproducible procedures for the preparation of protein samples isolated from human urines are essential for meaningful proteomic validation analyses. Recent studies in this field have concluded that organic solvent precipitation, followed by in-solution digestion provide the best performance for urine proteome analysis in the discovery and verification phases [403, 404]. We followed the same

strategy as these groups, in order to compare not only the different protein extraction procedures, but also to optimize the trypsin-based protein digestion that would be used in the quantification SRM-based assays (annex 4). We concluded that the utilization of acetonitrile precipitation at room temperature, followed by in-solution 1:10 (trypsin: protein) and overnight trypsin digestion gave high quantification values in the SRM-assay. Moreover, an incubation step with NAC, prior to enzymatic digestion, avoided N-terminal IAA over-alkylation and, thus, helped to obtain better quantification results.

Once the protein preparation procedure was set up, we moved on to the second qualification/verification step. We decided to increase our list of protein candidate biomarkers using the secreted proteins described in the literature. 42 proteins were quantified using single point quantification in an SRM-based assay on 107 post-PM urine samples (annex 4). In order to better understand the behavior of those proteins, further analysis on the obtained data was required. Those analyses are still ongoing; thus, no final conclusions have been included in this study at present. We hope to obtain interesting results, and though we will probably not be able to define any unique protein or peptide to serve as a diagnostic biomarker (depending on how selectively and sensitively it enables PCa assessment), we will probably end up with a panel of protein biomarkers that are able to work together and that will hopefully outperform and/or complement the results of PSA alone. In addition, this list of 42 candidate biomarkers may include promising biomarker candidates for identifying those cancers that are clinically significant, from the latent to the indolent (prognostic biomarkers).

So, the next step for urinary proteomics will be to move onward: first, making conclusions in the second qualification/verification phase and second, validating the "verified" biomarkers on a much larger cohort of samples. These findings must then be translated to a much easier format for the creation of diagnostic kits or tests that will be readily accessible and applicable in clinical practice [225].

In summary, the future of PCa biomarkers continues to move forward. It remains for us to validate the many exciting prostate biomarkers that have already been discovered. We also need to continue to discover markers that will help to minimize the number of unnecessary prostate biopsies and continue to develop markers that will identify those men with indolent PCa, who will not be affected by disease in their lifetimes and who do not need treatment. Markers still need to be developed that can identify men with aggressive disease, who will benefit from local therapy and those who are likely to fail local therapy and require adjuvant intervention. And finally, markers still need to be developed that may serve as surrogate end points for clinical progression or survival [405]. Another important point that needs to be

addressed is the necessity of the PM. In the future, we would like to know if urine samples without previous PM can provide enough material to correctly detect those biomarkers and, thus, enable a correct diagnosis. Although PM is part of the diagnostic tripod (PSA, DRE or "PM" and Biopsy), it is usually poorly tolerated by patients and always requires medical intervention. This fact may represent a limiting factor, since the urologist would need to have the facilities to freeze and store urine samples before sending them to the laboratory. In large trials, the question of whether and how to perform the PM to optimize sensitivity and specificity has to be addressed for each potential marker [240].

In conclusion, the data presented in this dissertation represent a significant advance in the standard care for PCa diagnosis. Our two approaches (RNA- and Protein-based) have begun to yield promising results, as both have levels of specificity that exceed those of PSA (Table 11). However, validation studies on larger, multi-centric cohorts of urine samples are needed to end up with a valid PCa biomarker. We also expect to obtain some interesting biomarkers for PCa diagnosis, as well as some prognostic markers. The obtained results should have a rapid application in the clinics and potentially influence, together with actual screening parameters (serum PSA and DRE), decisions that could improve the health system, as well as clinical, managerial and/or public practices for health outcomes in PCa (Figure 19). We are currently building a national highway system for personalized medicine, based on DNA, RNA, and protein profiles [406]. We look forward to doctors and patients navigating these roads to better healthy results.

**Table 11. Summary of the obtained results**

| Study | Marker | Study (PCa %) | Sens. | Spec. | AUC | PPV | NPV | Biopsies saved |
|---|---|---|---|---|---|---|---|---|
| *Rigau et al. 2010* | PSGR | 215 (34%) | 95% | 15% | 0.68 | 0.36 | 0.84 | 11.6% |
| | PCA3 | 215 (34%) | 95% | 17% | 0.66 | 0.37 | 0.86 | 13% |
| | PSGRvPCA3 | 215 (34%) | 95% | 34% | 0.73 | 0.42 | 0.92 | 24.2% |
| *Rigau et al. 2011* | 3M* | 154 (37%) | 96% | 34% | 0.74 | 0.46 | 0.94 | 22.7% |
| | 3M* v PSA | 154 (37%) | 96% | 34% | 0.77 | 0.47 | 0.94 | 22.6% |
| | 3M* v PSAD | 154 (37%) | 96% | 40% | 0.80 | 0.49 | 0.95 | 26.4% |
| | 3M*(diagnostic gray zone) | 82 (34%) | 96% | 50% | 0.82 | 0.51 | 0.96 | 33.7% |
| | 3M* v PSAD (diagnostic gray zone) | 77 (35%) | 96% | 62% | 0.89 | 0.58 | 0.97 | 41.6% |
| *Unpublished Proteomic data* | PROTEOMIC PANNEL** | 50 (38%) | 95% | 78% | 0.93 | ND | ND | ND |

3M* (PSMAvPSGRvPCA3)
** Proteomic pannel of 7 peptides

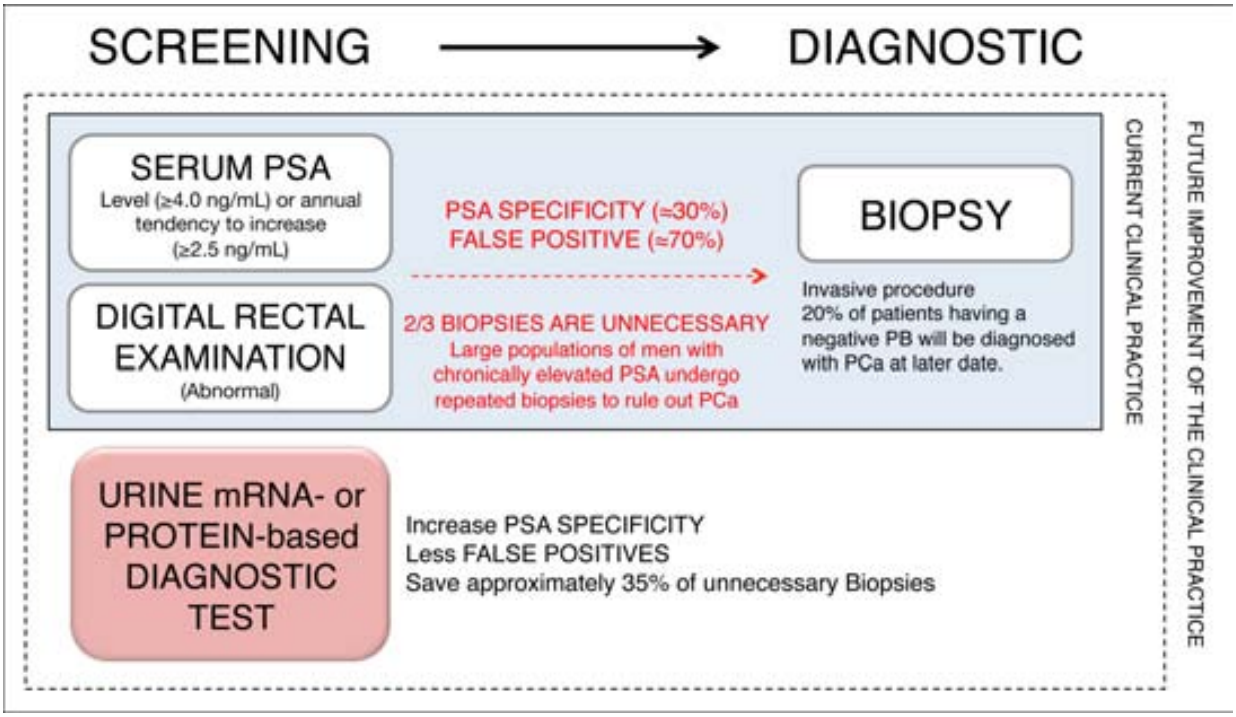



**Figure 19. Current and future improvement in PCa diagnostic squeme.**

# CONCLUSIONS

1. We have demonstrated that an RTqPCR assay on urine sediments from patients presenting for PB, based on PSGR (a new urine PCa biomarker), had a similar performance to the gold standard, urine-based RNA biomarker for PCa (PCA3).
2. We have also demonstrated that an RTqPCR assay on urine sediments from patients presenting for PB, based on the combination of PSGR and PCA3, improved the specificity of both individual biomarkers.
3. We have demonstrated that an RTqPCR assay on urine sediments from patients presenting for PB, based on PSMA (a marker previously described to be over-expressed in PCa), had a similar performance to PCA3 and PSGR.
4. We have demonstrated that an RTqPCR assay on urine sediments from patients presenting for PB with serum PSA levels between 4-10 ng/mL and no prior biopsy information, based on PSMA, could significantly improve the performance of PCA3 and PSGR.
5. We have shown that a multiplexed (PSMA, PSGR and PCA3) RTqPCR assay on urine sediments from patients who presented for PB could significantly improve the predictive ability, when compared to PCA3 or serum PSA alone.
6. We have shown that a multiplexed (PSMA, PSGR and PCA3) RTqPCR assay on urine sediments from patients who presented for PB with serum PSA levels between 4-10 ng/mL and no prior PB information could significantly improve the predictive ability, when compared to 3M, PCA3 or serum PSA alone.
7. Future studies need to be conducted with the multiplexed panel of biomarkers on patients with repeat biopsies and patients who have shown recurrence after prostatectomy, in order to establish specific thresholds and/or to generate predictive nomograms.

8. We found that the PCA3 score in men having a negative biopsy was statistically similar to that observed in men with isolated HGPIN. However, it was significantly lower than the score observed in men with PCa.
9. We demonstrated that the efficacy of post-PM urine PCA3 for identifying men with PCa was lower if we compared these levels with those of men with HGPIN to those without HGPIN.
10. We found that the PCA3 gene is not only a specific marker for PCa, since it is also over-expressed in HGPIN. However, future studies need to be conducted in HGPIN sub-populations. Therefore, we proposed that this fact should be taken into account, in order to establish specific cut-off points for the PCA3 score and to generate new biopsy nomograms for predicting PCa.

11. We demonstrated that 2D-DIGE analyses, using total urine protein and urine protein normalized for abundant proteins, are able to reveal novel biomarkers/diagnostic profiles for PCa in urine.
12. We qualified a panel of 7 peptides, within 5 different proteins (including urine PSA), by using an SRM-based relative quantification assay. Our panel of protein-based biomarkers was able to differentiate between PCa and benign urine samples with a sensitivity of 95% and a specificity of 78%.
13. We started a second qualification and verification phase in which 42 proteins were absolutely quantified using an SRM-based approach. However, these analyses are ongoing, and no final conclusions have been included in this study.
14. Future studies will need to be conducted with the new protein panel of biomarkers on patients with repeat biopsies and patients who have shown recurrence after prostatectomy, in order to establish specific thresholds and/or to generate predictive nomograms.
15. The future follow up of the analysed patients will reveal a possible prognostic potential of these panels of biomarkers.

16. The next step will be to move on towards the validation of these RNA- and protein-based biomarkers using a much larger cohort of samples. These findings will then be able to be translated to a much easier format for incorporation into diagnostic kits or tests that will be more accessible and applicable in clinical practice.
17. All together, transcriptomic and proteomic analysis constitutes an important step towards advancing the accurate diagnosis of PCa, which currently represents a setback in our ability to cure patients of the disease. Thus, RNA-based and protein-based biomarkers should have a rapid application in the clinics and, together with serum PSA and DRE, potentially influence decisions that could improve the health system, while reducing the number of unnecessary biopsies.

# BIBLIOGRAPHY

1. McNeal JE. Regional morphology and pathology of the prostate. Am J Clin Pathol 1968;49:347-57.
2. De Marzo AM, Platz EA, Sutcliffe S, et al. Inflammation in prostate carcinogenesis. Nat Rev Cancer 2007;7:256-69.
3. Myers RP. Structure of the adult prostate from a clinician's standpoint. Clin Anat 2000;13:214-5.
4. Abate-Shen C, Shen MM. Molecular genetics of prostate cancer. Genes Dev 2000;14:2410-34.
5. Tran CP, Lin C, Yamashiro J, Reiter RE. Prostate stem cell antigen is a marker of late intermediate prostate epithelial cells. Mol Cancer Res 2002;1:113-21.
6. Aboushwareb T, Atala A. Stem cells in urology. Nat Clin Pract Urol 2008;5:621-31.
7. Epstein JaN, GJ. Biopsy interpretation of the prostate. Lippincott Williams & Willkins, a Wolters Kluwer business ed. Philadelphia, PA, USA; 2008.
8. Weitzman SA, Gordon LI. Inflammation and cancer: role of phagocyte-generated oxidants in carcinogenesis. Blood 1990;76:655-63.
9. McNeal J. Pathology of benign prostatic hyperplasia. Insight into etiology. Urol Clin North Am 1990;17:477-86.
10. Thomas AW, Abrams P. Lower urinary tract symptoms, benign prostatic obstruction and the overactive bladder. BJU Int 2000;85 Suppl 3:57-68; discussion 70-1.
11. De Marzo AM, Coffey DS, Nelson WG. New concepts in tissue specificity for prostate cancer and benign prostatic hyperplasia. Urology 1999;53:29-39; discussion -42.
12. Gallardo-Arrieta F, Doll A, Rigau M, et al. A transcriptional signature associated with the onset of benign prostate hyperplasia in a canine model. Prostate 2010;70:1402-12.
13. Lee KL, Peehl DM. Molecular and cellular pathogenesis of benign prostatic hyperplasia. J Urol 2004;172:1784-91.
14. Gallardo-Arrieta F, Doll A, Rigau M, et al. A transcriptional signature associated with the onset of benign prostate hyperplasia in a canine model. Prostate;70:1402-12.
15. McNeal JE. Normal histology of the prostate. Am J Surg Pathol 1988;12:619-33.
16. De Marzo AM, Marchi VL, Epstein JI, Nelson WG. Proliferative inflammatory atrophy of the prostate: implications for prostatic carcinogenesis. Am J Pathol 1999;155:1985-92.
17. Tomlins SA, Mehra R, Rhodes DR, et al. Integrative molecular concept modeling of prostate cancer progression. Nat Genet 2007;39:41-51.
18. Montironi R, Scattoni V, Mazzucchelli R, Lopez-Beltran A, Bostwick DG, Montorsi F. Atypical foci suspicious but not diagnostic of malignancy in prostate needle biopsies (also referred to as "atypical small acinar proliferation suspicious for but not diagnostic of malignancy"). Eur Urol 2006;50:666-74.
19. Bostwick DG, Meiers I. Atypical small acinar proliferation in the prostate: clinical significance in 2006. Arch Pathol Lab Med 2006;130:952-7.
20. Ryu JH, Kim YB, Lee JK, Kim YJ, Jung TY. Predictive factors of prostate cancer at repeat biopsy in patients with an initial diagnosis of atypical small acinar proliferation of the prostate. Korean J Urol;51:752-6.
21. Haggman MJ, Macoska JA, Wojno KJ, Oesterling JE. The relationship between prostatic intraepithelial neoplasia and prostate cancer: critical issues. J Urol 1997;158:12-22.
22. Montironi R, Mazzucchelli R, Lopez-Beltran A, Cheng L, Scarpelli M. Mechanisms of disease: high-grade prostatic intraepithelial neoplasia and other proposed preneoplastic lesions in the prostate. Nat Clin Pract Urol 2007;4:321-32.
23. Bostwick DG, Shan A, Qian J, et al. Independent origin of multiple foci of prostatic intraepithelial neoplasia: comparison with matched foci of prostate carcinoma. Cancer 1998;83:1995-2002.
24. Joshua AM, Evans A, Van der Kwast T, et al. Prostatic preneoplasia and beyond. Biochim Biophys Acta 2008;1785:156-81.

25. Sakr WA. Prostatic intraepithelial neoplasia: A marker for high-risk groups and a potential target for chemoprevention. Eur Urol 1999;35:474-8.
26. Sakr WA, Haas GP, Cassin BF, Pontes JE, Crissman JD. The frequency of carcinoma and intraepithelial neoplasia of the prostate in young male patients. J Urol 1993;150:379-85.
27. Drago Jea. Introductory remarks and workshop summary. Urology 1992;39 (Suppl): S2-S8.
28. Bostwick DG, Liu L, Brawer MK, Qian J. High-grade prostatic intraepithelial neoplasia. Rev Urol 2004;6:171-9.
29. Montironi R, Mazzucchelli R, Algaba F, Lopez-Beltran A. Morphological identification of the patterns of prostatic intraepithelial neoplasia and their importance. J Clin Pathol 2000;53:655-65.
30. Sakr WA, Grignon DJ, Haas GP, Heilbrun LK, Pontes JE, Crissman JD. Age and racial distribution of prostatic intraepithelial neoplasia. Eur Urol 1996;30:138-44.
31. Qian J, Wollan P, Bostwick DG. The extent and multicentricity of high-grade prostatic intraepithelial neoplasia in clinically localized prostatic adenocarcinoma. Hum Pathol 1997;28:143-8.
32. Helpap B. The significance of the P504S expression pattern of high-grade prostatic intraepithelial neoplasia (HGPIN) with and without adenocarcinoma of the prostate in biopsy and radical prostatectomy specimens. Virchows Arch 2006;448:480-4.
33. Montironi R, Mazzucchelli R, Stramazzotti D, Pomante R, Thompson D, Bartels PH. Expression of pi-class glutathione S-transferase: two populations of high grade prostatic intraepithelial neoplasia with different relations to carcinoma. Mol Pathol 2000;53:122-8.
34. Morote J, Fernandez S, Alana L, et al. PTOV1 expression predicts prostate cancer in men with isolated high-grade prostatic intraepithelial neoplasia in needle biopsy. Clin Cancer Res 2008;14:2617-22.
35. Epstein JI, Potter SR. The pathological interpretation and significance of prostate needle biopsy findings: implications and current controversies. J Urol 2001;166:402-10.
36. Morote J, Raventos CX, Encabo G, Lopez M, de Torres IM. Effect of high-grade prostatic intraepithelial neoplasia on total and percent free serum prostatic-specific antigen. Eur Urol 2000;37:456-9.
37. Epstein JI, Herawi M. Prostate needle biopsies containing prostatic intraepithelial neoplasia or atypical foci suspicious for carcinoma: implications for patient care. J Urol 2006;175:820-34.
38. Zlotta AR, Raviv G, Schulman CC. Clinical prognostic criteria for later diagnosis of prostate carcinoma in patients with initial isolated prostatic intraepithelial neoplasia. Eur Urol 1996;30:249-55.
39. Herawi M, Kahane H, Cavallo C, Epstein JI. Risk of prostate cancer on first re-biopsy within 1 year following a diagnosis of high grade prostatic intraepithelial neoplasia is related to the number of cores sampled. J Urol 2006;175:121-4.
40. Moore CK, Karikehalli S, Nazeer T, Fisher HA, Kaufman RP, Jr., Mian BM. Prognostic significance of high grade prostatic intraepithelial neoplasia and atypical small acinar proliferation in the contemporary era. J Urol 2005;173:70-2.
41. Loeb S, Roehl KA, Yu X, Han M, Catalona WJ. Use of prostate-specific antigen velocity to follow up patients with isolated high-grade prostatic intraepithelial neoplasia on prostate biopsy. Urology 2007;69:108-12.
42. Bostwick DG, Brawer MK. Prostatic intra-epithelial neoplasia and early invasion in prostate cancer. Cancer 1987;59:788-94.
43. Putzi MJ, De Marzo AM. Morphologic transitions between proliferative inflammatory atrophy and high-grade prostatic intraepithelial neoplasia. Urology 2000;56:828-32.
44. DeMarzo AM, Nelson WG, Isaacs WB, Epstein JI. Pathological and molecular aspects of prostate cancer. Lancet 2003;361:955-64.

45. Brown ML, Lipscomb J, Snyder C. The burden of illness of cancer: economic cost and quality of life. Annu Rev Public Health 2001;22:91-113.
46. Jemal A, Siegel R, Ward E, Hao Y, Xu J, Thun MJ. Cancer statistics, 2009. CA Cancer J Clin 2009.
47. Ferlay J, Parkin DM, Steliarova-Foucher E. Estimates of cancer incidence and mortality in Europe in 2008. Eur J Cancer 2010;46:765-81.
48. Jemal A, Siegel R, Xu J, Ward E. Cancer statistics, 2010. CA Cancer J Clin;60:277-300.
49. American Cancer Society. Cancer Facts & Figures 2010. Atlanta: American Cancer Society. 2010.
50. Curado M, Edwards, B., Shin, HR., Store, H., Ferlay, J., Heanue, M. Cancer Incidence in Five Continents; http://www-dep.iarc.fr. Lió: IARC Scientific Publications 2007;Vol. IX. No. 160.
51. Legler JM, Feuer EJ, Potosky AL, Merrill RM, Kramer BS. The role of prostate-specific antigen (PSA) testing patterns in the recent prostate cancer incidence decline in the United States. Cancer Causes Control 1998;9:519-27.
52. Parkin DM, Whelan, S.L., Ferlay, J., Store, H. Cancer Incidence in Five Continents; http://www-dep.iarc.fr. Lió: IARC Cancer Base 2005;Vol. I a VIII, No. 7.
53. WHO Mortality Database; http://www.who.int/whosis/en/. In; 2008.
54. Parkin DM, Bray FI, Devesa SS. Cancer burden in the year 2000. The global picture. Eur J Cancer 2001;37 Suppl 8:S4-66.
55. Kolonel LN, Hankin JH, Whittemore AS, et al. Vegetables, fruits, legumes and prostate cancer: a multiethnic case-control study. Cancer Epidemiol Biomarkers Prev 2000;9:795-804.
56. Haenszel W, Kurihara M. Studies of Japanese migrants. I. Mortality from cancer and other diseases among Japanese in the United States. J Natl Cancer Inst 1968;40:43-68.
57. Platz EaG, E. . Prostate cancer. In: D S, JF F, eds. Cancer Epidemiology and Prevention 3ed. Oxford. Oxford University Press 2006;1128-1150.
58. Carter BS, Beaty TH, Steinberg GD, Childs B, Walsh PC. Mendelian inheritance of familial prostate cancer. Proceedings of the National Academy of Sciences of the United States of America 1992;89:3367-71.
59. Bratt O. Hereditary prostate cancer: clinical aspects. J Urol 2002;168:906-13.
60. Cancer risks in BRCA2 mutation carriers. The Breast Cancer Linkage Consortium. J Natl Cancer Inst 1999;91:1310-6.
61. Agalliu I, Gern R, Leanza S, Burk RD. Associations of high-grade prostate cancer with BRCA1 and BRCA2 founder mutations. Clin Cancer Res 2009;15:1112-20.
62. Foulkes WD. Inherited susceptibility to common cancers. N Engl J Med 2008;359:2143-
53.
63. Amundadottir LT, Sulem P, Gudmundsson J, et al. A common variant associated with prostate cancer in European and African populations. Nat Genet 2006;38:652-8.
64. Gudmundsson J, Sulem P, Manolescu A, et al. Genome-wide association study identifies a second prostate cancer susceptibility variant at 8q24. Nat Genet 2007;39:631-7.
65. Kolonel LN, Altshuler D, Henderson BE. The multiethnic cohort study: exploring genes, lifestyle and cancer risk. Nat Rev Cancer 2004;4:519-27.
66. DevCan: Probability of Developing or Dying of Cancer Software, Version 6.4.1. . Statistical Research and Applications Branch, National Cancer Institute srabcancergov/devcan 2009.
67. Edwards BK, Ward E, Kohler BA, et al. Annual report to the nation on the status of cancer, 1975-2006, featuring colorectal cancer trends and impact of interventions (risk factors, screening, and treatment) to reduce future rates. Cancer;116:544-73.
68. Kushi LH, Byers T, Doyle C, et al. American Cancer Society Guidelines on Nutrition and Physical Activity for cancer prevention: reducing the risk of cancer with healthy food choices and physical activity. CA Cancer J Clin 2006;56:254-81; quiz 313-4.

69. Gann PH, Ma J, Giovannucci E, et al. Lower prostate cancer risk in men with elevated plasma lycopene levels: results of a prospective analysis. Cancer Res 1999;59:1225-30.
70. Hoque A, Albanes D, Lippman SM, et al. Molecular epidemiologic studies within the Selenium and Vitamin E Cancer Prevention Trial (SELECT). Cancer Causes Control 2001;12:627-33.
71. Hanahan D, Weinberg RA. The hallmarks of cancer. Cell 2000;100:57-70.
72. Hanahan D, Weinberg RA. Hallmarks of cancer: the next generation. Cell;144:646-74.
73. Shen MM, Abate-Shen C. Molecular genetics of prostate cancer: new prospects for old challenges. Genes Dev;24:1967-2000.
74. Shah RB, Mehra R, Chinnaiyan AM, et al. Androgen-independent prostate cancer is a heterogeneous group of diseases: lessons from a rapid autopsy program. Cancer Res 2004;64:9209-16.
75. Liu W, Laitinen S, Khan S, et al. Copy number analysis indicates monoclonal origin of lethal metastatic prostate cancer. Nat Med 2009;15:559-65.
76. Humphrey PA. Diagnosis of adenocarcinoma in prostate needle biopsy tissue. J Clin Pathol 2007;60:35-42.
77. Grisanzio C, Signoretti S. p63 in prostate biology and pathology. J Cell Biochem 2008;103:1354-68.
78. Luo J, Zha S, Gage WR, et al. Alpha-methylacyl-CoA racemase: a new molecular marker for prostate cancer. Cancer Res 2002;62:2220-6.
79. Taylor BS, Schultz N, Hieronymus H, et al. Integrative genomic profiling of human prostate cancer. Cancer Cell;18:11-22.
80. Grignon DJ. Unusual subtypes of prostate cancer. Mod Pathol 2004;17:316-27.
81. Roy-Burman P, Wu H, Powell WC, Hagenkord J, Cohen MB. Genetically defined mouse models that mimic natural aspects of human prostate cancer development. Endocr Relat Cancer 2004;11:225-54.
82. Tao Z, Shen M, Zheng Y, et al. PCA3 gene expression in prostate cancer tissue in a Chinese population: quantification by real-time FQ-RT-PCR based on exon 3 of PCA3. Exp Mol Pathol;89:58-62.
83. Rossi DJ, Bryder D, Seita J, Nussenzweig A, Hoeijmakers J, Weissman IL. Deficiencies in DNA damage repair limit the function of haematopoietic stem cells with age. Nature 2007;447:725-9.
84. Isaacs JT. Control of cell proliferation and cell death in the normal and neoplastic prostate: A stem cell model. In Benign prostatic hyperplasia (ed. Rodgers CH, et al.). Department of Helath and Human Services, Washington, DC 1985:pp. 85–94.
85. English HF, Santen RJ, Isaacs JT. Response of glandular versus basal rat ventral prostatic epithelial cells to androgen withdrawal and replacement. Prostate 1987;11:229-42.
86. Evans GS, Chandler JA. Cell proliferation studies in the rat prostate: II. The effects of castration and androgen-induced regeneration upon basal and secretory cell proliferation. Prostate 1987;11:339-51.
87. Begley L, Monteleon C, Shah RB, Macdonald JW, Macoska JA. CXCL12 overexpression and secretion by aging fibroblasts enhance human prostate epithelial proliferation in vitro. Aging Cell 2005;4:291-8.
88. Bavik C, Coleman I, Dean JP, Knudsen B, Plymate S, Nelson PS. The gene expression program of prostate fibroblast senescence modulates neoplastic epithelial cell proliferation through paracrine mechanisms. Cancer Res 2006;66:794-802.
89. Bethel CR, Chaudhary J, Anway MD, Brown TR. Gene expression changes are age-dependent and lobe-specific in the brown Norway rat model of prostatic hyperplasia. Prostate 2009;69:838-50.
90. Sfanos KS, Sauvageot J, Fedor HL, Dick JD, De Marzo AM, Isaacs WB. A molecular analysis of prokaryotic and viral DNA sequences in prostate tissue from patients with

prostate cancer indicates the presence of multiple and diverse microorganisms. Prostate 2008;68:306-20.
91. De Marzo AM, Meeker AK, Zha S, et al. Human prostate cancer precursors and pathobiology. Urology 2003;62:55-62.
92. Meeker AK, Hicks JL, Platz EA, et al. Telomere shortening is an early somatic DNA alteration in human prostate tumorigenesis. Cancer Res 2002;62:6405-9.
93. Courtois-Cox S, Jones SL, Cichowski K. Many roads lead to oncogene-induced senescence. Oncogene 2008;27:2801-9.
94. Berger MF, Lawrence MS, Demichelis F, et al. The genomic complexity of primary human prostate cancer. Nature;470:214-20.
95. Kondo Y, Shen L, Cheng AS, et al. Gene silencing in cancer by histone H3 lysine 27 trimethylation independent of promoter DNA methylation. Nat Genet 2008;40:741-50.
96. Nelson WG, Yegnasubramanian S, Agoston AT, et al. Abnormal DNA methylation, epigenetics, and prostate cancer. Front Biosci 2007;12:4254-66.
97. Nelson WG, De Marzo AM, Yegnasubramanian S. Epigenetic alterations in human prostate cancers. Endocrinology 2009;150:3991-4002.
98. Nakayama M, Gonzalgo ML, Yegnasubramanian S, Lin X, De Marzo AM, Nelson WG. GSTP1 CpG island hypermethylation as a molecular biomarker for prostate cancer. J Cell Biochem 2004;91:540-52.
99. Varambally S, Dhanasekaran SM, Zhou M, et al. The polycomb group protein EZH2 is involved in progression of prostate cancer. Nature 2002;419:624-9.
100. Markowski MC, Bowen C, Gelmann EP. Inflammatory cytokines induce phosphorylation and ubiquitination of prostate suppressor protein NKX3.1. Cancer Res 2008;68:6896-901.
101. Abate-Shen C, Shen MM, Gelmann E. Integrating differentiation and cancer: the Nkx3.1 homeobox gene in prostate organogenesis and carcinogenesis. Differentiation 2008;76:717-27.
102. Chen HJ, Xu M, Zhang L, Zhang YK, Wang GM. [Prostate sarcoma: a report of 14 cases]. Zhonghua Nan Ke Xue 2005;11:683-5.
103. Casey G, Neville PJ, Plummer SJ, et al. RNASEL Arg462Gln variant is implicated in up to 13% of prostate cancer cases. Nat Genet 2002;32:581-3.
104. Kelley MR, Cheng L, Foster R, et al. Elevated and altered expression of the multifunctional DNA base excision repair and redox enzyme Ape1/ref-1 in prostate cancer. Clin Cancer Res 2001;7:824-30.
105. Chen Z, Trotman LC, Shaffer D, et al. Crucial role of p53-dependent cellular senescence in suppression of Pten-deficient tumorigenesis. Nature 2005;436:725-30.
106. Sommerfeld HJ, Meeker AK, Piatyszek MA, Bova GS, Shay JW, Coffey DS. Telomerase activity: a prevalent marker of malignant human prostate tissue. Cancer Res 1996;56:218-22.
107. Zhang W, Kapusta LR, Slingerland JM, Klotz LH. Telomerase activity in prostate cancer, prostatic intraepithelial neoplasia, and benign prostatic epithelium. Cancer Res 1998;58:619-21.
108. Gurel B, Iwata T, Koh CM, et al. Nuclear MYC protein overexpression is an early alteration in human prostate carcinogenesis. Mod Pathol 2008;21:1156-67.
109. Foster BA, Kaplan PJ, Greenberg NM. Peptide growth factors and prostate cancer: new models, new opportunities. Cancer Metastasis Rev 1998;17:317-24.
110. Djakiew D. Dysregulated expression of growth factors and their receptors in the development of prostate cancer. Prostate 2000;42:150-60.
111. Umbas R, Schalken JA, Aalders TW, et al. Expression of the cellular adhesion molecule E-cadherin is reduced or absent in high-grade prostate cancer. Cancer Res 1992;52:5104-9.

112. Morton RA, Ewing CM, Nagafuchi A, Tsukita S, Isaacs WB. Reduction of E-cadherin levels and deletion of the alpha-catenin gene in human prostate cancer cells. Cancer Res 1993;53:3585-90.
113. Kleinerman DI, Troncoso P, Lin SH, et al. Consistent expression of an epithelial cell adhesion molecule (C-CAM) during human prostate development and loss of expression in prostate cancer: implication as a tumor suppressor. Cancer Res 1995;55:1215-20.
114. Pisters LL, Troncoso P, Zhau HE, Li W, von Eschenbach AC, Chung LW. c-met proto-oncogene expression in benign and malignant human prostate tissues. J Urol 1995;154:293-8.
115. Cress AE, Rabinovitz I, Zhu W, Nagle RB. The alpha 6 beta 1 and alpha 6 beta 4 integrins in human prostate cancer progression. Cancer Metastasis Rev 1995;14:219-28.
116. Wang L, Liu R, Li W, et al. Somatic single hits inactivate the X-linked tumor suppressor FOXP3 in the prostate. Cancer Cell 2009;16:336-46.
117. Mellinghoff IK, Vivanco I, Kwon A, Tran C, Wongvipat J, Sawyers CL. HER2/neu kinase-dependent modulation of androgen receptor function through effects on DNA binding and stability. Cancer Cell 2004;6:517-27.
118. Saramaki OR, Tammela TL, Martikainen PM, Vessella RL, Visakorpi T. The gene for polycomb group protein enhancer of zeste homolog 2 (EZH2) is amplified in late-stage prostate cancer. Genes Chromosomes Cancer 2006;45:639-45.
119. Bachmann IM, Halvorsen OJ, Collett K, et al. EZH2 expression is associated with high proliferation rate and aggressive tumor subgroups in cutaneous melanoma and cancers of the endometrium, prostate, and breast. J Clin Oncol 2006;24:268-73.
120. Gaspar-Maia A, Alajem A, Polesso F, et al. Chd1 regulates open chromatin and pluripotency of embryonic stem cells. Nature 2009;460:863-8.
121. Tilley WD, Buchanan G, Hickey TE, Bentel JM. Mutations in the androgen receptor gene are associated with progression of human prostate cancer to androgen independence. Clin Cancer Res 1996;2:277-85.
122. Niu Y, Altuwaijri S, Lai KP, et al. Androgen receptor is a tumor suppressor and proliferator in prostate cancer. Proceedings of the National Academy of Sciences of the United States of America 2008;105:12182-7.
123. Bookstein R, MacGrogan D, Hilsenbeck SG, Sharkey F, Allred DC. p53 is mutated in a subset of advanced-stage prostate cancers. Cancer Res 1993;53:3369-73.
124. Chan JM, Stampfer MJ, Giovannucci E, et al. Plasma insulin-like growth factor-I and prostate cancer risk: a prospective study. Science 1998;279:563-6.
125. Dong JT, Lamb PW, Rinker-Schaeffer CW, et al. KAI1, a metastasis suppressor gene for prostate cancer on human chromosome 11p11.2. Science 1995;268:884-6.
126. Colombel M, Symmans F, Gil S, et al. Detection of the apoptosis-suppressing oncoprotein bc1-2 in hormone-refractory human prostate cancers. Am J Pathol 1993;143:390-400.
127. Huggins C, Hodges CV. Studies on prostatic cancer. I. The effect of castration, of estrogen and androgen injection on serum phosphatases in metastatic carcinoma of the prostate. CA Cancer J Clin 1972;22:232-40.
128. Feldman BJ, Feldman D. The development of androgen-independent prostate cancer. Nat Rev Cancer 2001;1:34-45.
129. Gregory CW, Hamil KG, Kim D, et al. Androgen receptor expression in androgen-independent prostate cancer is associated with increased expression of androgen-regulated genes. Cancer Res 1998;58:5718-24.
130. Isaacs JT, Coffey DS. Adaptation versus selection as the mechanism responsible for the relapse of prostatic cancer to androgen ablation therapy as studied in the Dunning R-3327-H adenocarcinoma. Cancer Res 1981;41:5070-5.
131. Liao S, Hung SC, Tymoczko JL, Liang T. Active forms and biodynamics of the androgen-receptor in various target tissues. Curr Top Mol Endocrinol 1976;4:139-51.

132. Bruchovsky N, Rennie PS, Lesser B, Sutherland DJ. The influence of androgen receptors on the concentration of androgens in nuclei of hormone-responsive cells. J Steroid Biochem 1975;6:551-60.
133. Krieg M, Bartsch W, Janssen W, Voigt KD. A comparative study of binding, metabolism and endogenous levels of androgens in normal, hyperplastic and carcinomatous human prostate. J Steroid Biochem 1979;11:615-24.
134. Logothetis CJ, Lin SH. Osteoblasts in prostate cancer metastasis to bone. Nat Rev Cancer 2005;5:21-8.
135. Liotta LA, Rao CN. Tumor invasion and metastasis. Monogr Pathol 1986:183-92.
136. Clarke NW, Hart CA, Brown MD. Molecular mechanisms of metastasis in prostate cancer. Asian J Androl 2009;11:57-67.
137. Holcomb IN, Grove DI, Kinnunen M, et al. Genomic alterations indicate tumor origin and varied metastatic potential of disseminated cells from prostate cancer patients. Cancer Res 2008;68:5599-608.
138. Scher HI, Jia X, de Bono JS, et al. Circulating tumour cells as prognostic markers in progressive, castration-resistant prostate cancer: a reanalysis of IMMC38 trial data. Lancet Oncol 2009;10:233-9.
139. Eschwege P, Moutereau S, Droupy S, et al. Prognostic value of prostate circulating cells detection in prostate cancer patients: a prospective study. Br J Cancer 2009;100:608-10.
140. Cunha GR, Donjacour AA, Cooke PS, et al. The endocrinology and developmental biology of the prostate. Endocr Rev 1987;8:338-62.
141. Kong D, Banerjee S, Ahmad A, et al. Epithelial to mesenchymal transition is mechanistically linked with stem cell signatures in prostate cancer cells. PLoS One;5:e12445.
142. Gleason DF. Classification of prostatic carcinomas. Cancer Chemother Rep 1966;50:125-8.
143. Gleason DF, Mellinger GT. Prediction of prognosis for prostatic adenocarcinoma by combined histological grading and clinical staging. J Urol 1974;111:58-64.
144. Epstein JI, Allsbrook WC, Jr., Amin MB, Egevad LL. The 2005 International Society of Urological Pathology (ISUP) Consensus Conference on Gleason Grading of Prostatic Carcinoma. Am J Surg Pathol 2005;29:1228-42.
145. Amin M, Boccon-Gibod L, Egevad L, et al. Prognostic and predictive factors and reporting of prostate carcinoma in prostate needle biopsy specimens. Scand J Urol Nephrol Suppl 2005:20-33.
146. Lotan TL, Epstein JI. Clinical implications of changing definitions within the Gleason grading system. Nat Rev Urol;7:136-42.
147. Ohori M, Wheeler TM, Scardino PT. The New American Joint Committee on Cancer and International Union Against Cancer TNM classification of prostate cancer. Clinicopathologic correlations. Cancer 1994;74:104-14.
148. Ludwig JA, Weinstein JN. Biomarkers in cancer staging, prognosis and treatment selection. Nat Rev Cancer 2005;5:845-56.
149. Sobin LH, Gospodariwicz, M. and Wittekind, C. TNM Classification of malignant tumours. UICC International Union Against Cancer. Wiley-Blackwell 2009;2009 Dec; pp. 243-248.
150. Guidelines EAoU. European Association of Urology; 2010.
151. Crawford ED, Andriole GL, Marberger M, Rittmaster RS. Reduction in the risk of prostate cancer: future directions after the Prostate Cancer Prevention Trial. Urology;75:502-9.
152. Choo R, Klotz L, Danjoux C, et al. Feasibility study: watchful waiting for localized low to intermediate grade prostate carcinoma with selective delayed intervention based on prostate specific antigen, histological and/or clinical progression. J Urol 2002;167:1664-9.

153. Strope SA, Andriole GL. Prostate cancer screening: current status and future perspectives. Nat Rev Urol;7:487-93.
154. Rigau M, Morote J, Mir MC, et al. PSGR and PCA3 as biomarkers for the detection of prostate cancer in urine. Prostate;70:1760-7.
155. Loeb S, Gashti SN, Catalona WJ. Exclusion of inflammation in the differential diagnosis of an elevated prostate-specific antigen (PSA). Urol Oncol 2009;27:64-6.
156. Pannek J, Partin AW. Prostate-specific antigen: what's new in 1997. Oncology (Williston Park) 1997;11:1273-8; discussion 9-82.
157. Catalona WJ, Smith DS, Ratliff TL, et al. Measurement of prostate-specific antigen in serum as a screening test for prostate cancer. N Engl J Med 1991;324:1156-61.
158. Makinen T, Tammela TL, Stenman UH, et al. Second round results of the Finnish population-based prostate cancer screening trial. Clin Cancer Res 2004;10:2231-6.
159. Cervera Deval J, Morales Olaya FJ, Jornet Fayos J, Gonzalez Anon M. [Diagnostic value of the second prostate biopsies in males of risk. Study stratified by value of PSA]. Actas Urol Esp 2004;28:666-71.
160. Papsidero LD, Wang MC, Valenzuela LA, Murphy GP, Chu TM. A prostate antigen in sera of prostatic cancer patients. Cancer Res 1980;40:2428-32.
161. Lilja H, Ulmert D, Vickers AJ. Prostate-specific antigen and prostate cancer: prediction, detection and monitoring. Nat Rev Cancer 2008;8:268-78.
162. Ung JO, Richie JP, Chen MH, Renshaw AA, D'Amico AV. Evolution of the presentation and pathologic and biochemical outcomes after radical prostatectomy for patients with clinically localized prostate cancer diagnosed during the PSA era. Urology 2002;60:458-63.
163. Thompson IM, Ankerst DP, Chi C, et al. Assessing prostate cancer risk: results from the Prostate Cancer Prevention Trial. J Natl Cancer Inst 2006;98:529-34.
164. Bostwick DG, Burke HB, Djakiew D, et al. Human prostate cancer risk factors. Cancer 2004;101:2371-490.
165. Thompson IM, Pauler DK, Goodman PJ, et al. Prevalence of prostate cancer among men with a prostate-specific antigen level < or =4.0 ng per milliliter. N Engl J Med 2004;350:2239-46.
166. Herrala AM, Porvari KS, Kyllonen AP, Vihko PT. Comparison of human prostate specific glandular kallikrein 2 and prostate specific antigen gene expression in prostate with gene amplification and overexpression of prostate specific glandular kallikrein 2 in tumor tissue. Cancer 2001;92:2975-84.
167. Qiu SD, Young CY, Bilhartz DL, et al. In situ hybridization of prostate-specific antigen mRNA in human prostate. J Urol 1990;144:1550-6.
168. Kulasingam V, Diamandis EP. Strategies for discovering novel cancer biomarkers through utilization of emerging technologies. Nat Clin Pract Oncol 2008;5:588-99.
169. Morote J, Raventos CX, Lorente JA, et al. Comparison of percent free prostate specific antigen and prostate specific antigen density as methods to enhance prostate specific antigen specificity in early prostate cancer detection in men with normal rectal examination and prostate specific antigen between 4.1 and 10 ng./ml. J Urol 1997;158:502-4.
170. Stamey TA, Yang N, Hay AR, McNeal JE, Freiha FS, Redwine E. Prostate-specific antigen as a serum marker for adenocarcinoma of the prostate. N Engl J Med 1987;317:909-16.
171. Christensson A, Bjork T, Nilsson O, et al. Serum prostate specific antigen complexed to alpha 1-antichymotrypsin as an indicator of prostate cancer. J Urol 1993;150:100-5.
172. Carter HB, Pearson JD, Metter EJ, et al. Longitudinal evaluation of prostate-specific antigen levels in men with and without prostate disease. JAMA 1992;267:2215-20.
173. D'Amico AV, Chen MH, Roehl KA, Catalona WJ. Preoperative PSA velocity and the risk of death from prostate cancer after radical prostatectomy. N Engl J Med 2004;351:125-35.

174. D'Amico AV, Renshaw AA, Sussman B, Chen MH. Pretreatment PSA velocity and risk of death from prostate cancer following external beam radiation therapy. JAMA 2005;294:440-7.
175. Schmid HP, McNeal JE, Stamey TA. Observations on the doubling time of prostate cancer. The use of serial prostate-specific antigen in patients with untreated disease as a measure of increasing cancer volume. Cancer 1993;71:2031-40.
176. D'Amico AV, Kantoff P, Loffredo M, Renshaw AA, Loffredo B, Chen MH. Predictors of mortality after prostate-specific antigen failure. Int J Radiat Oncol Biol Phys 2006;65:656-60.
177. Heidenreich A. Identification of high-risk prostate cancer: role of prostate-specific antigen, PSA doubling time, and PSA velocity. Eur Urol 2008;54:976-7; discussion 8-9.
178. Ramirez ML, Nelson EC, Devere White RW, Lara PN, Jr., Evans CP. Current applications for prostate-specific antigen doubling time. Eur Urol 2008;54:291-300.
179. O'Brien MF, Cronin AM, Fearn PA, et al. Pretreatment prostate-specific antigen (PSA) velocity and doubling time are associated with outcome but neither improves prediction of outcome beyond pretreatment PSA alone in patients treated with radical prostatectomy. J Clin Oncol 2009;27:3591-7.
180. Vickers AJ, Savage C, O'Brien MF, Lilja H. Systematic review of pretreatment prostate-specific antigen velocity and doubling time as predictors for prostate cancer. J Clin Oncol 2009;27:398-403.
181. Catalona WJ, Bartsch G, Rittenhouse HG, et al. Serum pro prostate specific antigen improves cancer detection compared to free and complexed prostate specific antigen in men with prostate specific antigen 2 to 4 ng/ml. J Urol 2003;170:2181-5.
182. Nurmikko P, Pettersson K, Piironen T, Hugosson J, Lilja H. Discrimination of prostate cancer from benign disease by plasma measurement of intact, free prostate-specific antigen lacking an internal cleavage site at Lys145-Lys146. Clin Chem 2001;47:1415-23.
183. Slawin KM, Shariat S, Canto E. BPSA: A Novel Serum Marker for Benign Prostatic Hyperplasia. Rev Urol 2005;7 Suppl 8:S52-6.
184. Amling CL, Bergstralh EJ, Blute ML, Slezak JM, Zincke H. Defining prostate specific antigen progression after radical prostatectomy: what is the most appropriate cut point? J Urol 2001;165:1146-51.
185. Kupelian P, Kuban D, Thames H, et al. Improved biochemical relapse-free survival with increased external radiation doses in patients with localized prostate cancer: the combined experience of nine institutions in patients treated in 1994 and 1995. Int J Radiat Oncol Biol Phys 2005;61:415-9.
186. Hanlon AL, Moore DF, Hanks GE. Modeling postradiation prostate specific antigen level kinetics: predictors of rising postnadir slope suggest cure in men who remain biochemically free of prostate carcinoma. Cancer 1998;83:130-4.
187. Stamey TA, Kabalin JN, Ferrari M, Yang N. Prostate specific antigen in the diagnosis and treatment of adenocarcinoma of the prostate. IV. Anti-androgen treated patients. J Urol 1989;141:1088-90.
188. Ryan CJ, Smith A, Lal P, et al. Persistent prostate-specific antigen expression after neoadjuvant androgen depletion: an early predictor of relapse or incomplete androgen suppression. Urology 2006;68:834-9.
189. Benaim EA, Pace CM, Lam PM, Roehrborn CG. Nadir prostate-specific antigen as a predictor of progression to androgen-independent prostate cancer. Urology 2002;59:73-8.
190. Scher HI, Sawyers CL. Biology of progressive, castration-resistant prostate cancer: directed therapies targeting the androgen-receptor signaling axis. J Clin Oncol 2005;23:8253-61.

191. Catalona WJ, Richie JP, Ahmann FR, et al. Comparison of digital rectal examination and serum prostate specific antigen in the early detection of prostate cancer: results of a multicenter clinical trial of 6,630 men. J Urol 1994;151:1283-90.
192. Mistry K, Cable G. Meta-analysis of prostate-specific antigen and digital rectal examination as screening tests for prostate carcinoma. J Am Board Fam Pract 2003;16:95-101.
193. Chodak GW, Keller P, Schoenberg HW. Assessment of screening for prostate cancer using the digital rectal examination. J Urol 1989;141:1136-8.
194. Richie JP, Catalona WJ, Ahmann FR, et al. Effect of patient age on early detection of prostate cancer with serum prostate-specific antigen and digital rectal examination. Urology 1993;42:365-74.
195. Presti JC, Jr., O'Dowd GJ, Miller MC, Mattu R, Veltri RW. Extended peripheral zone biopsy schemes increase cancer detection rates and minimize variance in prostate specific antigen and age related cancer rates: results of a community multi-practice study. J Urol 2003;169:125-9.
196. Remzi M, Fong YK, Dobrovits M, et al. The Vienna nomogram: validation of a novel biopsy strategy defining the optimal number of cores based on patient age and total prostate volume. J Urol 2005;174:1256-60; discussion 60-1; author reply 61.
197. Merrimen JL, Jones G, Walker D, Leung CS, Kapusta LR, Srigley JR. Multifocal high grade prostatic intraepithelial neoplasia is a significant risk factor for prostatic adenocarcinoma. J Urol 2009;182:485-90; discussion 90.
198. Conrad S, Graefen M, Pichlmeier U, et al. Prospective validation of an algorithm with systematic sextant biopsy to predict pelvic lymph node metastasis in patients with clinically localized prostatic carcinoma. J Urol 2002;167:521-5.
199. Bostwick DG, Grignon DJ, Hammond ME, et al. Prognostic factors in prostate cancer. College of American Pathologists Consensus Statement 1999. Arch Pathol Lab Med 2000;124:995-1000.
200. You J, Cozzi P, Walsh B, et al. Innovative biomarkers for prostate cancer early diagnosis and progression. Crit Rev Oncol Hematol;73:10-22.
201. Tuma RS. New tests for prostate cancer may be nearing the clinic. J Natl Cancer Inst;102:752-4.
202. Thompson IM, Ankerst DP, Chi C, et al. Operating characteristics of prostate-specific antigen in men with an initial PSA level of 3.0 ng/ml or lower. JAMA 2005;294:66-70.
203. Schroder FH, van der Cruijsen-Koeter I, de Koning HJ, Vis AN, Hoedemaeker RF, Kranse R. Prostate cancer detection at low prostate specific antigen. J Urol 2000;163:806-12.
204. Wolf AM, Wender RC, Etzioni RB, et al. American Cancer Society guideline for the early detection of prostate cancer: update 2010. CA Cancer J Clin;60:70-98.
205. Albertsen PC, Hanley JA, Fine J. 20-year outcomes following conservative management of clinically localized prostate cancer. JAMA 2005;293:2095-101.
206. Telesca D, Etzioni R, Gulati R. Estimating lead time and overdiagnosis associated with PSA screening from prostate cancer incidence trends. Biometrics 2008;64:10-9.
207. Catalona WJ, Richie JP, deKernion JB, et al. Comparison of prostate specific antigen concentration versus prostate specific antigen density in the early detection of prostate cancer: receiver operating characteristic curves. J Urol 1994;152:2031-6.
208. Sartor AO, Hricak H, Wheeler TM, et al. Evaluating localized prostate cancer and identifying candidates for focal therapy. Urology 2008;72:S12-24.
209. D'Amico AV, Whittington R, Malkowicz SB, et al. Biochemical outcome after radical prostatectomy, external beam radiation therapy, or interstitial radiation therapy for clinically localized prostate cancer. JAMA 1998;280:969-74.
210. Andriole GL, Crawford ED, Grubb RL, 3rd, et al. Mortality results from a randomized prostate-cancer screening trial. N Engl J Med 2009;360:1310-9.

211. Schroder FH, Hugosson J, Roobol MJ, et al. Screening and prostate-cancer mortality in a randomized European study. N Engl J Med 2009;360:1320-8.
212. Biomarkers and surrogate endpoints: preferred definitions and conceptual framework. Clin Pharmacol Ther 2001;69:89-95.
213. Gerszten RE, Wang TJ. The search for new cardiovascular biomarkers. Nature 2008;451:949-52.
214. Boja E, Hiltke T, Rivers R, et al. Evolution of Clinical Proteomics and its Role in Medicine. J Proteome Res 2010.
215. Anderson L. Candidate-based proteomics in the search for biomarkers of cardiovascular disease. J Physiol 2005;563:23-60.
216. Gutman S, Kessler LG. The US Food and Drug Administration perspective on cancer biomarker development. Nat Rev Cancer 2006;6:565-71.
217. Rifai N, Gillette MA, Carr SA. Protein biomarker discovery and validation: the long and uncertain path to clinical utility. Nat Biotechnol 2006;24:971-83.
218. Pepe MS, Etzioni R, Feng Z, et al. Phases of biomarker development for early detection of cancer. J Natl Cancer Inst 2001;93:1054-61.
219. Barker PE. Cancer biomarker validation: standards and process: roles for the National Institute of Standards and Technology (NIST). Ann N Y Acad Sci 2003;983:142-50.
220. Madu CO, Lu Y. Novel diagnostic biomarkers for prostate cancer. J Cancer;1:150-77.
221. Iorio L, Avagliano F. Observations on the Liber medicine orinalibus by Hermogenes. American journal of nephrology 1999;19:185-8.
222. Moe OW, Preisig PA. A better understanding of the kidney in health and disease: what will it take? Current opinion in nephrology and hypertension 2000;9:381-4.
223. Gonzalez-Buitrago JM, Ferreira L, Lorenzo I. Urinary proteomics. Clin Chim Acta 2007;375:49-56.
224. Thongboonkerd V, Malasit P. Renal and urinary proteomics: current applications and challenges. Proteomics 2005;5:1033-42.
225. Thongboonkerd V. Urinary proteomics: towards biomarker discovery, diagnostics and prognostics. Mol Biosyst 2008;4:810-5.
226. Mischak H, Julian BA, Novak J. High-resolution proteome/peptidome analysis of peptides and low-molecular-weight proteins in urine. Proteomics Clin Appl 2007;1:792.
227. Graff VD. Human Anatomy. Sixth Edition ed; 2001.
228. Barratt J, Topham P. Urine proteomics: the present and future of measuring urinary protein components in disease. CMAJ 2007;177:361-8.
229. Omenn GS, States DJ, Adamski M, et al. Overview of the HUPO Plasma Proteome Project: results from the pilot phase with 35 collaborating laboratories and multiple analytical groups, generating a core dataset of 3020 proteins and a publicly-available database. Proteomics 2005;5:3226-45.
230. Kolch W, Neususs C, Pelzing M, Mischak H. Capillary electrophoresis-mass spectrometry as a powerful tool in clinical diagnosis and biomarker discovery. Mass spectrometry reviews 2005;24:959-77.
231. Fliser D, Wittke S, Mischak H. Capillary electrophoresis coupled to mass spectrometry for clinical diagnostic purposes. Electrophoresis 2005;26:2708-16.
232. Anderson NL, Anderson NG. The human plasma proteome: history, character, and diagnostic prospects. Mol Cell Proteomics 2002;1:845-67.
233. Edwards JJ, Tollaksen SL, Anderson NG. Proteins of human urine. III. Identification and two-dimensional electrophoretic map positions of some major urinary proteins. Clin Chem 1982;28:941-8.
234. Castagna A, Cecconi D, Sennels L, et al. Exploring the hidden human urinary proteome via ligand library beads. J Proteome Res 2005;4:1917-30.
235. Havanapan PO, Thongboonkerd V. Are protease inhibitors required for gel-based proteomics of kidney and urine? J Proteome Res 2009;8:3109-17.

236. Drake RR, White KY, Fuller TW, et al. Clinical collection and protein properties of expressed prostatic secretions as a source for biomarkers of prostatic disease. J Proteomics 2009;72:907-17.
237. Foot NC, Papanicolaou GN, Holmquist ND, Seybolt JF. Exfoliative cytology of urinary sediments; a review of 2,829 cases. Cancer 1958;11:127-37.
238. Krishnan B, Truong LD. Prostatic adenocarcinoma diagnosed by urinary cytology. Am J Clin Pathol 2000;113:29-34.
239. Koss LG. Cytology of specimens obtained by prostatic massage, Diagnostic Cytology and its Histopathologic Bases. JB Lippincott Company, Philadelphia 1992:pp.1001-5.
240. Ploussard G, de la Taille A. Urine biomarkers in prostate cancer. Nat Rev Urol;7:101-9.
241. Roupret M, Hupertan V, Yates DR, et al. Molecular detection of localized prostate cancer using quantitative methylation-specific PCR on urinary cells obtained following prostate massage. Clin Cancer Res 2007;13:1720-5.
242. Hoque MO, Topaloglu O, Begum S, et al. Quantitative methylation-specific polymerase chain reaction gene patterns in urine sediment distinguish prostate cancer patients from control subjects. J Clin Oncol 2005;23:6569-75.
243. Nilsson J, Skog J, Nordstrand A, et al. Prostate cancer-derived urine exosomes: a novel approach to biomarkers for prostate cancer. Br J Cancer 2009;100:1603-7.
244. Hessels D, Klein Gunnewiek JM, van Oort I, et al. DD3(PCA3)-based molecular urine analysis for the diagnosis of prostate cancer. Eur Urol 2003;44:8-15; discussion -6.
245. Groskopf J, Aubin SM, Deras IL, et al. APTIMA PCA3 molecular urine test: development of a method to aid in the diagnosis of prostate cancer. Clin Chem 2006;52:1089-95.
246. Xu LL, Stackhouse BG, Florence K, et al. PSGR, a novel prostate-specific gene with homology to a G protein-coupled receptor, is overexpressed in prostate cancer. Cancer Res 2000;60:6568-72.
247. Xu LL, Sun C, Petrovics G, et al. Quantitative expression profile of PSGR in prostate cancer. Prostate Cancer Prostatic Dis 2006;9:56-61.
248. Rigau M, Morote J, Mir MC, et al. PSGR and PCA3 as biomarkers for the detection of prostate cancer in urine. Prostate 2010;70:1760-7.
249. Horoszewicz JS, Kawinski E, Murphy GP. Monoclonal antibodies to a new antigenic marker in epithelial prostatic cells and serum of prostatic cancer patients. Anticancer Res 1987;7:927-35.
250. Beckett ML, Cazares LH, Vlahou A, Schellhammer PF, Wright GL, Jr. Prostate-specific membrane antigen levels in sera from healthy men and patients with benign prostate hyperplasia or prostate cancer. Clin Cancer Res 1999;5:4034-40.
251. Sreekumar A, Laxman B, Rhodes DR, et al. Humoral immune response to alpha-methylacyl-CoA racemase and prostate cancer. J Natl Cancer Inst 2004;96:834-43.
252. Zehentner BK, Secrist H, Zhang X, et al. Detection of alpha-methylacyl-coenzyme-A racemase transcripts in blood and urine samples of prostate cancer patients. Mol Diagn Ther 2006;10:397-403.
253. Tomlins SA, Rhodes DR, Perner S, et al. Recurrent fusion of TMPRSS2 and ETS transcription factor genes in prostate cancer. Science 2005;310:644-8.
254. Hessels D, Smit FP, Verhaegh GW, Witjes JA, Cornel EB, Schalken JA. Detection of TMPRSS2-ERG fusion transcripts and prostate cancer antigen 3 in urinary sediments may improve diagnosis of prostate cancer. Clin Cancer Res 2007;13:5103-8.
255. Wozny W, Schroer K, Schwall GP, et al. Differential radioactive quantification of protein abundance ratios between benign and malignant prostate tissues: cancer association of annexin A3. Proteomics 2007;7:313-22.
256. Gerke V, Creutz CE, Moss SE. Annexins: linking Ca2+ signalling to membrane dynamics. Nat Rev Mol Cell Biol 2005;6:449-61.

257. Pisitkun T, Shen RF, Knepper MA. Identification and proteomic profiling of exosomes in human urine. Proceedings of the National Academy of Sciences of the United States of America 2004;101:13368-73.
258. Schostak M, Schwall GP, Poznanovic S, et al. Annexin A3 in urine: a highly specific noninvasive marker for prostate cancer early detection. J Urol 2009;181:343-53.
259. Meid FH, Gygi CM, Leisinger HJ, Bosman FT, Benhattar J. The use of telomerase activity for the detection of prostatic cancer cells after prostatic massage. J Urol 2001;165:1802-5.
260. Vicentini C, Gravina GL, Angelucci A, et al. Detection of telomerase activity in prostate massage samples improves differentiating prostate cancer from benign prostatic hyperplasia. J Cancer Res Clin Oncol 2004;130:217-21.
261. Sreekumar A, Poisson LM, Rajendiran TM, et al. Metabolomic profiles delineate potential role for sarcosine in prostate cancer progression. Nature 2009;457:910-4.
262. Laxman B, Morris DS, Yu J, et al. A first-generation multiplex biomarker analysis of urine for the early detection of prostate cancer. Cancer Res 2008;68:645-9.
263. Varambally S, Laxman B, Mehra R, et al. Golgi protein GOLM1 is a tissue and urine biomarker of prostate cancer. Neoplasia 2008;10:1285-94.
264. Darson MF, Pacelli A, Roche P, et al. Human glandular kallikrein 2 (hK2) expression in prostatic intraepithelial neoplasia and adenocarcinoma: a novel prostate cancer marker. Urology 1997;49:857-62.
265. Haese A, Graefen M, Steuber T, et al. Human glandular kallikrein 2 levels in serum for discrimination of pathologically organ-confined from locally-advanced prostate cancer in total PSA-levels below 10 ng/ml. Prostate 2001;49:101-9.
266. Gutman ABaG, E.B. An "acid" phosphataseoccuring in the serum of patients with metastasizing carcinoma of the prostate gland. J Biol Chem 1938;136:1938.
267. Gu Z, Thomas G, Yamashiro J, et al. Prostate stem cell antigen (PSCA) expression increases with high gleason score, advanced stage and bone metastasis in prostate cancer. Oncogene 2000;19:1288-96.
268. Gu Z, Yamashiro J, Kono E, Reiter RE. Anti-prostate stem cell antigen monoclonal antibody 1G8 induces cell death in vitro and inhibits tumor growth in vivo via a Fc-independent mechanism. Cancer Res 2005;65:9495-500.
269. Shariat SF, Andrews B, Kattan MW, Kim J, Wheeler TM, Slawin KM. Plasma levels of interleukin-6 and its soluble receptor are associated with prostate cancer progression and metastasis. Urology 2001;58:1008-15.
270. Nakashima J, Tachibana M, Horiguchi Y, et al. Serum interleukin 6 as a prognostic factor in patients with prostate cancer. Clin Cancer Res 2000;6:2702-6.
271. Shariat SF, Kattan MW, Traxel E, et al. Association of pre- and postoperative plasma levels of transforming growth factor beta(1) and interleukin 6 and its soluble receptor with prostate cancer progression. Clin Cancer Res 2004;10:1992-9.
272. McCabe NP, Angwafo FF, 3rd, Zaher A, Selman SH, Kouinche A, Jankun J. Expression of soluble urokinase plasminogen activator receptor may be related to outcome in prostate cancer patients. Oncol Rep 2000;7:879-82.
273. Shariat SF, Roehrborn CG, McConnell JD, et al. Association of the circulating levels of the urokinase system of plasminogen activation with the presence of prostate cancer and invasion, progression, and metastasis. J Clin Oncol 2007;25:349-55.
274. Paul B, Dhir R, Landsittel D, Hitchens MR, Getzenberg RH. Detection of prostate cancer with a blood-based assay for early prostate cancer antigen. Cancer Res 2005;65:4097-100.
275. Leman ES, Cannon GW, Trock BJ, et al. EPCA-2: a highly specific serum marker for prostate cancer. Urology 2007;69:714-20.
276. Kallioniemi O. Functional genomics and transcriptomics of prostate cancer: promises and limitations. BJU Int 2005;96 Suppl 2:10-5.

277. Sun QP, Li LY, Chen Z, et al. Detection of TMPRSS2-ETS Fusions by a Multiprobe Fluorescence in Situ Hybridization Assay for the Early Diagnosis of Prostate Cancer. A Pilot Study. J Mol Diagn 2010.
278. Saramaki OR, Harjula AE, Martikainen PM, Vessella RL, Tammela TL, Visakorpi T. TMPRSS2:ERG fusion identifies a subgroup of prostate cancers with a favorable prognosis. Clin Cancer Res 2008;14:3395-400.
279. Mosquera JM, Mehra R, Regan MM, et al. Prevalence of TMPRSS2-ERG fusion prostate cancer among men undergoing prostate biopsy in the United States. Clin Cancer Res 2009;15:4706-11.
280. Laxman B, Tomlins SA, Mehra R, et al. Noninvasive detection of TMPRSS2:ERG fusion transcripts in the urine of men with prostate cancer. Neoplasia 2006;8:885-8.
281. Vener T, Derecho C, Baden J, et al. Development of a multiplexed urine assay for prostate cancer diagnosis. Clin Chem 2008;54:874-82.
282. Marks LS, Bostwick DG. Prostate Cancer Specificity of PCA3 Gene Testing: Examples from Clinical Practice. Rev Urol 2008;10:175-81.
283. Rigau M, Morote J, Mir MC, et al. PSGR and PCA3 as Biomarkers for the Detection of Prostate Cancer in Urine. Prostate 2010;in press.
284. Etzioni R, Kooperberg C, Pepe M, Smith R, Gann PH. Combining biomarkers to detect disease with application to prostate cancer. Biostatistics 2003;4:523-38.
285. Bussemakers MJ, van Bokhoven A, Verhaegh GW, et al. DD3: a new prostate-specific gene, highly overexpressed in prostate cancer. Cancer Res 1999;59:5975-9.
286. Clarke RA, Zhao Z, Guo AY, et al. New genomic structure for prostate cancer specific gene PCA3 within BMCC1: implications for prostate cancer detection and progression. PLoS One 2009;4:e4995.
287. de Kok JB, Verhaegh GW, Roelofs RW, et al. DD3(PCA3), a very sensitive and specific marker to detect prostate tumors. Cancer Res 2002;62:2695-8.
288. Costa FF. Non-coding RNAs: new players in eukaryotic biology. Gene 2005;357:83-94.
289. Mattick JS, Makunin IV. Non-coding RNA. Hum Mol Genet 2006;15 Spec No 1:R17-29.
290. Szell M, Bata-Csorgo Z, Kemeny L. The enigmatic world of mRNA-like ncRNAs: their role in human evolution and in human diseases. Semin Cancer Biol 2008;18:141-8.
291. Turner AM, Morris KV. Controlling transcription with noncoding RNAs in mammalian cells. Biotechniques;48:ix-xvi.
292. Meng FJ, Shan A, Jin L, Young CY. The expression of a variant prostate-specific antigen in human prostate. Cancer Epidemiol Biomarkers Prev 2002;11:305-9.
293. van Gils MP, Hessels D, van Hooij O, et al. The time-resolved fluorescence-based PCA3 test on urinary sediments after digital rectal examination; a Dutch multicenter validation of the diagnostic performance. Clin Cancer Res 2007;13:939-43.
294. Haese A, de la Taille A, van Poppel H, et al. Clinical utility of the PCA3 urine assay in European men scheduled for repeat biopsy. Eur Urol 2008;54:1081-8.
295. Deras IL, Aubin SM, Blase A, et al. PCA3: a molecular urine assay for predicting prostate biopsy outcome. J Urol 2008;179:1587-92.
296. Schilling D, Hennenlotter J, Munz M, Bokeler U, Sievert KD, Stenzl A. Interpretation of the prostate cancer gene 3 in reference to the individual clinical background: implications for daily practice. Urol Int;85:159-65.
297. Hessels D, Schalken JA. The use of PCA3 in the diagnosis of prostate cancer. Nat Rev Urol 2009;6:255-61.
298. Aubin SM, Reid J, Sarno MJ, et al. PCA3 molecular urine test for predicting repeat prostate biopsy outcome in populations at risk: validation in the placebo arm of the dutasteride REDUCE trial. J Urol;184:1947-52.
299. Andriole G, Bostwick D, Brawley O, et al. Chemoprevention of prostate cancer in men at high risk: rationale and design of the reduction by dutasteride of prostate cancer events (REDUCE) trial. J Urol 2004;172:1314-7.

300. Nakanishi H, Groskopf J, Fritsche HA, et al. PCA3 molecular urine assay correlates with prostate cancer tumor volume: implication in selecting candidates for active surveillance. J Urol 2008;179:1804-9; discussion 9-10.
301. Whitman EJ, Groskopf J, Ali A, et al. PCA3 score before radical prostatectomy predicts extracapsular extension and tumor volume. J Urol 2008;180:1975-8; discussion 8-9.
302. Day JR, Jost M, Reynolds MA, Groskopf J, Rittenhouse H. PCA3: from basic molecular science to the clinical lab. Cancer Lett;301:1-6.
303. Yuan TT, Toy P, McClary JA, Lin RJ, Miyamoto NG, Kretschmer PJ. Cloning and genetic characterization of an evolutionarily conserved human olfactory receptor that is differentially expressed across species. Gene 2001;278:41-51.
304. Weng J, Ma W, Mitchell D, Zhang J, Liu M. Regulation of human prostate-specific G-protein coupled receptor, PSGR, by two distinct promoters and growth factors. J Cell Biochem 2005;96:1034-48.
305. Neuhaus EM, Zhang W, Gelis L, Deng Y, Noldus J, Hatt H. Activation of an olfactory receptor inhibits proliferation of prostate cancer cells. J Biol Chem 2009;284:16218-25.
306. Cotton M, Claing A. G protein-coupled receptors stimulation and the control of cell migration. Cell Signal 2009;21:1045-53.
307. Xia C, Ma W, Wang F, Hua S, Liu M. Identification of a prostate-specific G-protein coupled receptor in prostate cancer. Oncogene 2001;20:5903-7.
308. Ashida S, Nakagawa H, Katagiri T, et al. Molecular features of the transition from prostatic intraepithelial neoplasia (PIN) to prostate cancer: genome-wide gene-expression profiles of prostate cancers and PINs. Cancer Res 2004;64:5963-72.
309. Weng J, Wang J, Cai Y, et al. Increased expression of prostate-specific G-protein-coupled receptor in human prostate intraepithelial neoplasia and prostate cancers. Int J Cancer 2005;113:811-8.
310. Cardillo MR, Di Silverio F. Prostate--specific G protein couple receptor genes and STAG1/PMEPA1 in peripheral blood from patients with prostatic cancer. Int J Immunopathol Pharmacol 2006;19:871-8.
311. Rigau M, Ortega I, Mir MC, et al. A Three-Gene panel on urine increases PSA specificity in the detection of prostate cancer. Prostate.
312. Israeli RS, Powell CT, Fair WR, Heston WD. Molecular cloning of a complementary DNA encoding a prostate-specific membrane antigen. Cancer Res 1993;53:227-30.
313. Su SL, Huang IP, Fair WR, Powell CT, Heston WD. Alternatively spliced variants of prostate-specific membrane antigen RNA: ratio of expression as a potential measurement of progression. Cancer Res 1995;55:1441-3.
314. Ghosh A, Heston WD. Tumor target prostate specific membrane antigen (PSMA) and its regulation in prostate cancer. J Cell Biochem 2004;91:528-39.
315. Silver DA, Pellicer I, Fair WR, Heston WD, Cordon-Cardo C. Prostate-specific membrane antigen expression in normal and malignant human tissues. Clin Cancer Res 1997;3:81-5.
316. Olson WC, Heston WD, Rajasekaran AK. Clinical trials of cancer therapies targeting prostate-specific membrane antigen. Rev Recent Clin Trials 2007;2:182-90.
317. Sokoloff RL, Norton KC, Gasior CL, Marker KM, Grauer LS. A dual-monoclonal sandwich assay for prostate-specific membrane antigen: levels in tissues, seminal fluid and urine. Prostate 2000;43:150-7.
318. Xiao Z, Adam BL, Cazares LH, et al. Quantitation of serum prostate-specific membrane antigen by a novel protein biochip immunoassay discriminates benign from malignant prostate disease. Cancer Res 2001;61:6029-33.
319. Talesa VN, Antognelli C, Del Buono C, et al. Diagnostic potential in prostate cancer of a panel of urinary molecular tumor markers. Cancer Biomark 2009;5:241-51.
320. Zhang L, Wang CY, Yang R, et al. Real-time quantitative RT-PCR assay of prostate-specific antigen and prostate-specific membrane antigen in peripheral blood for detection of prostate cancer micrometastasis. Urol Oncol 2008;26:634-40.

321. Elgamal AA, Holmes EH, Su SL, et al. Prostate-specific membrane antigen (PSMA): current benefits and future value. Semin Surg Oncol 2000;18:10-6.
322. Mitchell PJ, Welton J, Staffurth J, et al. Can urinary exosomes act as treatment response markers in prostate cancer? J Transl Med 2009;7:4.
323. Landers KA, Burger MJ, Tebay MA, et al. Use of multiple biomarkers for a molecular diagnosis of prostate cancer. Int J Cancer 2005;114:950-6.
324. Schmidt U, Fuessel S, Koch R, et al. Quantitative multi-gene expression profiling of primary prostate cancer. Prostate 2006;66:1521-34.
325. Kreunin P, Zhao J, Rosser C, Urquidi V, Lubman DM, Goodison S. Bladder cancer associated glycoprotein signatures revealed by urinary proteomic profiling. J Proteome Res 2007;6:2631-9.
326. Masters JR. Prostate cancer proteomics. OMICS;15:169-71.
327. Hamburg MA, Collins FS. The path to personalized medicine. N Engl J Med 2010;363:301-4.
328. Anderson NL, Anderson NG, Pearson TW, et al. A human proteome detection and quantitation project. Mol Cell Proteomics 2009;8:883-6.
329. Pavlou MP, Diamandis EP. The cancer cell secretome: A good source for discovering biomarkers? J Proteomics 2010.
330. Makridakis M, Vlahou A. Secretome proteomics for discovery of cancer biomarkers. J Proteomics;73:2291-305.
331. May M. From cells, secrets of the secretome leak out. Nat Med 2009;15:828.
332. Xue H, Lu B, Zhang J, et al. Identification of serum biomarkers for colorectal cancer metastasis using a differential secretome approach. J Proteome Res;9:545-55.
333. Sardana G, Jung K, Stephan C, Diamandis EP. Proteomic analysis of conditioned media from the PC3, LNCaP, and 22Rv1 prostate cancer cell lines: discovery and validation of candidate prostate cancer biomarkers. J Proteome Res 2008;7:3329-38.
334. Sardana G, Marshall J, Diamandis EP. Discovery of candidate tumor markers for prostate cancer via proteomic analysis of cell culture-conditioned medium. Clin Chem 2007;53:429-37.
335. Martin DB, Gifford DR, Wright ME, et al. Quantitative proteomic analysis of proteins released by neoplastic prostate epithelium. Cancer Res 2004;64:347-55.
336. Goo YA, Liu AY, Ryu S, et al. Identification of secreted glycoproteins of human prostate and bladder stromal cells by comparative quantitative proteomics. Prostate 2009;69:49-61.
337. Domon B, Aebersold R. Mass spectrometry and protein analysis. Science 2006;312:212-7.
338. Decramer S, Gonzalez de Peredo A, Breuil B, et al. Urine in clinical proteomics. Mol Cell Proteomics 2008;7:1850-62.
339. Fliser D, Novak J, Thongboonkerd V, et al. Advances in urinary proteome analysis and biomarker discovery. J Am Soc Nephrol 2007;18:1057-71.
340. Wang H, Wong CH, Chin A, Kennedy J, Zhang Q, Hanash S. Quantitative serum proteomics using dual stable isotope coding and nano LC-MS/MSMS. J Proteome Res 2009;8:5412-22.
341. Wilkins MR, Pasquali C, Appel RD, et al. From proteins to proteomes: large scale protein identification by two-dimensional electrophoresis and amino acid analysis. Biotechnology (N Y) 1996;14:61-5.
342. O'Farrell PH. High resolution two-dimensional electrophoresis of proteins. J Biol Chem 1975;250:4007-21.
343. Alban A, David SO, Bjorkesten L, et al. A novel experimental design for comparative two-dimensional gel analysis: two-dimensional difference gel electrophoresis incorporating a pooled internal standard. Proteomics 2003;3:36-44.

344. Lafitte D, Dussol B, Andersen S, et al. Optimized preparation of urine samples for two-dimensional electrophoresis and initial application to patient samples. Clin Biochem 2002;35:581-9.
345. Ong SE, Mann M. Mass spectrometry-based proteomics turns quantitative. Nature chemical biology 2005;1:252-62.
346. Aebersold R, Mann M. Mass spectrometry-based proteomics. Nature 2003;422:198-207.
347. Hubbard SJ. Computational approaches to peptide identification via tandem MS. Methods Mol Biol 2010;604:23-42.
348. Allet N, Barrillat N, Baussant T, et al. In vitro and in silico processes to identify differentially expressed proteins. Proteomics 2004;4:2333-51.
349. Karas M, Hillenkamp F. Laser desorption ionization of proteins with molecular masses exceeding 10,000 daltons. Anal Chem 1988;60:2299-301.
350. Issaq HJ, Conrads TP, Prieto DA, Tirumalai R, Veenstra TD. SELDI-TOF MS for diagnostic proteomics. Anal Chem 2003;75:148A-55A.
351. Li J, White N, Zhang Z, et al. Detection of prostate cancer using serum proteomics pattern in a histologically confirmed population. J Urol 2004;171:1782-7.
352. Brun V, Masselon C, Garin J, Dupuis A. Isotope dilution strategies for absolute quantitative proteomics. J Proteomics 2009;72:740-9.
353. Engvall E, Perlmann P. Enzyme-linked immunosorbent assay (ELISA). Quantitative assay of immunoglobulin G. Immunochemistry 1971;8:871-4.
354. Kozarova A, Petrinac S, Ali A, Hudson JW. Array of informatics: Applications in modern research. J Proteome Res 2006;5:1051-9.
355. Coons AC, HJ. Jones, RN. Immunologic properties of an antibody containing a fluorescent group. Proc Soc Exp Biol 1941;47.
356. Schiess R, Wollscheid B, Aebersold R. Targeted proteomic strategy for clinical biomarker discovery. Mol Oncol 2009;3:33-44.
357. Lange V, Picotti P, Domon B, Aebersold R. Selected reaction monitoring for quantitative proteomics: a tutorial. Mol Syst Biol 2008;4:222.
358. Gallien S, Duriez E, Domon B. Selected reaction monitoring applied to proteomics. J Mass Spectrom;46:298-312.
359. Baty JD, Robinson PR. Single and multiple ion recording techniques for the analysis of diphenylhydantoin and its major metabolite in plasma. Biomed Mass Spectrom 1977;4:36-41.
360. Stahl-Zeng J, Lange V, Ossola R, et al. High sensitivity detection of plasma proteins by multiple reaction monitoring of N-glycosites. Mol Cell Proteomics 2007;6:1809-17.
361. Theodorescu D, Mischak H. Mass spectrometry based proteomics in urine biomarker discovery. World J Urol 2007;25:435-43.
362. Pisitkun T, Johnstone R, Knepper MA. Discovery of urinary biomarkers. Mol Cell Proteomics 2006;5:1760-71.
363. Thongboonkerd V. Recent progress in urinary proteomics. Proteomics Clin Appl 2007;1:780-91.
364. Spahr CS, Davis MT, McGinley MD, et al. Towards defining the urinary proteome using liquid chromatography-tandem mass spectrometry. I. Profiling an unfractionated tryptic digest. Proteomics 2001;1:93-107.
365. Pieper R, Gatlin CL, McGrath AM, et al. Characterization of the human urinary proteome: a method for high-resolution display of urinary proteins on two-dimensional electrophoresis gels with a yield of nearly 1400 distinct protein spots. Proteomics 2004;4:1159-74.
366. Adachi J, Kumar C, Zhang Y, Olsen JV, Mann M. The human urinary proteome contains more than 1500 proteins, including a large proportion of membrane proteins. Genome Biol 2006;7:R80.

367. Coon JJ, Zurbig P, Dakna M, et al. CE-MS analysis of the human urinary proteome for biomarker discovery and disease diagnostics. Proteomics Clin Appl 2008;2:964.
368. Zolg W. The proteomic search for diagnostic biomarkers: lost in translation? Mol Cell Proteomics 2006;5:1720-6.
369. Nagaraj N, Mann M. Quantitative analysis of the intra- and inter-individual variability of the normal urinary proteome. J Proteome Res;10:637-45.
370. Selevsek N, Matondo M, Carbayo MS, Aebersold R, Domon B. Systematic quantification of peptides/proteins in urine using selected reaction monitoring. Proteomics;11:1135-47.
371. States DJ, Omenn GS, Blackwell TW, et al. Challenges in deriving high-confidence protein identifications from data gathered by a HUPO plasma proteome collaborative study. Nat Biotechnol 2006;24:333-8.
372. Sun WC, Y. Li, F.X. Zhang, L. Yang, R.F. Zhang, Z. Zheng, D.X. Gao, Y.H. Dynamic urinary proteomic analysis reveals stable proteins to be potential biomarkers. Proteomics Clin Appl 2009;3:370-82.
373. Domon B, Aebersold R. Challenges and opportunities in proteomics data analysis. Mol Cell Proteomics 2006;5:1921-6.
374. Bolduc S, Lacombe L, Naud A, Gregoire M, Fradet Y, Tremblay RR. Urinary PSA: a potential useful marker when serum PSA is between 2.5 ng/mL and 10 ng/mL. Can Urol Assoc J 2007;1:377-81.
375. Takayama TK, Vessella RL, Brawer MK, True LD, Noteboom J, Lange PH. Urinary prostate specific antigen levels after radical prostatectomy. J Urol 1994;151:82-7.
376. DeVere White RW, Meyers FJ, Soares SE, Miller DG, Soriano TF. Urinary prostate specific antigen levels: role in monitoring the response of prostate cancer to therapy. J Urol 1992;147:947-51.
377. Pannek J, Rittenhouse HG, Evans CL, et al. Molecular forms of prostate-specific antigen and human kallikrein 2 (hK2) in urine are not clinically useful for early detection and staging of prostate cancer. Urology 1997;50:715-21.
378. Irani J, Salomon L, Soulie M, et al. Urinary/serum prostate-specific antigen ratio: comparison with free/total serum prostate-specific antigen ratio in improving prostate cancer detection. Urology 2005;65:533-7.
379. Malavaud B, Salama G, Miedouge M, et al. Influence of digital rectal massage on urinary prostate-specific antigen: interest for the detection of local recurrence after radical prostatectomy. Prostate 1998;34:23-8.
380. Rehman I, Azzouzi AR, Catto JW, et al. Proteomic analysis of voided urine after prostatic massage from patients with prostate cancer: a pilot study. Urology 2004;64:1238-43.
381. Theodorescu D, Fliser D, Wittke S, et al. Pilot study of capillary electrophoresis coupled to mass spectrometry as a tool to define potential prostate cancer biomarkers in urine. Electrophoresis 2005;26:2797-808.
382. Theodorescu D, Schiffer E, Bauer HW, et al. Discovery and validation of urinary biomarkers for prostate cancer. Proteomics Clin Appl 2008;2:556-70.
383. M'Koma AE, Blum DL, Norris JL, et al. Detection of pre-neoplastic and neoplastic prostate disease by MALDI profiling of urine. Biochem Biophys Res Commun 2007;353:829-34.
384. Okamoto A, Yamamoto H, Imai A, et al. Protein profiling of post-prostatic massage urine specimens by surface-enhanced laser desorption/ionization time-of-flight mass spectrometry to discriminate between prostate cancer and benign lesions. Oncol Rep 2009;21:73-9.
385. Djavan B, Remzi M, Schulman CC, Marberger M, Zlotta AR. Repeat prostate biopsy: who, how and when?. a review. Eur Urol 2002;42:93-103.

386. Iwakiri J, Granbois K, Wehner N, Graves HC, Stamey T. An analysis of urinary prostate specific antigen before and after radical prostatectomy: evidence for secretion of prostate specific antigen by the periurethral glands. J Urol 1993;149:783-6.
387. Jamaspishvili T, Kral M, Khomeriki I, Student V, Kolar Z, Bouchal J. Urine markers in monitoring for prostate cancer. Prostate Cancer Prostatic Dis;13:12-9.
388. Mischak H, Julian BA, Novak J. High-resolution proteome/peptidome analysis of peptides and low-molecular-weight proteins in urine. Proteomics Clin Appl 2007;1:792-804.
389. Bourdoumis A, Papatsoris AG, Chrisofos M, Efstathiou E, Skolarikos A, Deliveliotis C. The novel prostate cancer antigen 3 (PCA3) biomarker. Int Braz J Urol;36:665-8; discussion 9.
390. Pepe MS, Feng Z, Janes H, Bossuyt PM, Potter JD. Pivotal evaluation of the accuracy of a biomarker used for classification or prediction: standards for study design. J Natl Cancer Inst 2008;100:1432-8.
391. Shultz EK. Multivariate receiver-operating characteristic curve analysis: prostate cancer screening as an example. Clin Chem 1995;41:1248-55.
392. Steyerberg EW, Harrell FE, Jr., Borsboom GJ, Eijkemans MJ, Vergouwe Y, Habbema JD. Internal validation of predictive models: efficiency of some procedures for logistic regression analysis. J Clin Epidemiol 2001;54:774-81.
393. Harma V, Virtanen J, Makela R, et al. A comprehensive panel of three-dimensional models for studies of prostate cancer growth, invasion and drug responses. PLoS One;5:e10431.
394. Zlotta A, Ravic G, Schulman C. Clinical prognostic criteria for later diagnosis of prostate carcinoma in patients with initial isolated intraephitelial neoplasia. Eur Urol 1996;30:249-55.
395. Herawi M, Kahane H, Cavallo C. Risk of prostate cancer on first re-biopsy within 1 year following a diagnosis of high grade prostatic intraepithelial neoplasia is related to the number of cores
sampled. J Urol 2006;175:121-4.
396. Moore C, Karikehalli S, Nazeer T. Prognostic significance of high grade prostatic intraepithelial neoplasia and atypical small acinar proliferation in the contemporary era . J Urol 2005;173:70-2.
397. Morote J, Raventós C, Encabo G, López M, Torres Id. Effect of high grade prostatic intraepithelial neoplasia on total and percent free prostatic specific antigen. Eur Urol 2000;37:456-9.
398. Boja ES, Rodriguez H. The path to clinical proteomics research: integration of proteomics, genomics, clinical laboratory and regulatory science. Korean J Lab Med;31:61-71.
399. Lawlor K, Nazarian A, Lacomis L, Tempst P, Villanueva J. Pathway-based biomarker search by high-throughput proteomics profiling of secretomes. J Proteome Res 2009;8:1489-503.
400. Di Girolamo F, Boschetti E, Chung MC, Guadagni F, Righetti PG. "Proteomineering" or not? The debate on biomarker discovery in sera continues. J Proteomics;74:589-94.
401. Boschetti E, Righetti PG. The ProteoMiner in the proteomic arena: A non-depleting tool for discovering low-abundance species. J Proteomics 2008.
402. Kiyonami R, Schoen A, Prakash A, et al. Increased selectivity, analytical precision, and throughput in targeted proteomics. Mol Cell Proteomics;10:M110 002931.
403. Afkarian M, Bhasin M, Dillon ST, et al. Optimizing a proteomics platform for urine biomarker discovery. Mol Cell Proteomics 2010.
404. Court M, Selevsek N, Matondo M, et al. Toward a standardized urine proteome analysis methodology. Proteomics;11:1160-71.
405. Makarov DV, Loeb S, Getzenberg RH, Partin AW. Biomarkers for prostate cancer. Annu Rev Med 2009;60:139-51.
406. Hamburg MA, Collins FS. The path to personalized medicine. N Engl J Med;363:301-4.

# PUBLICATIONS

The Prostate

# PSGR and PCA3 as Biomarkers for the Detection of Prostate Cancer in Urine

Marina Rigau,[1] Juan Morote,[2,3] Maria Carmen Mir,[2] Carlos Ballesteros,[2] Israel Ortega,[4] Alex Sanchez,[4,5] Eva Colás,[1] Marta Garcia,[1] Anna Ruiz,[1] Miguel Abal,[6] Jacques Planas,[2] Jaume Reventós,[1,3] and Andreas Doll[1]*

[1]Biomedical Research Unit, Research Institute, Vall d'Hebron University Hospital, Barcelona, Spain
[2]Department of Urology, Vall d'Hebron University Hospital, Barcelona, Spain
[3]School of Medicine, Autonomous University of Barcelona, Spain
[4]Statistics and Bioinformatics Unit, Vall d'Hebron University Hospital, Barcelona, Spain
[5]Statistics Department, University of Barcelona, Barcelona, Spain
[6]Translational Oncology Laboratory, Instituto de Investigaciones Sanitarias, Complejo Hospitalario Universitario de Santiago (CHUS), Santiago de Compostela, Spain

**BACKGROUND.** Several studies have demonstrated the usefulness of monitoring an RNA transcript in urine, such as PCA3, for prostate cancer (PCa) diagnosis. PCa screening would benefit from additional biomarkers of higher specificity and could be used in conjunction with prostate-specific antigen (PSA) testing, in order to better determine biopsy candidates.
**METHODS.** We used urine sediments after prostate massage (PM) from 215 consecutive patients, who presented for prostate biopsy. We tested whether prostate-specific G-protein coupled receptor (PSGR), a biomarker previously described to be over-expressed in PCa tissue, could also be detected by quantitative real-time PCR in post-PM urine sediment. We combined these findings with prostate cancer gene 3 (PCA3), the current gold standard for PCa diagnosis in urine, to test if a combination of both biomarkers could improve the sensitivity of PCA3 alone.
**RESULTS.** By univariate analysis we found that PSGR and PCA3 were significant predictors of PCa. Receiver operator characteristic curve analysis and its multivariate extension, multivariate ROC (MultiROC), were used to assess the outcome predictive values of the individual and the paired biomarkers. We obtained the following area under the curve values: PSA (0.602), PSGR (0.681), PCA3 (0.656), and PSGRvPCA3 (0.729). Then, we tested whether a combination of PSGR and PCA3 could improve specificity by fixing the sensitivity at 95%. We obtained specificities of 15% (PSGR), 17% (PCA3), and 34% (PSGRvPCA3).
**CONCLUSIONS.** A multiplexed model including PSGR and PCA3 improves the specificity for the detection of PCa, especially in the area of high sensitivity. This could be clinically

Abbreviations: PCa, prostate cancer; PSA, prostate-specific antigen; PM, prostate massage; PSGR, prostate-specific G-protein coupled receptor; qPCR, quantitative real-time PCR; PCA3, prostate cancer gene 3; ROC, receiver operating characteristic; MultiROC, multivariate ROC; AUC, area under the curve; DRE, digital rectal exams; TRUS, transrectal ultrasound guided prostate biopsies; BPH, benign prostatic hyperplasia; OR, odorant receptor; G-proteins, GTP-binding proteins; GPCRs, G protein-coupled receptors; Ct, Threshold cycle; AUCm, area under each MultiROC curve; +PV, positive predictive value; −PV, negative predictive value.Jaume Reventós and Andreas Doll contributed equally to this study.

Grant sponsor: Instituto de Salud Carlos III, Ministerio de Ciéncia e Innovación; Grant number: RTICC RD06/0020/0058 PS 09-00496; Grant sponsor: Red de Genómica del Cáncer y Genotipado de tumores; Grant number: C03/10; Grant sponsor: Fundación para la Investigación en Urología (FIU); Grant sponsor: Recerca i de Societat de la Informació de la Generalitat de Catalunya; Grant number: SGR00487.

*Correspondence to: Andreas Doll, Biomedical Research Unit, Research Institute Vall d'Hebron University Hospital, Pg. Vall d'Hebron 119-129, 08035 Barcelona, Spain.
E-mail: adoll@ir.vhebron.net

Received 19 January 2010; Accepted 17 May 2010
DOI 10.1002/pros.21211
Published online in Wiley InterScience (www.interscience.wiley.com).

useful for determining which patients should undergo biopsy. *Prostate*


*KEY WORDS:* **prostate cancer; urine; biomarker; PSGR; PCA3**

## INTRODUCTION

Prostate cancer (PCa) has become the most common type of cancer in the western male population, where it is responsible for more male deaths than any other cancer, except lung cancer. Even though the introduction of the prostate-specific antigen (PSA) test in the late 1980s of the past century has led to a dramatic increase its detection [1], the risk of developing this type of cancer during a lifetime is estimated at one in six men, and the risk of death due to metastatic PCa is 1 in 30.

The diagnostic tools for detecting PCa can be separated into those that screen for the disease, such as PSA and digital rectal exams (DRE), and the decisive diagnosis set of transrectal ultrasound guided prostate biopsies (TRUS).

One of the limitations of serum PSA as a tumor marker is its lack of specificity, which results in a high rate of negative biopsies. There is a substantial overlap in the serum PSA values of men with benign prostatic hyperplasia (BPH) and those with PCa. Elevated PSA levels can also be attributed to other factors such as prostatitis, prostate irritation, and recent ejaculation [2,3]. As a consequence of the current screening parameters, around 2/3 of the approximately 1,300,000 biopsies made yearly in the United States (390,000 in Europe) are unnecessary [4,5]. In contrast, the false positive rate of a biopsy is about zero, although the false negative rate in the first biopsy may oscillate between 12%– and 32% [6,7]. Consequently, large populations of men with chronically elevated serum PSA undergo repeated biopsies to rule out PCa [2,3]. Various attempts have been made to overcome the limitations of PSA screening, including the use of ratio PSA, age-adjusted PSA cut-off points, free-PSA, PSA-density, PSA-velocity, PSA-slope and PSA-doubling time, and (−2) proPSA [8–10]. All have been proposed as a means of improving serum PSA specificity in the detection of PCa. Nevertheless, there has been no evidence to suggest that any of these testing strategies improves health outcomes [11].

There is an urgent need for the ability to detect PCa at an early stage using non-invasive procedures. A wide range of promising PCa biomarkers that are not only prostate-specific, but are also over-expressed in prostate tumors, have been identified, including CpG hypermethylation of GSTP1, the TMPRSS2: ERG gene fusion, AMACR, Anexina 3, Sarcosine, and RNA urine biomarker Prostate Cancer gene 3 (PCA3) [12–15]. As prostate cells can be detected in the urine of men with PCa, urine-based diagnostic tests have the advantage of being non- or minimally invasive. In recent years, there has existed a tendency to combine several biomarkers to a diagnostic profile, in order to improve the diagnostic sensitivity, for example, PCA3 with TMPRSS2: ERG or AMACR [16,17].

In this study, we sought to determine a urine-based diagnostic test for PCa, which used, in the same experimental settings, a combination of the well-known biomarker PCA3 with the novel PCa urinary biomarker prostate-specific G-protein coupled receptor (PSGR). PCA3 is a prostate-specific, non-coding RNA that is highly over-expressed in PCa, as compared to the normal prostate, and that is negative in other normal and malignant tissues. This gene has been regarded as a potential diagnostic tool in the detection of prostate tumor cells in tissue biopsies and bodily fluids [18–21]. PSGR is a member of the G-protein-coupled odorant receptor (OR) family, and it also corresponds to the sub-family Gα12 receptor, the expression of which is restricted to the human prostate [22]. Receptors coupled to heterotrimeric GTP-binding proteins (G-proteins) are integral membrane proteins involved in the transmission of signals from the extracellular environment to the cytoplasm [22–25]. A variety of external stimuli, including neurotransmitters, hormones, phospholipids, photons, odorants, taste ligands, and growth factors can activate specific members of the G protein-coupled receptors (GPCRs) [22]. PCa patients with advanced disease express elevated levels of GPCRs and GPCR ligands, suggesting that the GPCR system may be activated in the cancerous gland and may contribute to tumor growth [26]. PSGR maps to chromosome 11p15 and its cDNA have encoded a 320 aminoacid protein with a predicted molecular mass of 35.4 kDa [24]. PSGR shows a highly prostate tissue-specific expression and tumor associated over-expression that are restricted to prostate epithelial cells [22–25].

## MATERIALS AND METHODS

### Patients and Urine Collection

This study was approved by the institutional review board of the Vall d'Hebron Hospital. All urine samples were obtained from the Department of Urology of the Vall d'Hebron Hospital in Barcelona and were taken from patients subjected to prostate biopsy because of their increased serum PSA levels (>4.0 ng/ml) and/or

an abnormal diagnostic DRE. Patients with other known tumors and/or previous PCa therapies were excluded from the study. Written informed consent was obtained from all patients.

Two hundred sixty-two urine sediments were collected immediately after prostate massage (PM). PM were performed by systematically applying severe pressure to the prostate from the base to apex and from the lateral to the median line of each lobe. The definitive diagnosis of all patients was achieved by prostate biopsy. Biopsies were performed using an end-fire ultrasound transducer (Falcon 2101; B-K Medical, Inc.) and an automatic 18-gauge needle (Bard, Inc.). The minimum number of cores removed in every procedure was 10, and between 1 and 8 additional cores were removed, according to the Vienna nomogram [27]. The study population consisted of 215 men, 73(34%) were positive for PCa and 142 (66%) were negative (benign controls without cancer; Table I). The clinical and pathological information for these 215 patients are shown in Supplemental Table I.

### Sample Preparation

Urine samples (30–50 ml first catch) were collected after PM. Urine was collected in urine collection cups, kept on ice, transported to the lab, and processed within 30 min. The urine samples were centrifuged at 2,500 *g* for 10 min at 4°C, and then the pellets were washed twice with cold PBS (1×). Finally, the pellets were stored with 1:5 RNA Later (Ambion) at 80°C until RNA extraction.

**RNA Isolation and Preamplification.** Urine RNA was extracted with the QIAamp® Viral RNA Mini Kit (Qiagen). Single-stranded cDNA synthesis was carried out using the SuperScript III reverse transcriptase (Invitrogen) and incubated at 37°C for 20 min with RNaseH (Invitrogen). cDNA was stored at −20°C until preamplification with the TaqMan Preamp Master Mix Kit (Applied Biosystems, Foster City, CA).

**TABLE I. Clinical and Pathological Information of the Patients Used in This Study**

| Clinical information | All patients average (min, max) |
|---|---|
| No. of patients | 215 |
| Age (year) | 65.7 (44–85) |
| PSA level (ng/ml) | 11.86 (1.5–189) |
| Prostate volume (ml) | 48.52 (10.0–162.0) |
| PSA density (PSAD) | 0.3 (0.04–5.37) |
| Ratio free PSA/total PSA | 0.17 (0.00–0.26) |
| Positive DRE | Controls 30/142 (42.6%)<br>Cancers 27/73 (19.7%) |
| Prostate cancer (prevalence) | 73 (34%) |
| GLEASON < 7 | 18 (24.7%) |
| GLEASON = 7 | 42 (57.5%) |
| GLEASON > 7 | 13 (17.8%) |

Controls without prostate cancer (prevalence) 142 (66%).

### Quantitative PCR Analysis

Quantitative real-time PCR (qPCR) was used to detect the prostate derived transcripts PSGR, PCA3, and the control transcript PSA, all using the TaqMan® gene expression assay (Applied Biosystems). Reactions were carried out in triplicate on an ABI-Prism-7900 qPCR machine, and only those results with a standard deviation value of <0.38 were accepted (as recommended by the manufacturer). Threshold levels were set into the exponential phase of the qPCR. The data analysis was carried out using the ABI-Prism-7900SDS Software V2.3 (Applied Biosystems) with the same baseline and threshold set for each plate to generate threshold cycle ($C_t$) values for all the genes in each sample. We used PSA as a prostate housekeeping gene for normalization [19]. We established a qPCR cut-off point of $C_t$ (PSA) < 30 after preamplification to decide whether a specimen was informative and the amount of RNA present sufficient to yield an accurate result. We selected this cut-off point since higher $C_t$ values would correspond to approximately <10 prostatic cells per reaction. For each marker a score was calculated as $C_t$ (marker)/$C_t$ (PSA) × 1,000 [28–30]. A cut-off point for each marker was determined, and a specimen was counted positive when the marker was over the cut-off point.

### Statistical Analysis

In the cases of PCa and in the negative biopsy individuals, the characterization of candidate biomarkers was accomplished by comparing their mean values (univariate analysis). The *t*-test with Welch correction was used when the distribution of the data was normal. The *t*-test was carried out on log-transformed data, which were applied to stabilize the variances. ANOVA analysis was used to compare more than two groups of variables. Receiver operating characteristic (ROC) curves and the area under the curve (AUC) were used to evaluate the performance of each marker score as a measure to discriminate between patients in the PCa group and the others. ROC curves were calculated individually for each marker and for combinations of both markers, as well as for PSA, multivariate ROC (MultiROC). In order to combine more than one marker (k), the following procedure was followed: first, a k-component detection threshold was used, one for each biomarker; then, the new marker was declared positive if at least one of the

scores was above its detection threshold. Sensitivity and specificity values were calculated over the range of the k-component thresholds. This generated a cloud of points. Optimal ROC curve points for the new marker were obtained in the following way: for a fixed sensitivity value, the maximum value for specificity was selected from among the range of specificities in the cloud of points that matched that sensitivity value [31]. The Area under the MultiROC curve (AUCm) was measured to compare the outcome prediction performance between the paired combination and the individual markers. In order to determine if the combined model could predict the outcome significantly better than the individual biomarkers at a fixed sensitivity, we compared the corresponding false positive rates using a z-test for proportions. The computation of the variances of the false positive rates for the z-test was carried out as explained in Krzanowski et al. [32] for PSGR and PCA3 alone. We used a resampling approach [33] in the case of the combination of PSGR and PCA3 (PSGRvPCA3). *P*-values in these comparisons were corrected for multiple testing problems following the false discovery rate procedure [34]. Statistical analyses were performed using the PASW V17.0, Graphpad V4.0 and R.

## RESULTS

For the subject group studied, 215 of the 262 initial specimens yielded sufficient prostate derived RNA for analysis, corresponding to an informative specimen rate of 82.1% (benign 76.7% and PCa 23.3%).

PCA3 and PSGR were tested by univariate analysis to see whether they could differentiate between the patients with PCa and the patients with negative needle biopsies. We found that PSGR ($P = 0.008$) and PCA3 ($P < 0.001$) were significant predictors for PCa (Fig. 1).

Moreover, we tested the ability of PSGR and PCA3 to predict clinical risk groups based on patient variables. The clinical risk groups were determined by clinical patient data that directed the decision to pursue biopsy, to determine treatment, or to stratify patients for surveillance regimens. ANOVA analysis of PSGR and PCA3 did not show any association between the following clinical risk groups: positive versus negative digital rectal exam ($P = 0.296$ and $P = 0.502$, respectively), the GLEASON grade <7, 7, and >7 ($P = 0.838$ and $P = 0.362$, respectively) and the pathological stage 2a, 2b–c, and 3 ($P = 0.304$ and $P = 0.138$, respectively; Supplemental Fig. 1).

To visualize the diagnostic efficacy and to summarize the data of the gene-based qPCR assay of urine samples, a ROC curve analysis was performed for each marker. The ROC curve was used to determine the sensitivity and specificity at different score cut-offs (Fig. 2). ROC curves (PSGR, PCA3, and PSA) and MultiROC (PSGRvPCA3) analysis were used to assess the outcome predictive values of the individual and the paired biomarkers. We obtained the following AUC values: PSGR: 0.68 [0.61–0.76], PCA3: 0.66 [0.58–0.73], PSA: 0.602 [0.52–0.68], and PSGRvPCA3: 0.73 [0.65–0.80]. Comparisons were made between them with the following results: (PSGRvPCA3) versus PSGR $P = 0.124$, (PSGRvPCA3) versus PCA3 $P = 0.049$, and PCA3 versus PSGR $P = 0.663$.

By maximizing the sum of sensitivity and specificity for PSGR, the assay sensitivity was 0.59 [0.47–0.70] and the specificity 0.73 [0.65–0.80]. For PCA3 the sensitivity was 0.69 [0.56–0.79] and the specificity 0.59 [0.50–0.67]. For the combined marker model (PSGRvPCA3), we achieved a sensitivity of 0.77 [0.65–0.85] and a specificity of 0.60 [0.51–0.68]. The positive predictive value for the combined model was (+PV) = 0.50 and the negative predictive value (−PV) = 0.83.

When we compared PSGR, PCA3, and PSGRvPCA3 by fixing the sensitivity to the clinically interesting value of 95% [0.86–0.98], we obtained specificities of 0.15 [0.10–0.22] (PSGR), 0.17 [0.11–0.24] (PCA3), and

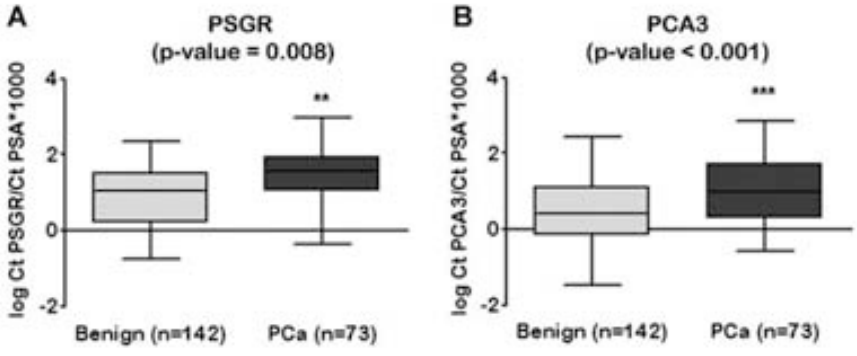



**Fig. 1.** Characterization of the urine based biomarker of prostate cancer (box and-whisker plots). **A**: Relative level of PSGR. **B**: Relative level of PCA3.

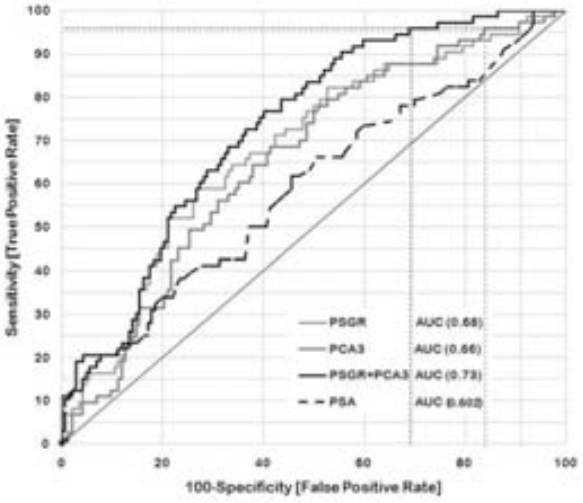



**Fig. 2.** Receiver-operating characteristic (ROC) curves for PSGR and PCA3 in men with detected prostate cancer and men with no prostate cancer detected at biopsy (n = 215). ROC curve for PSGR (light grey line), ROC curve for PCA3 (dark grey line), MultiROC curve for PSGR in combination with PCA3 (black line), and ROC curve for PSA (dashed black line).

0.34 [0.26–0.42] (PSGRvPCA3). Finally, a standard z-test, in a range of sensitivities (90–97.5%), was performed to check if these differences were significant. For the sensitivity of 95%, we obtained the following *P*-values: PSGRvPCA3 compared to PSGR; $P = 0.049$, PSGRvPCA3 compared to PCA; $P = 0.044$, and PSGR compared to PCA3; $P = 1$ (Table II).

## DISCUSSION

PCa is detected in only about one-third of all men biopsied [35]. Methods to enhance PSA specificity have assisted clinicians in deciding which patients should undergo biopsy; however, they have not necessarily improved diagnostic accuracy or facilitated optimal therapeutic decision-making. More accurate tests that can stratify patients according to their risk of developing PCa and identify those who require repeat prostate biopsy are needed [28]. Therefore, PCa screening would benefit from additional biomarkers of higher specificity that could be used in conjunction with PSA testing to better identify biopsy candidates. The use of proximal fluids, such as post-PM urine samples, could represent the best compromise between a minimally invasive technique and the possibility of obtaining enough cells for a correct diagnosis. The reasoning behind this approach is that a manipulation of the prostate will transport cancer cells via the prostatic ductal system into the urethra. Since a PM is part of the standard diagnostic tripod in PCa, urine specimens can be easily obtained after the routine examination process. The first portion of voided urine following PM contains the highest concentration of prostatic and urethral secretions. In urinary sediments not only PCa cells, but also non-neoplastic prostate and urothelial cells are present. To normalize our test for the number of prostate cells, PSA was used as a "housekeeping" gene [19]. PSA transcripts were quantitatively determined to correct for the number of prostate cells present in urine sediments. In order to improve the sensitivity of the PCA3 test [19–21], we combined it with PSGR, a prostate-specific integral membrane receptor that is highly over-expressed in PCa cells [23,36]. The relative quantification of PSGR was achieved by the use of the ratio $C_t$ (Gene)/$C_t$ (PSA) × 1,000 analogous to PCA3 as the diagnostic indicator [28–30]. Our data showed that PSGR is a useful biomarker in urine for the detection of PCa when compared to PCA3 as an alternative urine-based detection marker, which has previously been evaluated in similar groups of patients. In our study PSGR demonstrated a similar performance (AUC 0.681), as opposed to PCA3 (AUC 0.66–0.72) [29,30,37]. As these studies used different methodologies, such as target capture with magnetic particles, transcription-mediated amplification, and a hybridization protection assay to detect PCA3 transcripts in urine and/or used different disease prevalence, the direct comparison of these results and the AUCs would have been inappropriate. However, the determination

**TABLE II. Comparison of PSGR and PCA3 by z-Test for Proportions**

| | Specificity | | | *P*-value | | |
|---|---|---|---|---|---|---|
| | | | | PSGR vs. | PCA3 vs. | PSGR vs. |
| Sensitivity | PSGR | PCA3 | PSGRvPCA3 | 2M[a] | 2M[a] | PCA3 |
| 0.975 | 0.092 | 0.092 | 0.254 | 0.049 | 0.044 | 1 |
| 0.945 | 0.148 | 0.169 | 0.338 | 0.049 | 0.044 | 1 |
| 0.904 | 0.239 | 0.254 | 0.444 | 0.043 | 0.044 | 1 |

[a]2M: PSGR in combination with PCA3 (PSGRvPCA3).

of the cut-off score, PCA3 transcripts/PSA transcripts in urine, was identical. The main reasons for not using the commercial PCa urinary test system APTIMA® assay (Gen-Probe, San Diego, CA; PROGENSA™ for European countries) in this study were the high costs of the tests (between 250 and 300 per test) and the fact that this assay uses a closed system, which does not allow the user to test in parallel additional biomarkers. In contrast, the methodology used in our study allowed the combination of additional biomarkers to increase the sensibility and specificity for PCa detection.

We have found that PSGR could serve as an additional marker to improve PCa detection. By ROC curve analysis we were able to show that PSGR could present an alternative or complement to PCA3 testing. ROC curve analysis is a useful tool in the assessment of biomarker accuracy in two situations: acknowledging the strengths and the weaknesses of the method [38]. Instead of the traditional binary logistic regression approach, we used MultiROC curves to combine PSGR with PCA3. This analysis retains all the simplicity of interpretations using the traditional ROC curve analysis, but it additionally allows comparisons between the performance of multivariate combinations without being restricted to the display of a single variable's performance and comparisons of individual tests [39]. This kind of analysis has the advantage that the combinations of sensitivity and specificity are not fixed in advance.

In the early detection of disease, a combination of biomarkers promises improved discrimination over those diagnostic tests based on a single marker [40]. Because of this, in recent years, several studies followed the same approach of combining various biomarkers for the detection of PCa in urine. The principle that underlies the combined biomarker approach is consistent with tests offered for the detection of PCa in tissue specimens [41,42]. Ouyang et al. [16] developed a duplex quantitative polymerase chain reaction assay for the detection of PCa, based on the quantification of AMACR and PCA3 in urine sediments. For PCA3, they achieved a sensitivity of 72% and a specificity of 53%. Another important study came from Laxman et al., who combined PCA3 with GOLPH2, SPINK1, and TMPRSS2:ERG fusion status. They found a sensitivity of 62% and a specificity of 75% with an AUC of 0.76 for the combined model and for PCA3, 75% and 56%, respectively [17]. Vener et al. [13] used a combination of three methylation markers (GSTP1, RARB, and APC). They found an AUC of 0.69 with a sensitivity of 55% and a specificity of 80%. In the mid sensitivity range, these results were similar to our findings (sensitivity 77%, specificity 60%, AUC 0.73). In order to translate these findings to the clinics, it would be necessary to consider the relative importance of false negatives and false positives. If we want to use a test to rule in PCa patients, a false negative is worse than a false positive. Consequently, we propose to use a different combination of sensitivity and specificity, since in a diagnostic test for PCa the consequences of missing a case patient with cancer may be fatal [43]. With a sensitivity of 95%, Laxman et al. found a specificity of approximately 23%, while the specificity of Vener et al. dropped to >10%. Similarly, the specificity of PCA3 alone went down. Using the data of Marks et al. [29] a sensitivity of 95% would correspond to a specificity of 14%. These data were comparable to our findings. By fixing the sensitivity at 95%, we obtained specificities of 15% (PSGR) and 17% (PCA3), while in contrast to the other studies we found a specificity of 34% for the combination of PSGR and PCA3 (PSGRvPCA3) (Fig. 2). Therefore, the combination of both biomarkers could improve clinical decision making, whether or not to perform a biopsy, especially in a range of high sensitivities (≥95%).

Finally, we analyzed PSGR and PCA3 in order to test their ability to predict clinical risk groups based on patient variables. However, we could not observe any relationship between the clinical groups and PSGR or PCA3. These data confirm the findings of Hessels et al. [44] who could not find a correlation between clinical stage, biopsy Gleason score, radical prostatectomy Gleason score, tumor volume, and pathological stage and PCA3 in post-DRE urinary sediments in a study of 351 patients.

Another interesting benefit would be the improved ability to decide whether or not to perform a re-biopsy. In this case, sensitivity could be lower in order to get a higher specificity, as has been observed with the common cut-off value of 35 in the PCA3 test. In guiding repeat biopsy decisions, PCA3 alone has shown its clinical utility in men with ≥1 negative biopsy [37,41]. This is achieved with the common cut-off value of 35 in the PCA3 test. However, this was not an objective of this article. Therefore, in the future it would be interesting to analyze PSGR in patients with repeat biopsies and patients who have shown recurrence after prostatectomy, in order to establish specific thresholds or to generate predictive nomograms.

In summary, we have demonstrated a qPCR assay on urine sediment from patients presenting for prostate biopsy based on PSGR and PCA3 screening. Future studies will be conducted to further improve the performance of this test, maybe through the evaluation of additional markers. The improved performance of this test will fill important medical needs, such as helping to improve decision making for repeat biopsies in men who have elevated PSA levels and negative biopsies, resolving the so-called "PSA dilemma" and ruling out clinically significant PCa.

Thus, through the use of this test, risk stratification and patient counseling will be improved and the number of patients referred for unnecessary biopsies will be reduced.

**Human Genes**:

| Gene symbol | HUGO Gene Nomenclature Committee |
|---|---|
| PCA3 | PCA3 |
| PSGR | OR51E2 |
| PSA | KLK3 |

## ACKNOWLEDGMENTS

We want to thank Mª Dolors Agrás and Luis Ortega for their enduring helpfulness, flexibility, and disposition in the collection of the huge number of urine samples used in this study. The authors would also like to thank L. Piccione for correction of the manuscript. Marina Rigau is the recipient of a predoctoral fellowship from the Fundación Francisco Cobos. Marina Rigau is the recipient of a predoctoral fellowship from the Fundación Francisco Cobos (2008) and from the Catalan Government (2009-TEM 00066).

## REFERENCES

1. Jemal A, Siegel R, Ward E, Hao Y, Xu J, Thun MJ. Cancer statistics, 2009. CA Cancer J Clin 2009.
2. Pannek J, Partin AW. Prostate-specific antigen: What's new in 1997. Oncology (Williston Park) 1997;11(9):1273–1278, discussion 1279–1282.
3. Loeb S, Gashti SN, Catalona WJ. Exclusion of inflammation in the differential diagnosis of an elevated prostate-specific antigen (PSA). Urol Oncol 2009;27(1):64–66.
4. Catalona WJ, Smith DS, Ratliff TL, Dodds KM, Coplen DE, Yuan JJ, Petros JA, Andriole GL. Measurement of prostate-specific antigen in serum as a screening test for prostate cancer. N Engl J Med 1991;324(17):1156–1161.
5. Makinen T, Tammela TL, Stenman UH, Maattanen L, Aro J, Juusela H, Martikainen P, Hakama M, Auvinen A. Second round results of the Finnish population-based prostate cancer screening trial. Clin Cancer Res 2004;10(7):2231–2236.
6. Cervera Deval J, Morales Olaya FJ, Jornet Fayos J, Gonzalez Anon M. Diagnostic value of the second prostate biopsies in males of risk. Study stratified by value of PSA. Actas Urol Esp 2004;28(9):666–671.
7. Raber M, Scattoni V, Salonia A, Consonni P, Rigatti P. Repeated ultrasound-guided transrectal prostate biopsy in patients with negative histologic test. Arch Ital Urol Androl 2000;72(4):197–199.
8. Mikolajczyk SD, Song Y, Wong JR, Matson RS, Rittenhouse HG. Are multiple markers the future of prostate cancer diagnostics? Clin Biochem 2004;37(7):519–528.
9. Sokoll LJ, Wang Y, Feng Z, Kagan J, Partin AW, Sanda MG, Thompson IM, Chan DW. [−2]Proenzyme prostate specific antigen for prostate cancer detection: A national cancer institute early detection research network validation study. J Urol 2008; 180(2):539–543, discussion 543.
10. Morote J, Raventos CX, Lorente JA, Lopez-Pacios MA, Encabo G, de Torres I, Andreu J. Comparison of percent free prostate specific antigen and prostate specific antigen density as methods to enhance prostate specific antigen specificity in early prostate cancer detection in men with normal rectal examination and prostate specific antigen between 4.1 and 10 ng/ml. J Urol 1997; 158(2):502–504.
11. Harris R, Lohr KN. Screening for prostate cancer: An update of the evidence for the U.S. Preventive services task force. Ann Intern Med 2002;137(11):917–929.
12. Sreekumar A, Poisson LM, Rajendiran TM, Khan AP, Cao Q, Yu J, Laxman B, Mehra R, Lonigro RJ, Li Y, Nyati MK, Ahsan A, Kalyana-Sundaram S, Han B, Cao X, Byun J, Omenn GS, Ghosh D, Pennathur S, Alexander DC, Berger A, Shuster JR, Wei JT, Varambally S, Beecher C, Chinnaiyan AM. Metabolomic profiles delineate potential role for sarcosine in prostate cancer progression. Nature 2009;457(7231):910–914.
13. Vener T, Derecho C, Baden J, Wang H, Rajpurohit Y, Skelton J, Mehrotra J, Varde S, Chowdary D, Stallings W, Leibovich B, Robin H, Pelzer A, Schafer G, Auprich M, Mannweiler S, Amersdorfer P, Mazumder A. Development of a multiplexed urine assay for prostate cancer diagnosis. Clin Chem 2008;54(5): 874–882.
14. Saramaki OR, Harjula AE, Martikainen PM, Vessella RL, Tammela TL, Visakorpi T. TMPRSS2: ERG fusion identifies a subgroup of prostate cancers with a favorable prognosis. Clin Cancer Res 2008;14(11):3395–3400.
15. Marks LS, Bostwick DG. Prostate cancer specificity of PCA3 gene testing: Examples from clinical practice. Rev Urol 2008;10(3): 175–181.
16. Ouyang B, Bracken B, Burke B, Chung E, Liang J, Ho SM. A duplex quantitative polymerase chain reaction assay based on quantification of alpha-methylacyl-CoA racemase transcripts and prostate cancer antigen 3 in urine sediments improved diagnostic accuracy for prostate cancer. J Urol 2009;181(6):2508–2513, discussion 2513–2504.
17. Laxman B, Morris DS, Yu J, Siddiqui J, Cao J, Mehra R, Lonigro RJ, Tsodikov A, Wei JT, Tomlins SA, Chinnaiyan AM. A first-generation multiplex biomarker analysis of urine for the early detection of prostate cancer. Cancer Res 2008;68(3):645–649.
18. de Kok JB, Verhaegh GW, Roelofs RW, Hessels D, Kiemeney LA, Aalders TW, Swinkels DW, Schalken JA. DD3(PCA3), a very sensitive and specific marker to detect prostate tumors. Cancer Res 2002;62(9):2695–2698.
19. Hessels D, Klein Gunnewiek JM, van Oort I, Karthaus HF, van Leenders GJ, van Balken B, Kiemeney LA, Witjes JA, Schalken JA. DD3(PCA3)-based molecular urine analysis for the diagnosis of prostate cancer. Eur Urol 2003;44(1):8–15, discussion 15–16.
20. Fradet Y, Saad F, Aprikian A, Dessureault J, Elhilali M, Trudel A, Piche L, Chypre C. UPM3, a nex molecular urine test for the detection of prostate cancer. Urology 2004;64:311–315.
21. Tinzl M, Marberger M, Horvath S, Chypre C. DD3PCA3 RNA analysis in urine—A new perspective for detecting prostate cancer. Eur Urol 2004;46(2):182–186, discussion 187.
22. Xia C, Ma W, Wang F, Hua S, Liu M. Identification of a prostate-specific G-protein coupled receptor in prostate cancer. Oncogene 2001;20(41):5903–5907.
23. Xu LL, Sun C, Petrovics G, Makarem M, Furusato B, Zhang W, Sesterhenn IA, McLeod DG, Sun L, Moul JW, Srivastava S. Quantitative expression profile of PSGR in prostate cancer. Prostate Cancer Prostatic Dis 2006;9(1):56–61.
24. Xu LL, Stackhouse BG, Florence K, Zhang W, Shanmugam N, Sesterhenn IA, Zou Z, Srikantan V, Augustus M, Roschke V,

Carter K, McLeod DG, Moul JW, Soppett D, Srivastava S. PSGR a novel prostate-specific gene with homology to a G protein-coupled receptor, is overexpressed in prostate cancer. Cancer Res 2000;60(23):6568–6572.

25. Yuan TT, Toy P, McClary JA, Lin RJ, Miyamoto NG, Kretschmer PJ. Cloning and genetic characterization of an evolutionarily conserved human olfactory receptor that is differentially expressed across species. Gene 2001;278(1–2):41–51.
26. Daaka Y. G proteins in cancer: The prostate cancer paradigm. Sci STKE 2004;216:re2. 2004
27. Remzi M, Fong YK, Dobrovits M, Anagnostou T, Seitz C, Waldert M, Harik M, Marihart S, Marberger M, Djavan B. The Vienna nomogram: Validation of a novel biopsy strategy defining the optimal number of cores based on patient age and total prostate volume. J Urol 2005;174(4 Pt 1):1256–1260, discussion 1260–1251; author reply 1261.
28. Hessels D, Schalken JA. The use of PCA3 in the diagnosis of prostate cancer. Nat Rev Urol 2009;6(5):255–261.
29. Marks LS, Fradet Y, Deras IL, Blase A, Mathis J, Aubin SM, Cancio AT, Desaulniers M, Ellis WJ, Rittenhouse H, Groskopf J. PCA3 molecular urine assay for prostate cancer in men undergoing repeat biopsy. Urology 2007;69(3):532–535.
30. van Gils MP, Hessels D, van Hooij O, Jannink SA, Peelen WP, Hanssen SL, Witjes JA, Cornel EB, Karthaus HF, Smits GA, Dijkman GA, Mulders PF, Schalken JA. The time-resolved fluorescence-based PCA3 test on urinary sediments after digital rectal examination; a Dutch multicenter validation of the diagnostic performance. Clin Cancer Res 2007;13(3):939–943.
31. Baker SG. The central role of receiver operating characteristic (ROC) curves in evaluating tests for the early detection of cancer. J Natl Cancer Inst 2003;95(7):511–515.
32. Krzanowski WJ, Hand DJ. ROC curves for continuous data. J Biopharm Stat 2009;20(2):485–487.
33. Good P. Resampling methods. Boston: Birkhauser 1999.
34. Benjamini Y, Hochberg Y. Controlling the false discovery rate: A practical and powerful approach to multiple testing. J R Stat Soc B 1995;55(No.1):289–300.
35. Djavan B, Remzi M, Schulman CC, Marberger M, Zlotta AR. Repeat prostate biopsy: Who, how and when? A review. Eur Urol 2002;42(2):93–103.
36. Wang J, Weng J, Cai Y, Penland R, Liu M, Ittmann M. The prostate-specific G-protein coupled receptors PSGR and PSGR2 are prostate cancer biomarkers that are complementary to alpha-methylacyl-CoA racemase. Prostate 2006;66(8):847–857.
37. Haese A, de la Taille A, van Poppel H, Marberger M, Stenzl A, Mulders PF, Huland H, Abbou CC, Remzi M, Tinzl M, Feyerabend S, Stillebroer AB, van Gils MP, Schalken JA. Clinical utility of the PCA3 urine assay in European men scheduled for repeat biopsy. Eur Urol 2008;54(5):1081–1088.
38. Soreide K. Receiver-operating characteristic curve analysis in diagnostic, prognostic and predictive biomarker research. J Clin Pathol 2009;62(1):1–5.
39. Shultz EK. Multivariate receiver-operating characteristic curve analysis: Prostate cancer screening as an example. Clin Chem 1995;41(8 Pt 2):1248–1255.
40. Etzioni R, Kooperberg C, Pepe M, Smith R, Gann PH. Combining biomarkers to detect disease with application to prostate cancer. Biostatistics 2003;4(4):523–538.
41. Landers KA, Burger MJ, Tebay MA, Purdie DM, Scells B, Samaratunga H, Lavin MF, Gardiner RA. Use of multiple biomarkers for a molecular diagnosis of prostate cancer. Int J Cancer 2005;114(6):950–956.
42. Schmidt U, Fuessel S, Koch R, Baretton GB, Lohse A, Tomasetti S, Unversucht S, Froehner M, Wirth MP, Meye A. Quantitative multi-gene expression profiling of primary prostate cancer. Prostate 2006;66(14):1521–1534.
43. Pepe MS, Feng Z, Janes H, Bossuyt PM, Potter JD. Pivotal evaluation of the accuracy of a biomarker used for classification or prediction: Standards for study design. J Natl Cancer Inst 2008;100(20):1432–1438.
44. Hessels D, van Gils MP, van Hooij O, Jannink SA, Witjes JA, Verhaegh GW, Schalken JA. Predictive value of PCA3 in urinary sediments in determining clinico-pathological characteristics of prostate cancer. Prostate; 70(1):10–16.

The Prostate

# A Three-Gene Panel on Urine Increases PSA Specificity in the Detection of Prostate Cancer

Marina Rigau,[1,2] Israel Ortega,[3] Maria Carmen Mir,[4] Carlos Ballesteros,[4] Marta Garcia,[1] Marta Llauradó,[1] Eva Colás,[1] Núria Pedrola,[1] Melania Montes,[1] Tamara Sequeiros,[1] Tugce Ertekin,[1] Blanca Majem,[1] Jacques Planas,[4] Anna Ruiz,[1] Miguel Abal,[5] Alex Sánchez,[3,6] Juan Morote,[4,7] Jaume Reventós,[1,7] and Andreas Doll[1]*

[1]*Research Unit in Biomedicine and Translational and Pediatric Oncology, Vall d'Hebron Research Institute, Barcelona, Spain*
[2]*Universitat Autònoma de Barcelona, Barcelona, Spain*
[3]*Statistics and Bioinformatics Unit, Vall d'Hebron Research Institute, Barcelona, Spain*
[4]*Department of Urology, Vall d'Hebron University Hospital, Barcelona, Spain*
[5]*Department of Medical Oncology, Complexo Hospitalario Universitario SERGAS, Santiago de Compostela, Spain*
[6]*Statistics Department, University of Barcelona, Barcelona, Spain*
[7]*School of Medicine, Universitat Autònoma de Barcelona, Barcelona, Spain*

**BACKGROUND.** Several studies have demonstrated the usefulness of monitoring an RNA transcript, such as PCA3, in post-prostate massage (PM) urine for increasing the specificity of prostate-specific antigen (PSA) in the detection of prostate cancer (PCa). However, a single marker may not necessarily reflect the multifactorial nature of PCa.

**METHODS.** We analyzed post-PM urine samples from 154 consecutive patients, who presented for prostate biopsies because of elevated serum PSA (>4 ng/ml) and/or abnormal digital rectal exam. We tested whether the putative PCa biomarkers PSMA, PSGR, and PCA3 could be detected by quantitative real-time PCR in post-PM urine sediment. We combined these findings to test if a combination of these biomarkers could improve the specificity of actual diagnosis. Afterwards, we specifically tested our model for clinical usefulness in the PSA diagnostic "gray zone" (4–10 ng/ml) on a target subset of 82 men with no prior biopsy.

**RESULTS.** By univariate analysis, we found that the PSMA, PSGR, and PCA3 scores were significant predictors of PCa. Using a multiplex model, the area under the multi receiver-operating characteristic curve was 0.74 versus 0.82 in the diagnostic "gray zone." Fixing the sensitivity at 96%, we obtained a specificity of 34% and 50% in the gray zone.

**CONCLUSIONS.** Taken together, these results provide a strategy for the development of a more accurate model for PCa diagnosis. In the future, a multiplexed, urine-based diagnostic test for PCa with a higher specificity, but the same sensitivity as the serum-PSA test, could be used to determine better which patients should undergo biopsy. *Prostate* © 2011 Wiley-Liss, Inc.

*KEY WORDS:* **prostate cancer; urine; biomarkers; PSA; PSMA; PSGR; PCA3**

Additional Supporting Information may be found in the online version of this article.

Grant sponsor: Instituto de Salud Carlos III, Ministerio de Ciéncia e Innovación; Grant number: RTICC RD06/0020/0058, CP10/00355, PS 09-00496, Red de Genómica del Cáncer y Genotipado de tumores C03/10; Grant sponsor: Fundación para la Investigación en Urología; Grant number: FIU-2007; Grant sponsor: Departament d'Universitats, Recerca I Societat de la Informació de la Generalitat de Catalunya; Grant number: SGR00487; Grant sponsor: Marina Rigau is the recipient of a predoctoral fellowship from the Fundación Francisco Cobos; Grant number: Jovenes Licenciados 2008; Grant sponsor: Departament d'Univeristats, Recerca i Societat de la Informació de la Generalitat de Catalunya; Grant number: 2009-TEM 00066.

Juan Morote, Jaume Reventós, and Andreas Doll contributed equally to this study.

*Correspondence to: Dr. Andreas Doll, Research Unit in Biomedicine and Translational and Pediatric Oncology, Research Institute Vall d'Hebron University Hospital, Pg. Vall d'Hebrón 119-129, 08035 Barcelona, Spain. E-mail: adoll@ir.vhebron.net

Received 18 January 2011; Accepted 7 March 2011
DOI 10.1002/pros.21390
Published online in Wiley Online Library (wileyonlinelibrary.com).

## INTRODUCTION

Prostate cancer (PCa) has emerged as the most common neoplasia among American men (28% of all cases), followed by lung cancer (15%) and colorectal and rectum cancer (9%). PCa is the second leading cause of death from cancer in American men (32,050 deaths) after lung cancer (86,200) [1] and the third one in the European Union (89,300 deaths in 2008) after lung cancer (255,300 deaths) and colorectal cancer (110,500 deaths) [2]. When PCa is detected in its initial stages, it is still curable by various therapies, but once disseminated, only palliative approaches are attempted.

The suspicion of PCa is established through elevated serum prostate-specific antigen (PSA) and/or abnormal digital rectal exam (DRE). Diagnosis is confirmed through prostate-needle biopsy (PB). PCa screening is widely utilized and considered to be an effective detection method for PCa. However, due to a substantial overlap in the serum PSA values of men with benign prostatic hyperplasia (BPH) or chronic prostatitis and those with PCa, only about 30–40% of the biopsied men are diagnosed with PCa. As a result of their persistently elevated PSA levels, but negative PB results, these men undergo repeated biopsies. This situation is called the ''PSA dilemma'' [3,4].

As both normal prostate and PCa epithelial cells can be detected in the urine of men, their use in the detection of PCa is very attractive. Urine-based testing for PCA3 expression and GSTP1 methylation has already been documented in large PCa screening programs. However, only a few studies have taken into account the heterogeneity of cancer development based on a diagnostic profile [5–7]. Recently, several studies followed the approach of combining various biomarkers for the detection of PCa in urine. The principle that underlies the combined biomarker approach is consistent with tests offered for the detection of PCa in tissue specimens [8,9] and takes into consideration the heterogeneity of cancer development based on a diagnostic profile. Ouyang et al. [10] developed a duplex quantitative real time PCR (qPCR) assay for the detection of PCa, based on AMACR and PCA3 in urine sediments. Laxman et al. [7] combined PCA3 with GOLPH2, SPINK1, and TMPRSS2:ERG fusion status, and Vener et al. [11] used a combination of three methylation markers (GSTP1, RARB, and APC).

Unfortunately, two important aspects have not been sufficiently addressed in the current studies. First, there needs to be an analysis of patients with prior PB information whose results fall within the serum PSA diagnostic ''gray zone'' (4–10 ng/ml). Second, focus must be placed on the relative importance of false negatives and false positives, since in a diagnostic test for PCa, the consequences of missing a cancer case may be fatal to the patient.

Consequently, the aim of this study was to develop a multiplexed, urine-based diagnostic test for PCa with a higher specificity, but the same sensitivity as the serum PSA test, which could be useful in determining which patients should undergo PB. We chose three genes, known to be over-expressed in PCa. Prostate specific membrane antigen (PSMA) is an integral, non-shed type2 membrane protein that is highly and specifically expressed in prostate epithelial cells and strongly up-regulated in PCa. PSMA is also present in the neovasculature of other solid tumors [12]. Prostate-specific G-protein coupled receptor (PSGR) is a member of the G-protein coupled odorant receptor family, which is highly prostate tissue-specific and presents tumor-associated over-expression in tissue and urine [13–15]. And finally, Prostate cancer antigen 3 (PCA3) is a prostate-specific non-coding RNA, which is highly over-expressed in more than 95% of primary PCa specimens, PCa metastasis, and urine [16,17].

## MATERIALS AND METHODS

### Patients and Urine Collection

This study obtained approval from the institutional review board. Urine samples were collected from 198 consecutive men, who had been referred for PB, immediately post-prostate massage (PM). The indications for PB were an abnormal DRE and/or serum PSA levels higher than 4 ng/ml. Patients with other known tumors and/or previous PCa therapies were excluded from the study. Written informed consent was obtained from all patients. The study population consisted of 154 men, 57 (37%) were positive for PCa and 97 (63%) were negative (benign controls without cancer). The clinical and pathological information data for these 154 patients are shown in Table I.

**PM methodology.** PM was performed by systematically applying severe digital pressure to the prostate from the base to the apex and from the lateral to the median line of each lobe.

**PB methodology.** The biopsies were performed using an end-fire ultrasound transducer (Falcon 2101, B-K Medical, Inc.) and an automatic 18gauge needle (Bard, Inc.). The minimum number of cores removed in every procedure was 10, and between 1 and 8 additional cores were removed, according to the Vienna nomogram [18].

**TABLE I. Clinical and Pathological Information of the Informative Patients**

| | All patients average (min, max, or %) | Patients PSA 4–10 ng/ml and first biopsy average (min, max, or %) |
|---|---|---|
| No of patients | 154 | 82 |
| Age (year) | 65.5 (44–85) | 64.8 (44–85) |
| PSA level (range) | 10.9 (2.5–189) | 6.6 (4–10) |
| Prostate Volume (ml)[a] | 51.1 (10.0–162.0) | 46.2 (16.0–120.0) |
| PSA density (PSAD)[a] | 1.61 (0.29–43.7) | 0.167 (0.06–0.34) |
| Ratio free PSA/total PSA[b] | 0.17 (0.00–0.80) | 0.18 (0.01–0.80) |
| Prostate cancer | 57 (37%) | 28 (34.1%) |
| GLEASON < 7 | 7 (12.3%) | 4 (14.3%) |
| GLEASON = 7 | 41 (71.9%) | 20 (71.4%) |
| GLEASON > 7 | 9 (15.8%) | 4 (14.3%) |
| Benign | 97 (63%) | 54 (65.9%) |

[a]In 7 (5) patients in the prostate volume/PSAD was not determined.
[b]In 4 (2) patients fPSA was not determined.

### Sample Preparation

Urine samples (50 ml first catch) were collected in urine collection cups, kept on ice, transported to the lab and processed within 30 min. The urine samples were centrifuged at 2,500*g* for 10 min at 4°C, and then the pellets were washed twice with cold 1× PBS. Finally, the pellets were stored with 1:5-RNA Later (Ambion, Foster City, CA) at −80°C until RNA extraction. Urine RNA was extracted with the QIAamp® Viral RNA Mini Kit (Qiagen, Hilden, Germany). Single-stranded cDNA synthesis was carried out using the SuperScript III Reverse Transcriptase (Invitrogen, Carlsbad, CA) and stored at −80°C until pre-amplification with the TaqMan Preamp Master Mix Kit (Applied Biosystems, Foster City, CA).

### Quantitative PCR Analysis

We selected three putative PCa urinary biomarkers, PSMA, PSGR, PCA3 and the control transcript, PSA, all from the TaqMan® Gene Expression Assay (Applied Biosystems) to be analyzed by qPCR. Reactions were carried out in triplicate on an ABI-Prism-7900 qPCR machine, and only those results with a standard deviation <0.38 value were accepted (as recommended by the manufacturer). Threshold levels were set into the exponential phase of the qPCR. The data analysis was carried out using the ABI-Prism-7900SDS Software V2.3 (Applied Biosystems) with the same baseline and threshold set for each plate to generate threshold cycle (Ct) values for all the genes in each sample. To exclude the possibility that these markers might also be expressed in non-cancer cells normally found in urine sediments, such as cells from the urothelium, kidney, bladder, or blood, the content in the clinical specimens was normalized to the amount of prostate derived RNA. Since only a relatively small number of prostate cells are to be found in urine, we performed a cDNA pre-amplification step before the qPCR. We established a qPCR cut-off of Ct of PSA ≤30 after pre-amplification, in order to determine whether a specimen was informative, that is, the amount of RNA present was sufficient to yield an accurate result. For each marker, a score was calculated as Ct(marker)/Ct(PSA) × 1,000 [19–21]. A cut-off value for each marker was determined, and a specimen was counted positive when the marker was over the cut-off value.

### Statistical Analysis

The characterization of candidate biomarkers was accomplished by comparing the mean values (univariate analysis) between the different groups analyzed. The *t*-test was used when the data were normally distributed; otherwise, a Mann–Whitney test was used. Normality was verified with the Shapiro–Wilk normality test. ANOVA analysis was used to compare more than two groups of variables. All tests were carried out on log-transformed data.

Receiver-operating characteristic (ROC) curves and the area under the curve (AUC) were used to evaluate the performance of each marker score as a way to discriminate between patients in the PCa group and the others. ROC curves were calculated individually for each marker and multivariate ROC (MultiROC) for combinations of both markers, as well as for PSA, as we described before [15]. Briefly, a detection threshold was used, one for each biomarker; then, the new marker was declared positive if at least one of the scores was above its detection threshold. Sensitivity and specificity values were calculated over the

range of thresholds. MultiROC curve points for the new marker were obtained in the following way: for a fixed sensitivity value, the maximum value for specificity was selected from the range of specificities in the cloud of points that matched the sensitivity value. This approach has already been used when combining biomarkers [22,23].

The area under the MultiROC curve (AUCm) was measured to compare the outcome prediction performance between the combined and the individual markers. In order to determine if the combined model could predict the outcome significantly better than the individual biomarkers at a fixed sensitivity, we compared the corresponding false positive rates using a Z-test for proportions. The computation of the variances of the false positive rates for the Z-test was carried out as explained in Krzanowski et al. [24] for PSMA, PSGR, and PCA3 alone. We used a resampling approach in the case of the combination of PSMA, PSGR, PCA3, and PSA ((PSMA ∨ PSGR ∨ PCA3) ∨ PSA). *P*-values in these comparisons were corrected for multiple testing problems following the false discovery rate procedure [25].

A major concern in creating a ROC curve to represent the performance of a combination of biomarkers is over-fitting. To control this bias, we used an approach based on the split-half (50%) method to validate the results [26]. Briefly, 50% of the sample was used as a training set to compute a ROC curve. This ROC curve was used to determine the cut-off values needed to obtain a defined set of sensitivities (90–100%). Thus, with the determined cut-offs, the test sample was used to estimate the specificity corresponding to each sensitivity. This procedure was repeated 500 times and final estimations were obtained by computing the mean values produced after all the repetitions. Statistical analyses were performed using the PASW V17.0 and the free statistical language R and the ROC-package from the Bioconductor project (http://www.bioconductor.org).

## RESULTS

### Results Obtained in the Overall Group

The PCa detection rate by prostate biopsy was 37% (57/154). For the total subject group studied, 154 of the 198 specimens yielded sufficient RNA for analysis, corresponding to an informative specimen rate of 78.8% (95.0% for patients with PCa and 70.3% for benign patients).

PSMA, PSGR, PCA3 and additional clinical parameters were tested on the post-PM urine sediments obtained from 154 patients, in order to determine whether they could differentiate patients with PCa from patients with negative needle biopsies. PSMA ($P = 0.016$), PSGR ($P < 0.001$), PCA3 ($P = 0.018$), and PSAD ($P = 0.027$) were significant predictors for PCa (Fig. 1), while PSA ($P = 0.55$), fPSA ($P = 0.60$), and prostate volume ($P = 0.053$) did not

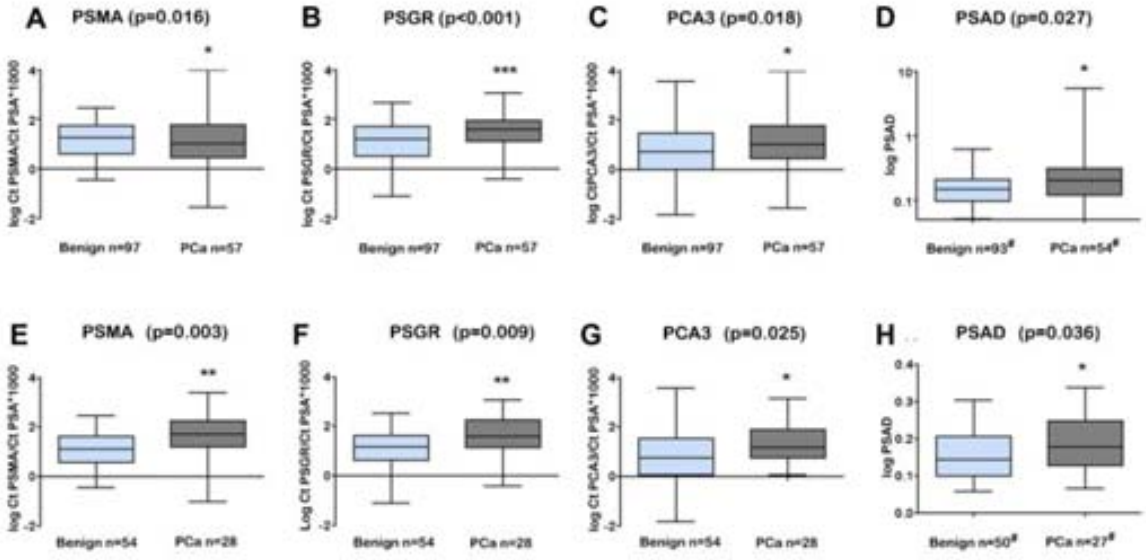



**Fig. 1.** Characterization of urine based biomarkers for prostate cancer (box-and-whisker plots). Relative level of PSMA (**A**), PSGR (**B**), PCA3 (**C**), and PSAD (**D**) in men with PCa versus benign (Overall group). Relative level of PSMA (**E**), PSGR (**F**), PCA3 (**G**), and PSAD (**H**) in men within PSA 4–10 ng/ml and first biopsy, with PCa versus benign (Clinical interest group). # Patients with PSAD determination.

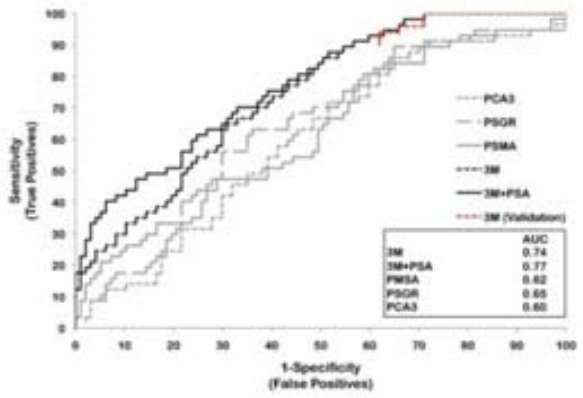



**Fig. 2.** ROC curves comparing individual markers, 3M (PSMA ∨ PSGR ∨ PCA3) and 3M with PSA (3M ∨ PSA) in men with detected prostate cancer and men with no detected cancer at biopsy. ROC-curve for PSMA (gray line), ROC-curve for PSGR (dashed gray line) and ROC-curve for PCA3 (gray dotted line), 3M ∨ PSA -ROC (black line), 3M-ROC (black dotted line), and ROC curve for split-half validation (red line).

show statistical significance (Supplemental Fig. 1). After multivariate analysis we chose PSMA, PSGR, and PCA3 for ROC-curve analysis to visualize the diagnostic efficacy and to summarize the data of the gene-based qPCR assay on the urine samples. We obtained the following AUC values [95% CI]: PSMA 0.62 [0.53–0.72], PSGR 0.65 [0.56–0.73], PCA3 0.60 [0.51–0.69] (Fig. 2).

To determine if the combination of various biomarkers could improve performance over single biomarkers, we combined PSMA, PSGR, and PCA3 (3M) using a MultiROC approach (PSMA ∨ PSGR ∨ PCA3) PCA3) and a traditional binary logistic regression analysis (Fig. 2). The performance of the MultiROC model was superior to the logistic regression analysis (Supplementary Fig. 2). We then combined the 3 Marker model (3M) with serum PSA to a new combined marker to evaluate if including these parameters could improve the prediction of PCa. The area under the MultiROC for 3M (PSMA ∨ PSGR ∨ PCA3) was 0.74 [0.68–0.80] and for 3M ∨ PSA 0.77 [0.71–0.83] (Fig. 2). Comparisons were made between them with the following results: 3M versus PCA3 $P = 0.0071$, 3M versus PSGR $P = 0.0493$, 3M versus PSMA $P = 0.0253$, PSMA versus PSGR $P = 0.75$, PSMA versus PCA3 $P = 0.69$, and PSGR versus PCA3 $P = 0.46$ (Table II).

By maximizing the sum of sensitivity and specificity for PSMA, the assay sensitivity was 81% [68–90] and the specificity 41% [31–52]. For PSGR, the assay sensitivity was 63% [49–77] and the specificity 64% [54–73], and for PCA3 the sensitivity was 86% [74–94]

**TABLE II. Comparison of PSMA, PSGR, PCA3, and 3M (PSMA ∨ PSGR ∨ PCA3) by a Z-Test for proportions**

| | Overall group | Clinical interest group (PSA 4–10 ng/ml first biopsy) |
|---|---|---|
| PSGR vs. PCA3 | 0.4623 | 0.5407 |
| PSGR vs. PSMA | 0.7500 | 0.3601 |
| PSMA vs. PCA3 | 0.6885 | 0.1214 |
| PCA3 vs. 3M | 0.0071 | 0.0024 |
| PSGR vs. 3M | 0.0493 | 0.0199 |
| PSMA vs. 3M | 0.0253 | 0.1542 |

and the specificity 35% [27–45]. For 3M, we achieved a sensitivity of 89% [79–95] and a specificity of 45% [35–57]. The statistical comparison of the ROC-curves is shown in Table II. When we compared the PSMA, PSGR, PCA3, and 3M by fixing the sensitivity to the clinically interesting values range of 90–100% while the specificity for the individual markers drops to 0% at 100% sensitivity the combined models maintained a specificity of 29% [21–39] (Fig. 2). A standard Z-test, was performed to check if these differences were significant. For example, for a sensitivity of 96%, we obtained the following *P*-values: 3M compared to PSMA; $P = 0.0001$, 3M compared to PSGR; $P = 0.0002$, 3M compared to PCA3; $P = 0.0006$ (Supplemental Table Ia). In order to validate this data we performed a validation by split-half method with 500 iterations. We constructed a validation ROC-curve in the range of sensitivities from 90% to 100% (Table III).

### Results Obtained in the Group of Clinical Interest (Gray Zone)

In order to decide, whether a biopsy should be performed or not, a subgroup of special clinical interest was examined. This group was comprised of 82 patients with no prior PB, whose results were in the PSA diagnostic "gray zone" between 4 and 10 ng/ml. The PCa detection rate by PB was 34% (28/82) (Table I). For the total subject group studied, 82 of the 93 specimens yielded sufficient RNA for analysis, corresponding to an informative specimen rate of 88.2%.

All markers were tested by univariate analysis to see if they could differentiate patients with PCa from patients with negative PB. The following results were obtained: PSMA ($P = 0.003$), PSGR ($P = 0.009$), PCA3 ($P = 0.025$), PSAD ($P = 0.036$) (Fig. 1), and prostate volume ($P = 0.02$) were statistically significant, while PSA ($P = 0.26$) and fPSA ($P = 0.67$), did not show any statistical significance (Supplemental Fig. 1). We obtained the following AUC [95% CI] values: PSMA

**TABLE III. Validation by Split-Half Method With 500 Iterations**

| | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 3M (Overall group) | | | | | | | | | | | |
| Sensitivity (%) | 90 | 91 | 92 | 93 | 94 | 95 | 96 | 97 | 98 | 99 | 100 |
| Specificity (%) | 38 | 38 | 38 | 38 | 34 | 34 | 34 | 29 | 29 | 29 | 29 |
| 3M (Clinical interest group; PSA 4–10 ng/ml, first biopsy) | | | | | | | | | | | |
| Sensitivity (%) | 90 | 91 | 92 | 93 | 94 | 95 | 96 | 97 | 98 | 99 | 100 |
| Specificity (%) | 49 | 49 | 49 | 45 | 45 | 45 | 45 | 45 | 45 | 45 | 45 |

0.74 [0.63–0.86], PSGR 0.66 [0.54–0.79], and PCA3 0.61 [0.48–0.74]. The AUC for 3M was 0.82 [0.77–0.86] (Fig. 3). The statistical comparison of the ROC-curves is shown in Table II. By maximizing the sum of sensitivity and specificity for PSMA, the sensitivity was 64% [44–81] and the specificity 70% [56–82]. For PSGR the sensitivity was 61% [41–79] and the specificity 70% [56–82], and for PCA3 the assay sensitivity was 71% [51–87] and the specificity 54% [40–67]. For 3M, we achieved a sensitivity of 89% [72–96] and a specificity of 57% [44–70]. Then, we compared PSMA, PSGR, PCA3, and 3M by fixing the sensitivity to the clinically interesting values range of 90–100%, while the specificity for the individual markers drops to 0% at 100% sensitivity the combined model maintained a specificity of 46% [37–63]. We obtained a specificity of 4% [1–13] for PSMA, 12% [6–24] for PSGR and 0% [0–7] for PCA3 (Fig. 3). A standard test for two proportions was performed in a range of sensitivities (90–100%) to check if these differences were significant (Supplemental Table Ib). Validation ROC-curve in the range of sensitivities from 90% to 100% is shown in Table III. For example, for a sensitivity of 96%, we obtained the following *P*-values: 3M compared to PSMA, $P = 0.0005$, 3M compared to PSGR, $P < 0.0001$, 3M compared to PCA3, $P = 0.0001$ (Supplemental Table Ib).

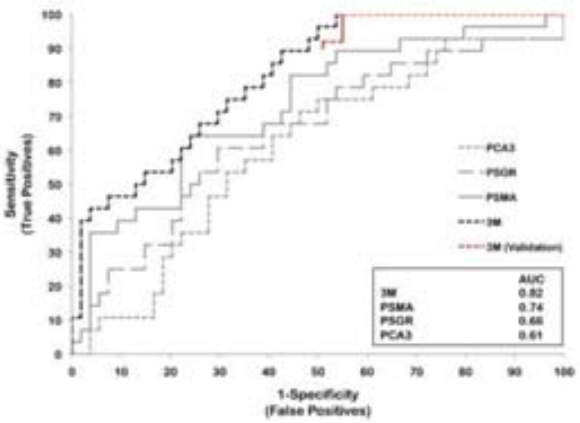



**Fig. 3.** Receiver-operating-characteristic (ROC) curves comparing individual markers with 3M, in men with PSA 4–10 ng/ml and no previous biopsy. ROC-curve for PSMA (gray line), ROC-curve for PSGR (gray dashed line) and ROC-curve for PCA3 (gray dotted line), and 3M-ROC (black line). ROC curve for split-half validation (red line).

Finally, we calculated the number of biopsies that could be avoided by using the combined marker model (3M) on the overall group and on the PSA 4–10 ng/ml group as the % saved biopsies = true negatives + false negatives. For a sensitivity of 96%, 23% (with 3M ∨ PSA 23%) of the biopsies could be avoided in the overall group, including the complete range of PSA and repeated biopsies. The positive (PPV) and negative predictive values (NPV) in this case were 46% and 94%, respectively. In the analysis of the clinical risk group of special interest with no prior biopsy and results in the PSA "gray zone" between 4 and 10ng/ml, 34% of the biopsies could be avoided with a sensitivity of 96%. The PPV and NPV were 51% and 96%, respectively.

## DISCUSSION

Screening by serum PSA is widely utilized and considered to be an effective detection method for PCa, but only about 30–40% of the biopsied men are diagnosed with PCa [27]. As a result of their persistently elevated PSA levels, but negative biopsy results, these men undergo repeated PB. This situation is called the "PSA dilemma." Various attempts have been made to overcome the limitations of PSA screening; nevertheless, there has been no evidence to suggest that any of these methods improve health outcomes. Furthermore, current evidence suggests that the PSA cut-off level should be reduced and that the number of punctures in the PB should not be fewer than 10 [18]. With these standards, the specificity of PSA will be reduced in the future and the number of negative biopsies will increase, as well as the more frequent diagnosis of clinically insignificant carcinomas. Taking all these factors into consideration, there is obviously a need for an additional diagnostic test that better identifies the patients who will indeed have a positive biopsy [19]. This would help to reduce the number of initial and repeat PB. For this reason, it is a priority to search for markers specifically linked to PCa that make its diagnosis more

effective. The use of proximal fluids, such as urine samples post-PM, could be the best compromise between a minimally invasive technique and the possibility of obtaining sufficient cells for a correct diagnosis. Therefore, in this study, we developed a multiplexed qPCR-based test for PCa on post-PM urinary sediments. Because the first portion of a voided urine sample contains the highest concentration of prostatic and urethral secretions [28], post-PM voided urine sample collection was selected. In addition, this type of sample is a more readily accepted specimen for men to provide, rather than ejaculate or expressed prostatic secretions. Urine-based assays would have a great advantage, since the cells from multiple cancerous foci may be released and collected from the whole prostate [29]. As PSA is not expressed in other cells present in urine, such as transitional epithelium, blood cells, or kidney epithelium, it can be used as a prostate gland housekeeping gene [30]. PSMA, for example, is expressed predominantly in the prostate, but it has also been detected in other tissues such as the kidney, the proximal small intestine, and the salivary glands [31]. The relative quantification of a prostate cancer biomarker, can be achieved by the use of the ratio Ct(Gene)/Ct(PSA) × 1,000 [19–21]. The PSA mRNA yield can also be used to decide whether a specimen is informative, that is, if the amount of RNA present is sufficient to yield an accurate result. Interestingly, in our study this yield was much higher in the PCa patients (95%) than in their benign counterparts (70%). This may reflect changes in focal contacts, a loss of cell–cell contacts, or enhanced cell motility in PCa cells.

A wide range of promising PCa biomarkers that are not only prostate-specific, but are also overexpressed in prostate tumors have been identified, including CpG hypermethylation of GSTP1, the TMPRSS2:ERG gene fusions and RNA biomarkers PCA3 [11,15,32,33] and PSGR [15]. Although, urine-based testing for PCA3 expression and GSTP1 methylation has already been documented in large screening programs, there are only a few studies that take into account the heterogeneity of cancer development based on a diagnostic profile. A combination of various biomarkers should clearly improve performance over single biomarkers [22], as a single marker may not necessarily reflect the multifactorial nature of PCa.

We have demonstrated here that PSMA, a marker previously described to be over-expressed in PCa, has a similar performance (AUC 0.62) to the known urine markers PCA3 (AUC 0.66–0.72), and PSGR (0.68) [15,20,21,34]. In order to improve the diagnostic efficacy of these biomarkers one by one, we combined these three biomarkers (PSMA ∨ PCA3 ∨ PSGR).

Instead of the traditional binary logistic regression approach, we used MultiROC curves. This analysis retains all the simplicity of interpretations using the traditional ROC curve analysis, but it additionally allows comparisons between the performance of multivariate combinations without being restricted to the display of a single variable's performance and comparisons of individual tests [35]. This kind of analysis has the advantage that the combinations of sensitivity and specificity are not fixed in advance. We found that the AUCm (0.74) was notably improved, compared to the logistic regression approach (AUC 0.65) (Supplemental Fig. 2) or the AUC for individual markers alone: PSMA (0.62), PSGR (0.65), and PCA3 (0.60). Because of the multifactorial nature of cancer, it is possible that marker A may be positive in one patient and marker B in another. The logistic model fits a logistic curve to the probability of disease for a patient. Using this technique, the combination of both markers is essentially linear, assigning an individual weight to each one. This is valid under the assumption that the expression of both markers is more significant than each marker alone. However, there is a loss of information when the expression of either A or B is, by itself, enough to classify the patients. For that reason, in our model the combined marker was declared positive if at least one of the scores was above its detection threshold.

The next step was to evaluate the behavior of these markers in the PSA diagnostic "gray zone" of 4–10 ng/ml when no prior biopsy information was available. We demonstrated that the prediction of PCa improved significantly for PSMA (0.74), while PSGR (0.66) and PCA3 (0.61) showed a similar performance. For the combined model, we found 0.82. A combination with PSA was not analyzed, since it is not considered predictive within the "gray zone."

In order to translate these findings to the clinics, it is necessary to consider the relative importance of false negatives and false positives. Consequently, we proposed to fix the sensitivity at a high range (90–100%) and calculate the corresponding specificity. We chose a high percentage, because in a decisive test for whether a biopsy should be performed or not, a false negative would be worse than a false positive. This is because the consequences of missing a cancer case could end up being fatal to a patient [36]. Close to 100% sensitivity, the specificity for the individual marker often drops dramatically. In this study at 100% sensitivity, the specificity for PSMA, PSGR, and PCA3 was 0% while the combined model (3M) maintained a specificity of 29%. There was no further advantage using 3M in combination with PSA (29%). The data are consistent with a previous study, where we found a specificity for PSGR of 2.8% and for PCA3

of 2.1% and 13.4% for a combination of both markers (PCA3 ∨ PSGR) [15].

There are several studies that followed the same approach of combining PCA3 with other biomarkers in order to increase the detection of PCa. Hessels et al. [5] performed a study using urine sediments from 108 patients where the authors combined PCA3 with the TMPRSS2:ERG fusion status. The combination of both markers remarkably increased the sensitivity for the detection of PCa from 62% (PCA3 test alone) to 73%. Another important study came from Laxman et al. [7], who combined PCA3 with GOLPH2, SPINK1, and TMPRSS2:ERG fusion status. They founded a sensitivity of 62% and a specificity of 75% with an AUC of 0.76 for the combined model and for PCA3, 75% and 56%, respectively. Ouyang et al. [10] developed a duplex quantitative polymerase chain reaction assay for the detection of PCa, based on the quantification of AMACR and PCA3 in urine sediments. For PCA3, they achieved a sensitivity of 72% and a specificity of 53%.

The principle that underlies the combined biomarker approach is consistent with tests offered for the detection of PCa in tissue specimens [8,9]. Our combined model results provided overall increased sensitivity without decreasing the specificity. Translated to the clinics, we achieved the same high sensitivity as the PSA-test alone, but we increased the specificity considerably. For example, using this method at a sensitivity of 96%, 33.7% of the biopsies performed on patients having a PSA value in the "gray zone" and no prior biopsy information, could have been saved from having the procedure. This would correspond to an approximate yearly number of 438,100 biopsies avoided throughout the United States. We used the formula of: % biopsies saved = true negatives (test negative and biopsy negative) + false negatives (test negative and biopsy positive)/all. Although this would imply that one could save a biopsy by incorrectly classifying (test negative and biopsy positive) a patient as not having PCa, the number of false negatives to obtain a sensitivity of 96% is negligible (in this study one or two patients respectively; NPV $\geq$ 95%).

Another interesting benefit would be the improved ability to decide whether to perform a re-biopsy or not. In this case, sensitivity could be lower in order to get a higher specificity, as it has been observed with the common cut-off value of 35 in the PCA3-test. In guiding repeat biopsy decisions, PCA3 alone has shown its clinical utility in men with $\geq$1 negative biopsy [34]. This is achieved with the common cut-off value of 35 in the PCA3-test. However, this was not an objective of this article. In the future, it would be interesting to analyze PSMA and PSGR in patients with repeat biopsies and in patients who have shown recurrence after prostatectomy, in order to establish specific thresholds or to generate predictive nomograms.

A possible limitation of this and other studies could be that the definition used for patients considered negative for PCa is based on recent negative biopsies. However, this definition can be problematic, since approximately 20% of the patients having a negative prostate biopsy will actually be diagnosed with PCa at a later date [37]. Consequently, many men with negative biopsy findings undergo repeated PB to rule out PCa, meaning that there are a significant number of individuals with a positive test who are positive, though their biopsy results were negative. An accurate, controlled approach is called for, which will provide a validation of the biomarker profile in a prospective study or the use of retrospective studies on age-matched patients with a 5-year follow up [10]. Another possible limitation is the fact, that patients with negative biopsy results, tended to gain less cells of prostatic origin in urines after PM than their malign counterparts. A possible explication could be the loss cell–cell contacts and epithelial polarization in PCa [38].

Despite the obvious limitations of our small cohort of patients, the data here support the applicability of this non-invasive assay toward an accurate diagnostic test for PCa. These results support the examination of larger cohorts, across multiple institutions, for further validation before they can be used as a guideline for the urologist to lower the rate of unnecessary biopsies.

## CONCLUSION

In summary, we have shown that a multiplexed qPCR assay on urine sediments from patients presenting for PB can significantly improve the predictive ability when compared to PCA3 or PSA alone. In the future, a multiplexed, urine-based diagnostic test for PCa with a higher specificity, but the same sensitivity as the serum-PSA test, could be used to determine better which patients should undergo biopsy.

## ACKNOWLEDGMENTS

We want to thank everyone in the Department of Urology at the Vall d'Hebron Hospital for their enduring helpfulness, flexibility, and disposition in the collection of the huge number of urine samples used in this study. The authors would also like to thank L. Piccione for correction of the manuscript.

## REFERENCES


1. Jemal A, Siegel R, Xu J, Ward E. Cancer statistics. CA Cancer J Clin 2010;60(5):277–300.
2. Ferlay J, Parkin DM, Steliarova-Foucher E. Estimates of cancer incidence and mortality in Europe in 2008. Eur J Cancer 2010;46(4):765–781.
3. Loeb S, Gashti SN, Catalona WJ. Exclusion of inflammation in the differential diagnosis of an elevated prostate-specific antigen (PSA). Urol Oncol 2009;27(1):64–66.
4. Pannek J, Partin AW. Prostate-specific antigen: What's new in 1997. Oncology (Williston Park) 1997;11(9):1273–1278; discussion 1279–1282.
5. Hessels D, Smit FP, Verhaegh GW, Witjes JA, Cornel EB, Schalken JA. Detection of TMPRSS2-ERG fusion transcripts and prostate cancer antigen 3 in urinary sediments may improve diagnosis of prostate cancer. Clin Cancer Res 2007;13(17):5103–5108.
6. Hoque MO, Topaloglu O, Begum S, Henrique R, Rosenbaum E, Van Criekinge W, Westra WH, Sidransky D. Quantitative methylation-specific polymerase chain reaction gene patterns in urine sediment distinguish prostate cancer patients from control subjects. J Clin Oncol 2005;23(27):6569–6575.
7. Laxman B, Morris DS, Yu J, Siddiqui J, Cao J, Mehra R, Lonigro RJ, Tsodikov A, Wei JT, Tomlins SA, Chinnaiyan AM. A first-generation multiplex biomarker analysis of urine for the early detection of prostate cancer. Cancer Res 2008;68(3):645–649.
8. Landers KA, Burger MJ, Tebay MA, Purdie DM, Scells B, Samaratunga H, Lavin MF, Gardiner RA. Use of multiple biomarkers for a molecular diagnosis of prostate cancer. Int J Cancer 2005;114(6):950–956.
9. Schmidt U, Fuessel S, Koch R, Baretton GB, Lohse A, Tomasetti S, Unversucht S, Froehner M, Wirth MP, Meye A. Quantitative multi-gene expression profiling of primary prostate cancer. Prostate 2006;66(14):1521–1534.
10. Ouyang B, Bracken B, Burke B, Chung E, Liang J, Ho SM. A duplex quantitative polymerase chain reaction assay based on quantification of alpha-methylacyl-CoA racemase transcripts and prostate cancer antigen 3 in urine sediments improved diagnostic accuracy for prostate cancer. J Urol 2009;181(6):2508–2513; discussion 2513–2504.
11. Vener T, Derecho C, Baden J, Wang H, Rajpurohit Y, Skelton J, Mehrotra J, Varde S, Chowdary D, Stallings W, Leibovich B, Robin H, Pelzer A, Schafer G, Auprich M, Mannweiler S, Amersdorfer P, Mazumder A. Development of a multiplexed urine assay for prostate cancer diagnosis. Clin Chem 2008;54(5):874–882.
12. Elsasser-Beile U, Buhler P, Wolf P. Targeted therapies for prostate cancer against the prostate specific membrane antigen. Curr Drug Targets 2009;10(2):118–125.
13. Xu LL, Stackhouse BG, Florence K, Zhang W, Shanmugam N, Sesterhenn IA, Zou Z, Srikantan V, Augustus M, Roschke V, Carter K, McLeod DG, Moul JW, Soppett D, Srivastava S. PSGR, a novel prostate-specific gene with homology to a G protein-coupled receptor, is overexpressed in prostate cancer. Cancer Res 2000;60(23):6568–6572.
14. Xu LL, Sun C, Petrovics G, Makarem M, Furusato B, Zhang W, Sesterhenn IA, McLeod DG, Sun L, Moul JW, Srivastava S. Quantitative expression profile of PSGR in prostate cancer. Prostate Cancer Prostatic Dis 2006;9(1):56–61.
15. Rigau M, Morote J, Mir MC, Ballesteros C, Colás E, Garcia M, Ruiz A, Abal M, Planas J, Reventós J, Doll A. PSGR and PCA3 as biomarkers for the detection of prostate cancer in urine. Prostate 2010;70(16):1760–1767.
16. Bussemakers MJ, van Bokhoven A, Verhaegh GW, Smit FP, Karthaus HF, Schalken JA, Debruyne FM, Ru N, Isaacs WB. DD3: A new prostate-specific gene, highly overexpressed in prostate cancer. Cancer Res 1999;59(23):5975–5979.
17. de Kok JB, Verhaegh GW, Roelofs RW, Hessels D, Kiemeney LA, Aalders TW, Swinkels DW, Schalken JA. DD3(PCA3), a very sensitive and specific marker to detect prostate tumors. Cancer Res 2002;62(9):2695–2698.
18. Remzi M, Fong YK, Dobrovits M, Anagnostou T, Seitz C, Waldert M, Harik M, Marihart S, Marberger M, Djavan B. The Vienna nomogram: Validation of a novel biopsy strategy defining the optimal number of cores based on patient age and total prostate volume. J Urol 2005;174(4 Pt 1):1256–1260; discussion 1260–1251; author reply 1261.
19. Hessels D, Schalken JA. The use of PCA3 in the diagnosis of prostate cancer. Nat Rev Urol 2009;6(5):255–261.
20. Marks LS, Fradet Y, Deras IL, Blase A, Mathis J, Aubin SM, Cancio AT, Desaulniers M, Ellis WJ, Rittenhouse H, Groskopf J. PCA3 molecular urine assay for prostate cancer in men undergoing repeat biopsy. Urology 2007;69(3):532–535.
21. van Gils MP, Hessels D, van Hooij O, Jannink SA, Peelen WP, Hanssen SL, Witjes JA, Cornel EB, Karthaus HF, Smits GA, Dijkman GA, Mulders PF, Schalken JA. The time-resolved fluorescence-based PCA3 test on urinary sediments after digital rectal examination; a Dutch multicenter validation of the diagnostic performance. Clin Cancer Res 2007;13(3):939–943.
22. Etzioni R, Kooperberg C, Pepe M, Smith R, Gann PH. Combining biomarkers to detect disease with application to prostate cancer. Biostatistics 2003;4(4):523–538.
23. Baker SG. Identifying combinations of cancer markers for further study as triggers of early intervention. Biometrics 2000;56(4):1082–1087.
24. Krzanowski WJ, Hand DJ. ROC curves for continuous data. J Biopharm Stat 2009;20(2):485–487.
25. Benjamini Y, Hochberg Y. Controlling the false discovery rate: A practical and powerful approach to multiple testing. J R Stat Soc Ser B (Methodol) 1995;57(1):289–300.
26. Steyerberg EW, Harrell FE Jr, Borsboom GJ, Eijkemans MJ, Vergouwe Y, Habbema JD. Internal validation of predictive models: Efficiency of some procedures for logistic regression analysis. J Clin Epidemiol 2001;54(8):774–781.
27. Djavan B, Remzi M, Schulman CC, Marberger M, Zlotta AR. Repeat prostate biopsy: Who, how and when? A review. Eur Urol 2002;42(2):93–103.
28. Iwakiri J, Granbois K, Wehner N, Graves HC, Stamey T. An analysis of urinary prostate specific antigen before and after radical prostatectomy: Evidence for secretion of prostate specific antigen by the periurethral glands. J Urol 1993;149(4):783–786.
29. Laxman B, Tomlins SA, Mehra R, Morris DS, Wang L, Helgeson BE, Shah RB, Rubin MA, Wei JT, Chinnaiyan AM. Noninvasive detection of TMPRSS2: ERG fusion transcripts in the urine of men with prostate cancer. Neoplasia 2006;8(10):885–888.
30. Hessels D, Klein Gunnewiek JM, van Oort I, Karthaus HF, van Leenders GJ, van Balken B, Kiemeney LA, Witjes JA, Schalken JA. DD3(PCA3)-based molecular urine analysis for the diagnosis of prostate cancer. Eur Urol 2003;44(1):8–15; discussion 15–16.

31. Ghosh A, Heston WD. Tumor target prostate specific membrane antigen (PSMA) and its regulation in prostate cancer. J Cell Biochem 2004;91(3):528–539.

32. Saramaki OR, Harjula AE, Martikainen PM, Vessella RL, Tammela TL, Visakorpi T. TMPRSS2: ERG fusion identifies a subgroup of prostate cancers with a favorable prognosis. Clin Cancer Res 2008;14(11):3395–3400.

33. Marks LS, Bostwick DG. Prostate cancer specificity of PCA3 gene testing: Examples from clinical practice. Rev Urol 2008;10(3):175–181.

34. Haese A, de la Taille A, van Poppel H, Marberger M, Stenzl A, Mulders PF, Huland H, Abbou CC, Remzi M, Tinzl M, Feyerabend S, Stillebroer AB, van Gils MP, Schalken JA. Clinical utility of the PCA3 urine assay in European men scheduled for repeat biopsy. Eur Urol 2008;54(5):1081–1088.

35. Shultz EK. Multivariate receiver-operating characteristic curve analysis: Prostate cancer screening as an example. Clin Chem 1995;41(8 Pt 2):1248–1255.

36. Pepe MS, Feng Z, Janes H, Bossuyt PM, Potter JD. Pivotal evaluation of the accuracy of a biomarker used for classification or prediction: Standards for study design. J Natl Cancer Inst 2008;100(20):1432–1438.

37. Cervera Deval J, Morales Olaya FJ, Jornet Fayos J, Gonzalez Anon M. Diagnostic value of the second prostate biopsies in males of risk. Study stratified by value of PSA. Actas Urol Esp 2004;28(9):666–671.

38. Harma V, Virtanen J, Makela R, Happonen A, Mpindi JP, Knuuttila M, Kohonen P, Lotjonen J, Kallioniemi O, Nees M. A comprehensive panel of three-dimensional models for studies of prostate cancer growth, invasion and drug responses. PLoS ONE 5(5):e10431.

# ANNEX

**SUPPLEMENTAL DATA "PSGR and PCA3 as Biomarkers for the Detection of Prostate Cancer in Urine"**

**Supplemental Table 1. Clinical Data of all patients used in this study.**

**Supplemental Figure 1. Characterization of PSGR as a urine-based biomarker for prostate cancer (box-and-whisker plots). A and D:** Relative levels of PSGR and PCA3 comparing positive and negative DRE. **B and E:** Relative levels of PSGR and PCA3 comparing Gleason < 7, 7 and > 7. **C and F:** Relative levels of PSGR and PCA3 comparing stages T2a, T2b-c and T3 of PSGR and PCA3.

**Supplemental Table 1. Clinical Data of all patients used in this study.**

| N° | Diagnosis | Diagnosis | Age | Re-biopsy | PSA | Free PSA | Ratio | PST Vol. | DRE (neg=0, pos=1) | PSAD | Gleason | Stage |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 | HG PIN | Benign | 55 | 0 | 6.60 | 1.03 | 0.16 | 35.0 | 1 | 0.19 | | |
| 2 | HG PIN | Benign | 63 | 0 | 6.30 | 1.01 | 0.16 | 73.0 | 0 | 0.09 | | |
| 3 | BPH | Benign | 52 | 0 | 6.40 | 1.12 | 0.17 | 27.0 | 0 | 0.24 | | |
| 4 | BPH | Benign | 73 | 0 | 8.20 | 1.25 | 0.15 | ND | 0 | ND | | |
| 5 | HG PIN | Benign | 57 | 2 | 15.20 | ND | ND | 41.0 | 0 | 0.37 | | |
| 6 | BPH/Prostatitis | Benign | 71 | 0 | 5.90 | 0.77 | 0.13 | 53.7 | 1 | 0.11 | | |
| 7 | Prostatitis | Benign | 65 | 1 | 5.10 | 3.24 | 0.15 | ND | 0 | ND | | |
| 8 | BPH | Benign | 69 | 0 | 8.70 | 1.26 | 0.14 | 36.0 | 0 | 0.24 | | |
| 9 | BPH | Benign | 69 | 1 | 4.70 | 1.50 | 0.32 | 59.4 | 1 | 0.08 | | |
| 10 | HG PIN | Benign | 65 | 0 | 6.6 | 2.26 | 0.34 | 91.0 | 0 | 0.07 | | |
| 11 | HG PIN | Benign | 54 | 0 | 4.8 | 0.98 | 0.21 | 31.4 | 0 | 0.15 | | |
| 12 | N | Benign | 63 | 1 | 5.9 | 0.87 | 0.15 | 54.0 | 1 | 0.11 | | |
| 13 | HG PIN | Benign | 61 | 0 | 8.80 | 1.55 | 0.18 | ND | 0 | ND | | |
| 14 | HG PIN | Benign | 62 | 0 | 4.50 | 0.92 | 0.21 | 65.4 | 1 | 0.07 | | |
| 15 | BPH/Prostatitis | Benign | 62 | 0 | 5.20 | 0.43 | 0.08 | 20.0 | 0 | 0.26 | | |
| 16 | HG PIN | Benign | 69 | 0 | 6.8 | 1.49 | 0.22 | 38.7 | 0 | 0.18 | | |
| 17 | N | Benign | 73 | 0 | 1.50 | 0.38 | 0.25 | 33.0 | 1 | 0.05 | | |
| 18 | HG PIN | Benign | 68 | 1 | 4.5 | 0.70 | 0.16 | 87.5 | 0 | 0.05 | | |
| 19 | PIA | Benign | 60 | 1 | 3.10 | 0.68 | 0.22 | 43.9 | 0 | 0.07 | | |
| 20 | HG PIN | Benign | 66 | 0 | 24.7 | 1.60 | 0.70 | 40.2 | 0 | 0.61 | | |
| 21 | BPH | Benign | 52 | 0 | 3.60 | 1.03 | 0.29 | 49.0 | 0 | 0.07 | | |
| 22 | BPH | Benign | 73 | 0 | 93.60 | 0.04 | 0.00 | 40.8 | 0 | 2.29 | | |
| 23 | N | Benign | 74 | 0 | 10.40 | 0.80 | 0.08 | 44.6 | 0 | 0.23 | | |
| 24 | HG PIN | Benign | 72 | 0 | 8.90 | 1.78 | 0.20 | 53.9 | 0 | 0.17 | | |
| 25 | BPH | Benign | 62 | 1 | 5.45 | 0.76 | 0.14 | 42.6 | 0 | 0.13 | | |
| 26 | PIA | Benign | 72 | 2 | 10 | 1.44 | 0.14 | 69.0 | 1 | 0.14 | | |
| 27 | PIA | Benign | 59 | 0 | 5.90 | 0.13 | 0.02 | 57.1 | 0 | 0.10 | | |
| 28 | BPH | Benign | 65 | 0 | 5.40 | 1.51 | 0.28 | 33.5 | 0 | 0.16 | | |
| 29 | HG PIN | Benign | 76 | 0 | 6.70 | 1.65 | 0.25 | 35.7 | 1 | 0.19 | | |
| 30 | HG PIN | Benign | 54 | 0 | 13.9 | 2.42 | 0.17 | 99.4 | 1 | 0.14 | | |

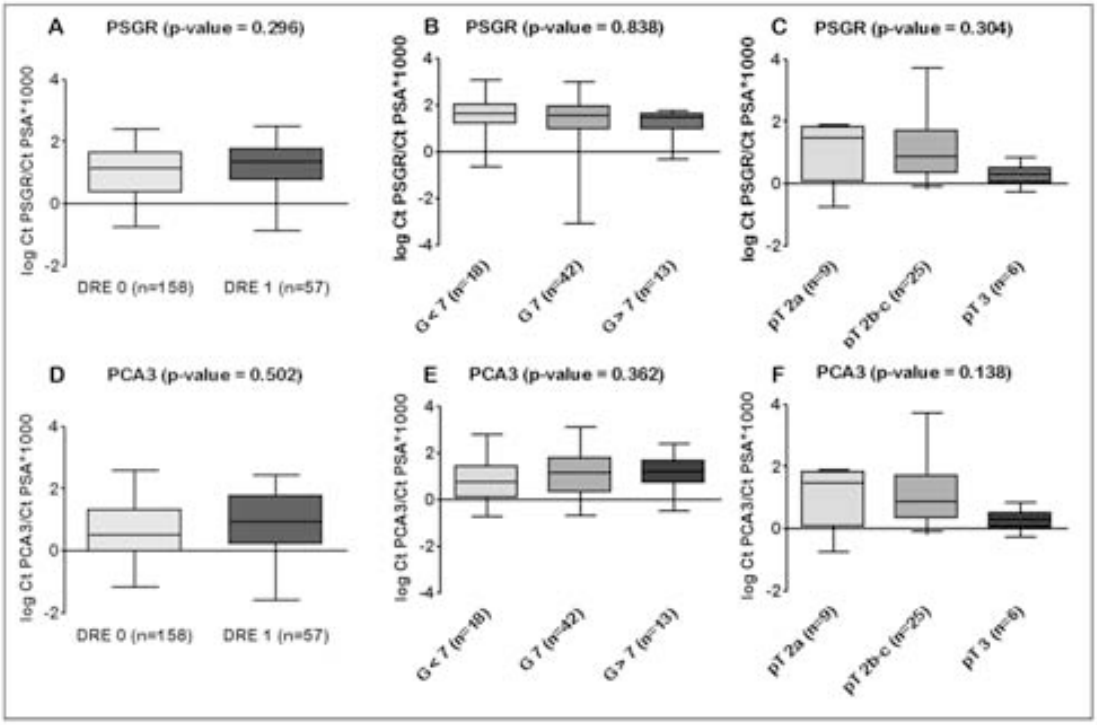
| | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| 31 | HG PIN | Benign | 65 | 0 | 4.9 | 0.86 | 0.18 | 46.5 | 0 | 0.11 |
| 32 | HG PIN | Benign | 66 | 1 | 8.90 | 0.26 | 0.03 | 49.0 | 0 | 0.18 |
| 33 | BPH | Benign | 61 | 1 | 5.7 | 0.90 | 0.16 | 41.4 | 0 | 0.14 |
| 34 | BPH | Benign | 63 | 0 | 4.40 | 0.97 | 0.22 | 46.4 | 0 | 0.09 |
| 35 | HG PIN | Benign | 63 | 1 | 11.10 | 1.70 | 0.15 | 60.0 | 0 | 0.19 |
| 36 | BPH/Prostatitis | Benign | 73 | 0 | 7.90 | 1.24 | 0.16 | 45.8 | 0 | 0.17 |
| 37 | N | Benign | 55 | 0 | 5.20 | 0.93 | 0.18 | 22.3 | 0 | 0.23 |
| 38 | HG PIN | Benign | 57 | 0 | 5.9 | 1.47 | 0.25 | 62.5 | 0 | 0.09 |
| 39 | HG PIN | Benign | 60 | 0 | 12 | 1.28 | 0.11 | 45.2 | 0 | 0.27 |
| 40 | N | Benign | 69 | 2 | 7.7 | 1.03 | 0.13 | 78.1 | 0 | 0.10 |
| 41 | HG PIN | Benign | 64 | 0 | 9 | 1.58 | 0.17 | 76.2 | 0 | 0.12 |
| 42 | HG PIN | Benign | 62 | 2 | 5.4 | 0.68 | 0.13 | 47.0 | 0 | 0.11 |
| 43 | BPH/Prostatitis | Benign | 69 | 1 | 11.6 | 1.80 | 0.15 | 67.3 | 0 | 0.17 |
| 44 | HG PIN | Benign | 64 | 0 | 9.6 | 0.88 | 0.09 | 46.0 | 0 | 0.21 |
| 45 | HG PIN | Benign | 63 | 0 | 13.7 | 1.92 | 0.14 | 100.0 | 0 | 0.14 |
| 46 | HG PIN | Benign | 73 | 1 | 4.5 | 0.90 | 0.20 | 48.0 | 0 | 0.09 |
| 47 | PIA | Benign | 68 | 0 | 5.7 | 0.82 | 0.14 | 76.2 | 0 | 0.07 |
| 48 | HG PIN | Benign | 63 | 0 | 15.2 | 1.14 | 0.07 | 41.1 | 0 | 0.37 |
| 49 | HG PIN | Benign | 67 | 0 | 7.3 | 0.73 | 0.10 | 35.6 | 0 | 0.21 |
| 50 | PIA | Benign | 64 | 1 | 7.2 | 1.37 | 0.19 | 48.2 | 0 | 0.15 |
| 51 | BPH/Prostatitis | Benign | 53 | 0 | 6.00 | 0.73 | 0.12 | ND | 0 | ND |
| 52 | HG PIN | Benign | 63 | 1 | 5.7 | 0.90 | 0.16 | 58.5 | 0 | 0.10 |
| 53 | N | Benign | 61 | 0 | 6.1 | 1.17 | 0.19 | 29.9 | 0 | 0.20 |
| 54 | N | Benign | 66 | 0 | 5.7 | 0.67 | 0.12 | 45.3 | 0 | 0.13 |
| 55 | N | Benign | 58 | 1 | 4.90 | 0.50 | 0.10 | 32.4 | 0 | 0.15 |
| 56 | BPH | Benign | 67 | 0 | 5.20 | 0.88 | 0.17 | 24.1 | 0 | 0.22 |
| 57 | BPH/Prostatitis | Benign | 58 | 0 | 4.6 | 0.78 | 0.17 | 44.0 | 1 | 0.10 |
| 58 | PIA | Benign | 78 | 0 | 28.2 | 4.75 | 0.17 | 94.5 | 1 | 0.30 |
| 59 | BPH | Benign | 63 | 2 | 5.90 | 0.82 | 0.14 | 60.8 | 0 | 0.10 |
| 60 | HG PIN | Benign | 71 | 0 | 8.1 | 2.01 | 0.25 | 95.1 | 1 | 0.09 |
| 61 | HG PIN | Benign | 78 | 0 | 8.3 | 1.47 | 0.18 | 94.0 | 0 | 0.09 |
| 62 | HG PIN | Benign | 66 | 0 | 8.80 | 0.09 | 0.01 | 27.6 | 0 | 0.32 |
| 63 | BPH/Prostatitis | Benign | 67 | 0 | 7.80 | 1.24 | 0.16 | 68.2 | 0 | 0.11 |
| 64 | HG PIN | Benign | 62 | 0 | 12.19 | 1.67 | 0.14 | 82.2 | 0 | 0.15 |
| 65 | BPH | Benign | 76 | 0 | 4.7 | 1.28 | 0.27 | 41.2 | 0 | 0.11 |
| 66 | BPH | Benign | 59 | 0 | 4.70 | 0.90 | 0.19 | 38.7 | 0 | 0.12 |
| 67 | BPH | Benign | 70 | 0 | 4.90 | 1.14 | 0.23 | 48.00 | 0 | 0.10 |
| 68 | HG PIN | Benign | 73 | 1 | 4.27 | 0.94 | 0.22 | 50.0 | 0 | 0.09 |
| 69 | N | Benign | 71 | 1 | 6.60 | 1.02 | 0.16 | 59.0 | 0 | 0.11 |
| 70 | BPH | Benign | 60 | 0 | 1.60 | 0.56 | 0.35 | 34.0 | 0 | 0.05 |
| 71 | HG PIN | Benign | 62 | 0 | 13.8 | 1.85 | 0.13 | 39.9 | 0 | 0.35 |
| 72 | PIA | Benign | 65 | 1 | 2.5 | 0.46 | 0.18 | 38.9 | 1 | 0.06 |
| 73 | HG PIN | Benign | 57 | 0 | 8.7 | 1.22 | 0.14 | 28.6 | 0 | 0.30 |
| 74 | HG PIN | Benign | 68 | 0 | 6.8 | 1.37 | 0.20 | 63.8 | 0 | 0.11 |
| 75 | HG PIN | Benign | 72 | 0 | 7.5 | 1.29 | 0.17 | 35.0 | 0 | 0.21 |
| 76 | N | Benign | 76 | 0 | 13.30 | 1.10 | 0.08 | 39.9 | 1 | 0.33 |
| 77 | BPH | Benign | 62 | 0 | 6.30 | 0.44 | 0.07 | 39.2 | 0 | 0.16 |
| 78 | BPH | Benign | 78 | 2 | 13.00 | 1.76 | 0.14 | 95.0 | 0 | 0.14 |
| 79 | BPH | Benign | 79 | 0 | 6.60 | 0.88 | 0.13 | 41.0 | 0 | 0.16 |
| 80 | BPH | Benign | 72 | 3 | 6.90 | 0.35 | 0.05 | 24.5 | 0 | 0.28 |
| 81 | PIA | Benign | 64 | 0 | 8.2 | 1.90 | 0.23 | 62.8 | 0 | 0.13 |
| 82 | N | Benign | 85 | 0 | 9.5 | 2.79 | 0.29 | 85.7 | 0 | 0.11 |
| 83 | HG PIN | Benign | 71 | 0 | 6.9 | 0.55 | 0.08 | 23.1 | 0 | 0.30 |
| 84 | BPH/Prostatitis | Benign | 74 | 0 | 4.40 | 0.92 | 0.21 | 62.6 | 1 | 0.07 |

| | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| 85 | PIA | Benign | 60 | 0 | 7.4 | 1.23 | 0.17 | 35.7 | 1 | 0.21 |
| 86 | HG PIN | Benign | 57 | 1 | 5.30 | 1.52 | 0.29 | 42.0 | 1 | 0.13 |
| 87 | HG PIN | Benign | 71 | 0 | 4.8 | 1.16 | 0.24 | 83.1 | 0 | 0.06 |
| 88 | HG PIN | Benign | 62 | 1 | 5.9 | 0.78 | 0.13 | 61.1 | 0 | 0.10 |
| 89 | BPH | Benign | 66 | 0 | 12.90 | 1.42 | 0.11 | 48.0 | 0 | 0.27 |
| 90 | BPH | Benign | 72 | 1 | 4.80 | 1.42 | 0.29 | 43.0 | 0 | 0.11 |
| 91 | BPH/Prostatitis | Benign | 67 | 0 | 5 | 0.53 | 0.11 | 61.5 | 0 | 0.08 |
| 92 | PIA | Benign | 60 | 1 | 5.4 | 0.65 | 0.12 | ND | 0 | ND |
| 93 | PIA | Benign | 63 | 1 | 6.50 | 1.40 | 0.21 | 63.3 | 0 | 0.10 |
| 94 | HG PIN | Benign | 74 | 0 | 7.1 | 1.17 | 0.17 | 36.3 | 0 | 0.20 |
| 95 | BPH/Prostatitis | Benign | 63 | 0 | 5.90 | 1.08 | 0.18 | 103.0 | 0 | 0.06 |
| 96 | BPH/Prostatitis | Benign | 49 | 0 | 3.92 | 4.65 | 0.16 | 40.1 | 0 | 0.10 |
| 97 | N | Benign | 64 | 0 | 3.8 | 1.06 | 0.28 | 61.8 | 0 | 0.06 |
| 98 | HG PIN | Benign | 57 | 2 | 6.7 | 1.27 | 0.19 | 31.2 | 0 | 0.21 |
| 99 | HG PIN | Benign | 44 | 0 | 4.3 | 0.29 | 0.01 | 16.0 | 0 | 0.27 |
| 100 | BPH | Benign | 64 | 0 | 6.80 | 0.91 | 0.13 | ND | 0 | ND |
| 101 | BPH/Prostatitis | Benign | 76 | 3 | 23.1 | 2.13 | 0.09 | 50.0 | 0 | 0.46 |
| 102 | BPH | Benign | 66 | 1 | 9.70 | 1.78 | 0.18 | 66.7 | 0 | 0.15 |
| 103 | HG PIN | Benign | 71 | 0 | 8.60 | 2.52 | 0.29 | 54.1 | 1 | 0.16 |
| 104 | PIA | Benign | 64 | 0 | 7.50 | 1.44 | 0.19 | 31.0 | 0 | 0.24 |
| 105 | PIA | Benign | 59 | 0 | 6.60 | 1.47 | 0.22 | 50.0 | 1 | 0.13 |
| 106 | HG PIN | Benign | 65 | 0 | 5.4 | 1.07 | 0.20 | 58.5 | 0 | 0.09 |
| 107 | PIA | Benign | 66 | 0 | 8.20 | 1.05 | 0.13 | 40.0 | 0 | 0.21 |
| 108 | HG PIN | Benign | 62 | 1 | 9.6 | 1.45 | 0.15 | 42.9 | 0 | 0.22 |
| 109 | N | Benign | 73 | 0 | 9.70 | 1.80 | 0.19 | 58.0 | 0 | 0.17 |
| 110 | HG PIN | Benign | 58 | 0 | 5.4 | 1.19 | 0.22 | 82.1 | 0 | 0.07 |
| 111 | PIA | Benign | 57 | 0 | 7.90 | 1.01 | 0.13 | 50.0 | 0 | 0.16 |
| 112 | HG PIN | Benign | 56 | 0 | 12.3 | 3.17 | 0.26 | 59.9 | 0 | 0.21 |
| 113 | HG PIN | Benign | 61 | 0 | 2.7 | 0.42 | 0.16 | 28.7 | 0 | 0.09 |
| 114 | HG PIN | Benign | 66 | 1 | 4.30 | 2.24 | 0.52 | 60.0 | 0 | 0.07 |
| 115 | HG PIN | Benign | 73 | 2 | 7.1 | 0.82 | 0.12 | 87.5 | 0 | 0.08 |
| 116 | HG PIN | Benign | 55 | 1 | 5.9 | 0.79 | 0.13 | 69.9 | 0 | 0.08 |
| 117 | HG PIN | Benign | 68 | 0 | 6.5 | 0.82 | 0.13 | 24.7 | 0 | 0.26 |
| 118 | PIA | Benign | 74 | 0 | 5.90 | 1.10 | 0.19 | 45.0 | 0 | 0.13 |
| 119 | HG PIN | Benign | 69 | 0 | 11.4 | 2.51 | 0.22 | 69.1 | 1 | 0.16 |
| 120 | HG PIN | Benign | 63 | 0 | 7.3 | 0.70 | 0.10 | 57.6 | 0 | 0.13 |
| 121 | BPH | Benign | 53 | 0 | 6.10 | ND | ND | 68.0 | 0 | 0.09 |
| 122 | HG PIN | Benign | 77 | 3 | 14.5 | 2.73 | 0.19 | 51.7 | 0 | 0.28 |
| 123 | N | Benign | 62 | 0 | 8.50 | 1.23 | 0.14 | 35.0 | 1 | 0.24 |
| 124 | HG PIN | Benign | 69 | 0 | 12.1 | 2.72 | 0.22 | 35.0 | 1 | 0.35 |
| 125 | PIA | Benign | 52 | 0 | 8.60 | 1.88 | 0.22 | 41.0 | 0 | 0.21 |
| 126 | BPH | Benign | 72 | 0 | 8.80 | 2.45 | 0.28 | 75.0 | 0 | 0.12 |
| 127 | PIA | Benign | 69 | 2 | 7.70 | 1.15 | 0.15 | 40.3 | 1 | 0.19 |
| 128 | HG PIN | Benign | 61 | 0 | 5.3 | 1.12 | 0.21 | 60.5 | 1 | 0.09 |
| 129 | PIA | Benign | 65 | 0 | 11.40 | 1.34 | 0.12 | 45.0 | 0 | 0.25 |
| 130 | HG PIN | Benign | 68 | 0 | 9.2 | 1.25 | 0.14 | 54.0 | 0 | 0.17 |
| 131 | HG PIN | Benign | 70 | 0 | 5 | 0.48 | 0.10 | 37.1 | 1 | 0.13 |
| 132 | HG PIN | Benign | 64 | 0 | 13.8 | 1.15 | 0.08 | 143.7 | 1 | 0.10 |
| 133 | HG PIN | Benign | 68 | 0 | 6.3 | 1.31 | 0.21 | ND | 1 | ND |
| 134 | BPH/Prostatitis | Benign | 62 | 0 | 5.40 | 0.99 | 0.18 | 65.0 | 1 | 0.08 |
| 135 | BPH | Benign | 71 | 0 | 13.20 | 2.82 | 0.21 | 45.0 | 1 | 0.29 |
| 136 | N | Benign | 76 | 0 | 4 | 0.59 | 0.15 | 37.3 | 0 | 0.11 |
| 137 | PIA | Benign | 66 | 1 | 9.40 | 1.12 | 0.12 | 50.0 | 0 | 0.19 |
| 138 | HG PIN | Benign | 60 | 1 | 12 | 0.86 | 0.07 | 29.0 | 0 | 0.41 |

| | | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 139 | HG PIN | Benign | 70 | 0 | 7.2 | 0.98 | 0.14 | 43.5 | 1 | 0.17 | | | | | | |
| 140 | HG PIN | Benign | 76 | 1 | 16.3 | 2.46 | 0.15 | 80.4 | 0 | 0.20 | | | | | | |
| 141 | PIA | Benign | 52 | 0 | 6.10 | 0.80 | 0.13 | 39.9 | 0 | 0.15 | | | | | | |
| 142 | HG PIN | Benign | 61 | 0 | 7.6 | 1.48 | 0.19 | 39.4 | 0 | 0.19 | | | | | | |
| 143 | PCa | PCa | 68 | 0 | 4.70 | 0.79 | 0.17 | 42.0 | 0 | 0.11 | 7 | 3 | + | 4 | T | 1c |
| 144 | PCa | PCa | 65 | 0 | 4.80 | 1.06 | 0.22 | 53.0 | 1 | 0.09 | 6 | 3 | + | 3 | T | 1c |
| 145 | PCa | PCa | 64 | 0 | 166.2 | 43.65 | 0.26 | 31.0 | 1 | 5.36 | 7 | 4 | + | 3 | T | 2b |
| 146 | PCa | PCa | 62 | 0 | 6.70 | 1.54 | 0.23 | 58.0 | 0 | 0.12 | 7 | 3 | + | 4 | T | 1c |
| 147 | PCa | PCa | 68 | 0 | 6.10 | 0.88 | 0.15 | 36.1 | 0 | 0.17 | 6 | 3 | + | 3 | T | 1c |
| 148 | PCa | PCa | 82 | 1 | 26.70 | 3.25 | 0.12 | ND | 0 | ND | 9 | 5 | + | 4 | | |
| 149 | PCa | PCa | 55 | 0 | 10.80 | 1.19 | 0.11 | 31.0 | 0 | 0.35 | 7 | 4 | + | 3 | T | 1c |
| 150 | PCa | PCa | 80 | 0 | 9.40 | 1.31 | 0.14 | 26.7 | 0 | 0.35 | 7 | 3 | + | 4 | T | 1c |
| 151 | PCa | PCa | 67 | 0 | 8.35 | 1.00 | 0.12 | 64.0 | 0 | 0.13 | 6 | 3 | + | 3 | T | 1c |
| 152 | PCa | PCa | 80 | 2 | 14.70 | 0.78 | 0.05 | 38.0 | 1 | 0.39 | 6 | 3 | + | 3 | | |
| 153 | PCa | PCa | 63 | 0 | 4.27 | 1.07 | 0.25 | 26.2 | 1 | 0.16 | 6 | 3 | + | 3 | T | 1c |
| 154 | PCa | PCa | 62 | 0 | 7.8 | 1.09 | 0.14 | ND | 0 | ND | 6 | 3 | + | 3 | T | 1c |
| 155 | PCa | PCa | 64 | 0 | 5.80 | 0.70 | 0.12 | 56.3 | 0 | 0.10 | 5 | 2 | + | 3 | T | 1c |
| 156 | PCa | PCa | 60 | 0 | 7.30 | 1.39 | 0.19 | 33.0 | 0 | 0.22 | 7 | 3 | + | 4 | T | 2b |
| 157 | PCa | PCa | 66 | 0 | 4.80 | 0.58 | 0.12 | 56.0 | 0 | 0.09 | 7 | 3 | + | 4 | T | 1c |
| 158 | PCa | PCa | 66 | 0 | 11.40 | 4.69 | 0.41 | 53.0 | 1 | 0.22 | 7 | 4 | + | 3 | T | 1c |
| 159 | PCa | PCa | 60 | 0 | 6.80 | 1.24 | 0.18 | 43.0 | 1 | 0.16 | 7 | 3 | + | 4 | T | 2a |
| 160 | PCa | PCa | 65 | 1 | 7.00 | 1.61 | 0.23 | 162.5 | 0 | 0.04 | 7 | 3 | + | 4 | T | 2 |
| 161 | PCa | PCa | 63 | 0 | 9.4 | 1.33 | 0.14 | 30.0 | 0 | 0.31 | 7 | 3 | + | 4 | T | 1c |
| 162 | PCa | PCa | 71 | 0 | 23.1 | 2.03 | 0.09 | ND | 0 | ND | 7 | 4 | + | 3 | T | 1c |
| 163 | PCa | PCa | 66 | 0 | 16.40 | ND | ND | 25.0 | 1 | 0.66 | 6 | 3 | + | 3 | T | 2a |
| 164 | PCa | PCa | 70 | 1 | 12.00 | 1.32 | 0.11 | 22.0 | 0 | 0.55 | 7 | 3 | + | 4 | T | 1c |
| 165 | PCa | PCa | 70 | 0 | 7.40 | 1.03 | 0.14 | 54.7 | 0 | 0.14 | 6 | 3 | + | 3 | T | 1c |
| 166 | PCa | PCa | 56 | 0 | 189.0 0 | ND | ND | 42.7 | 1 | 4.43 | 1 0 | 5 | + | 5 | | |
| 167 | PCa | PCa | 72 | 1 | 6.3 | 0.63 | 0.10 | 29.2 | 0 | 0.22 | 8 | 4 | + | 4 | T | 1c |
| 168 | PCa | PCa | 58 | 0 | 4.97 | 0.65 | 0.13 | 30.0 | 0 | 0.17 | 7 | 4 | + | 3 | T | 1c |
| 169 | PCa | PCa | 62 | 0 | 77.30 | 6.23 | 0.08 | 66.0 | 1 | 1.17 | 8 | 4 | + | 4 | T | 3a |
| 170 | PCa | PCa | 73 | 1 | 8.80 | 0.62 | 0.07 | 10.0 | 0 | 0.88 | 6 | 3 | + | 3 | T | 1c |
| 171 | PCa | PCa | 57 | 0 | 5.80 | 0.58 | 0.10 | 43.1 | 1 | 0.13 | 7 | 4 | + | 3 | T | 2c |
| 172 | PCa | PCa | 73 | 1 | 5.60 | 0.84 | 0.15 | 50.0 | 0 | 0.11 | 6 | 3 | + | 3 | T | 1c |
| 173 | PCa | PCa | 59 | 0 | 7.80 | 1.15 | 0.15 | 29.0 | 1 | 0.27 | 6 | 3 | + | 3 | T | 3b |
| 174 | PCa | PCa | 55 | 0 | 23.30 | ND | ND | 7.68 | 1 | 3.03 | 7 | 3 | + | 4 | T | 2b |
| 175 | PCa | PCa | 71 | 0 | 7.40 | 1.05 | 0.14 | 50.00 | 0 | 0.15 | 7 | 3 | + | 4 | T | 1c |
| 176 | PCa | PCa | 68 | 0 | 7.4 | 0.95 | 0.13 | 27.0 | 0 | 0.27 | 8 | 4 | + | 4 | T | 1c |
| 177 | PCa | PCa | 77 | 0 | 29.60 | 2.48 | 0.08 | 45.7 | 0 | 0.65 | 1 0 | 5 | + | 5 | T | 1c |
| 178 | PCa | PCa | 58 | 0 | 6.40 | 5.12 | 0.80 | 50.0 | 1 | 0.13 | 7 | 3 | + | 4 | | |
| 179 | PCa | PCa | 64 | 0 | 7.80 | 1.43 | 0.18 | 33.3 | 1 | 0.23 | 7 | 3 | + | 4 | T | 3b |
| 180 | PCa | PCa | 62 | 1 | 4.5 | 0.50 | 0.11 | 28.1 | 1 | 0.16 | 7 | 4 | + | 3 | T | 1c |
| 181 | PCa | PCa | 72 | 1 | 9.20 | 1.07 | 0.12 | 27.9 | 0 | 0.33 | 7 | 4 | + | 3 | T | 1c |
| 182 | PCa | PCa | 67 | 0 | 4.30 | 0.82 | 0.19 | ND | 1 | ND | 6 | 3 | + | 3 | T | 2a |
| 183 | PCa | PCa | 79 | 0 | 5.60 | ND | ND | 16.6 | 0 | 0.34 | 7 | 4 | + | 3 | T | 1c |
| 184 | PCa | PCa | 67 | 0 | 6.7 | 1.68 | 0.25 | 40.4 | 0 | 0.17 | 7 | 3 | + | 4 | T | 1c |
| 185 | PCa | PCa | 67 | 0 | 10.10 | 2.32 | 0.23 | 74.0 | 0 | 0.14 | 7 | 3 | + | 4 | T | 1c |
| 186 | PCa | PCa | 67 | 0 | 17.5 | 3.08 | 0.18 | 41.5 | 1 | 0.42 | 8 | 4 | + | 4 | T | 3a |
| 187 | PCa | PCa | 64 | 0 | 4.20 | 0.43 | 0.10 | 42.0 | 0 | 0.10 | 7 | 3 | + | 4 | T | 1c |
| 188 | PCa | PCa | 66 | 0 | 82.20 | 8.57 | 0.10 | 23.3 | 1 | 3.53 | 6 | 3 | + | 3 | T | 3a |
| 189 | PCa | PCa | 72 | 0 | 10.3 | 1.13 | 0.11 | 18.0 | 0 | 0.57 | 7 | 3 | + | 4 | T | 1c |
| 190 | PCa | PCa | 82 | 0 | 42.80 | 3.25 | 0.08 | ND | 0 | ND | 1 0 | 5 | + | 5 | | |
| 191 | PCa | PCa | 70 | 0 | 7.8 | 1.10 | 0.14 | 26.0 | 0 | 0.30 | 8 | 4 | + | 4 | T | 1c |
| 192 | PCa | PCa | 59 | 0 | 7.1 | 0.79 | 0.11 | 28.7 | 0 | 0.25 | 7 | 3 | + | 4 | | |
| 193 | PCa | PCa | 73 | 0 | 6.72 | ND | ND | 38.0 | 1 | 0.18 | 6 | 3 | + | 3 | T | 2b |
| 194 | PCa | PCa | 78 | 2 | 21.70 | 1.93 | 0.09 | 30.2 | 0 | 0.72 | 7 | 4 | + | 3 | T | 1c |

| | | | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 195 | PCa | PCa | 58 | 0 | 4.10 | 0.66 | 0.16 | 20.0 | 0 | 0.21 | 8 | 4 | + | 4 | T | 3a |
| 196 | PCa | PCa | 69 | 0 | 29.6 | 2.48 | 0.08 | 21.4 | 1 | 1.38 | 7 | 3 | + | 4 | T | 3a |
| 197 | PCa | PCa | 58 | 0 | 20.60 158.5 | 1.92 | 0.09 | 46.4 | 1 | 0.44 | 7 | 3 | + | 4 | T | 3a |
| 198 | PCa | PCa | 65 | 1 | 0 | 13.34 | 0.08 | 38.4 | 1 | 4.13 | 9 | 5 | + | 4 | T | 3a |
| 199 | PCa | PCa | 62 | 1 | 5.30 | 1.22 | 0.23 | 29.0 | 0 | 0.18 | 7 | 3 | + | 4 | | |
| 200 | PCa | PCa | 61 | 0 | 9.7 | 0.78 | 0.08 | 34.7 | 0 | 0.28 | 7 | 4 | + | 3 | T | 1c |
| 201 | PCa | PCa | 62 | 0 | 9.90 | 3.37 | 0.34 | 110.0 | 0 | 0.09 | 7 | 3 | + | 4 | | |
| 202 | PCa | PCa | 58 | 0 | 5.9 | 1.02 | 0.17 | 38.2 | 0 | 0.15 | 6 | 3 | + | 3 | T | 1c |
| 203 | PCa | PCa | 61 | 0 | 7.8 | 1.09 | 0.14 | 36.5 | 0 | 0.21 | 6 | 3 | + | 3 | T | 1c |
| 204 | PCa | PCa | 70 | 0 | 7.02 | 0.63 | 0.09 | 22.9 | 0 | 0.31 | 7 | 3 | + | 4 | T | 1c |
| 205 | PCa | PCa | 71 | 0 | 5.4 | 0.86 | 0.16 | 30.9 | 0 | 0.17 | 7 | 3 | + | 4 | T | |
| 206 | PCa | PCa | 70 | 0 | 6.3 | 0.63 | 0.10 | 26.8 | 1 | 0.24 | 8 | 4 | + | 4 | T | 3b |
| 207 | PCa | PCa | 82 | 0 | 8.90 | 1.22 | 0.14 | 40.0 | 0 | 0.22 | 7 | 3 | + | 4 | | |
| 208 | PCa | PCa | 58 | 1 | 12.30 | 0.74 | 0.06 | ND | 1 | ND | 7 | 3 | + | 4 | T | 3a |
| 209 | PCa | PCa | 61 | 0 | 7.20 | 0.86 | 0.12 | 29.4 | 1 | 0.24 | 8 | 4 | + | 4 | T | 1c |
| 210 | PCa | PCa | 77 | 0 | 20.9 | 2.71 | 0.13 | 27.0 | 1 | 0.77 | 7 | 4 | + | 3 | T | 2a |
| 211 | PCa | PCa | 64 | 0 | 6.15 | 1.17 | 0.19 | 84.0 | 0 | 0.07 | 6 | 3 | + | 3 | T | 1c |
| 212 | PCa | PCa | 70 | 0 | 4.6 | 0.68 | 0.15 | 51.4 | 1 | 0.09 | 7 | 4 | + | 3 | | |
| 213 | PCa | PCa | 69 | 1 | 10.80 | 1.79 | 0.17 | 59.4 | 0 | 0.18 | 7 | 4 | + | 3 | | |
| 214 | PCa | PCa | 67 | 0 | 4.40 | 1.06 | 0.24 | 22.7 | 0 | 0.19 | 7 | 3 | + | 4 | | |
| 215 | PCa | PCa | 55 | 0 | 4.70 | 1.36 | 0.29 | 27.0 | 0 | 0.17 | 7 | 3 | + | 4 | T | 1c |

**Supplemental Figure 1. Characterization of PSGR as a urine-based biomarker for prostate cancer (box-and-whisker plots).**



**SUPPLEMENTAL DATA "A Three-Gene Panel on Urine Increases PSA Specificity in the Detection of Prostate Cancer"**

**Supplemental Table Ia. Biomarker Comparison of 3M by z test for proportions**

**Supplemental Figure 1. Characterization of urine based biomarker of prostate cancer (box-and-whisker plots).** Relative level of PSA **(A)**, free PSA **(B)**, Prostate volume **(C)** in men with PCa vs. Benign. Relative level of PSA **(D)**, free PSA **(E)**, Prostate volume **(F)** in men within PSA 4-10ng/mL and first biopsy, with PCa vs. Benign.

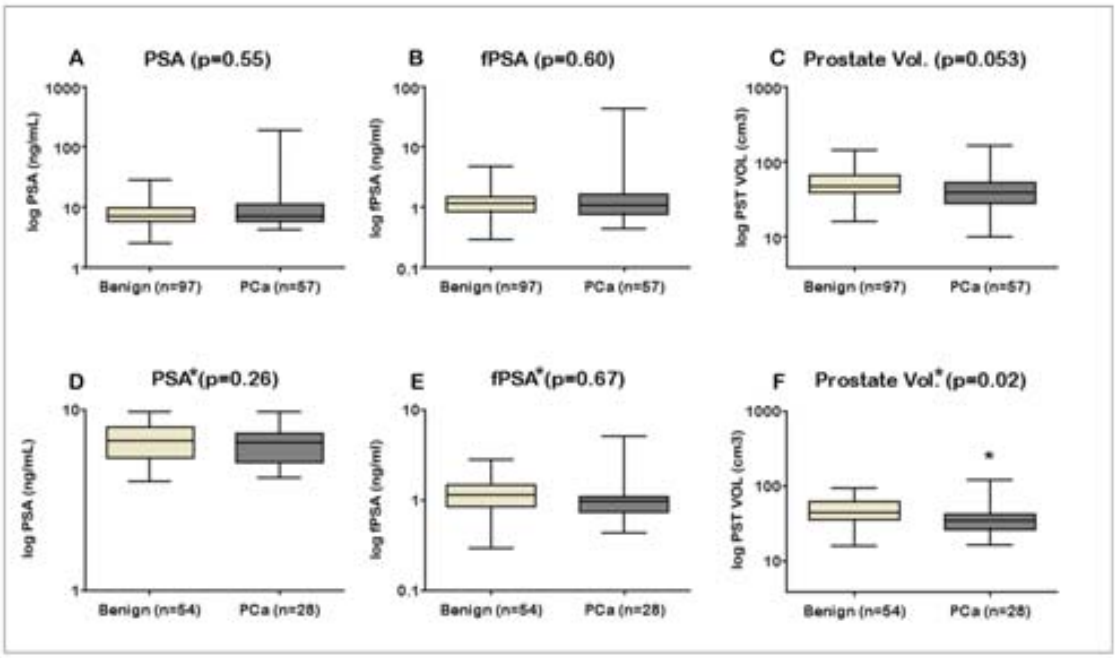
**Supplemental Figure 2. Comparison of logistic regression model with mROC model 3M (PSMAvPSGRvPCA3) in men with prostate cancer detected and men with no cancer detected at biopsy.** MutliROC-curve for 3M (black line) and Logistic regression model ROC-curve for 3M (grey line).

**Supplemental Table Ia. Biomarker Comparison of 3M by z test for proportions (Overall population)**

| | Specificity | | | | p-value | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| Sensitivity | 3M | PSMA | PSGR | PCA3 | PSGR vs. 3M* | PCA3 vs. 3M* | PSMA vs. 3M* | PSGR vs. PCA3 | PSGR vs. PSMA | PSMA vs. PCA3 |
| 98.2% | 33.0% | 3.1% | 0.0% | 0.0% | 0.0002 | 0.0000 | 0.0000 | > 0.999 | 0.9315 | > 0.999 |
| 98.1% | 33.0% | 3.1% | 0.0% | 0.0% | 0.0002 | 0.0000 | 0.0000 | > 0.999 | 0.9315 | > 0.999 |
| 96.5% | 34.0% | 3.1% | 0.0% | 3.1% | 0.0002 | 0.0006 | 0.0001 | > 0.999 | 0.9315 | > 0.999 |
| 96.4% | 34.0% | 3.1% | 0.0% | 3.1% | 0.0002 | 0.0006 | 0.0001 | > 0.999 | 0.9315 | > 0.999 |
| 96.3% | 34.0% | 3.1% | 0.0% | 3.1% | 0.0002 | 0.0006 | 0.0001 | > 0.999 | 0.9315 | > 0.999 |
| 94.7% | 37.1% | 14.4% | 18.6% | 7.2% | 0.0429 | 0.0013 | 0.0158 | > 0.999 | 0.9315 | > 0.999 |
| 94.4% | 37.1% | 14.4% | 18.6% | 7.2% | 0.0429 | 0.0013 | 0.0158 | > 0.999 | 0.9315 | > 0.999 |
| 93.0% | 40.2% | 19.6% | 21.6% | 14.4% | 0.0429 | 0.0084 | 0.0368 | > 0.999 | 0.9315 | > 0.999 |
| 92.9% | 40.2% | 19.6% | 21.6% | 14.4% | 0.0429 | 0.0084 | 0.0368 | > 0.999 | 0.9315 | > 0.999 |
| 92.6% | 40.2% | 19.6% | 21.6% | 14.4% | 0.0429 | 0.0084 | 0.0368 | > 0.999 | 0.9315 | > 0.999 |
| 91.2% | 43.3% | 23.7% | 28.9% | 27.8% | 0.0866 | 0.0811 | 0.0433 | > 0.999 | 0.9315 | > 0.999 |
| 90.7% | 43.3% | 23.7% | 28.9% | 27.8% | 0.0866 | 0.0811 | 0.0433 | > 0.999 | 0.9315 | > 0.999 |
| 89.5% | 45.4% | 28.9% | 35.1% | 29.9% | 0.1591 | 0.0811 | 0.0669 | > 0.999 | 0.9315 | > 0.999 |

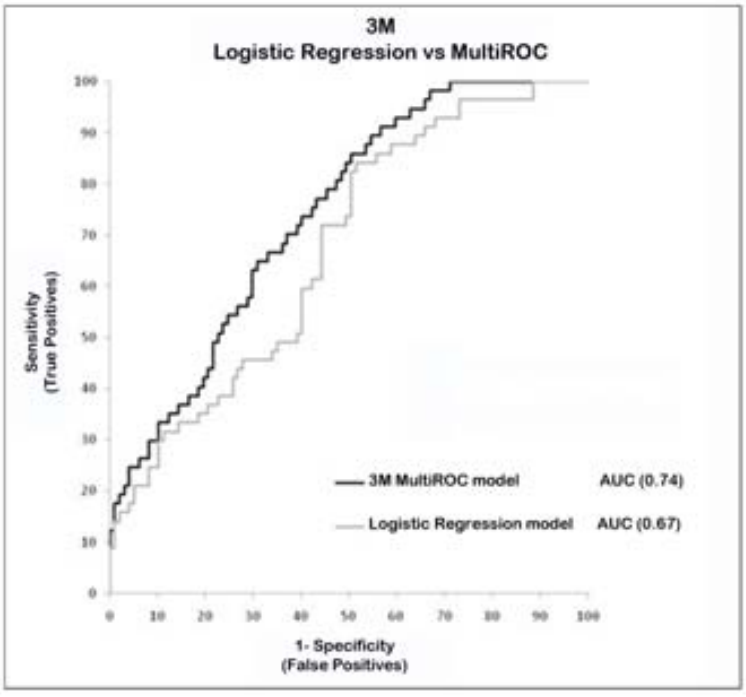
* p-values were adjusted for multiple testing

**Supplemental Figure 1. Characterization of urine based biomarker of prostate cancer (box-and-whisker plots).**




(*) Men within PSA 4-10ng/mL and no prior biospy

**Supplemental Figure 2. Comparison of logistic regression model with mROC model 3M (PSMAvPSGRvPCA3) in men with prostate cancer detected and men with no cancer detected at biopsy.**



## OTHER SUPPLEMENTARY DATA

**1. Overall patients:** After multivariate analysis, we chose PCA3, PSGR, PSMA, and PSAD for ROC-curve analysis to visualize the diagnostic efficacy and to summarize the data of the gene-based qPCR assay on the urine samples.

We obtained the following AUC values: PCA3 (0.61), PSGR (0.64), PSMA (0.63) and PSAD (0.61) (Supplemental Figure 1). To determine if the combination of various biomarkers could improve performance over single biomarkers, we first combined the 3 genes (PCA3, PSGR and PSMA). This combination was then recombined with PSAD to a new marker. The AUC for the combined marker model was 0.80 (0.74 without PSAD), having a sensitivity of 96% and a specificity of 40%. The positive and negative predictive values were 48% and 95%, respectively (Supplemental Figure 3).

**2. PSA gray zone (PSA between 4-10 ng/mL and with no prior biopsy):** The AUC was PCA3 (0.61), PSGR (0.67), PSMA (0.75) and PSAD (0.63). The AUC for the combined marker model was 0.89 (0.82 without PSAD), having a sensitivity of 96%, and specificity 62%. The positive and negative predictive values were 57% and 97%, respectively (Supplemental Figure 4).

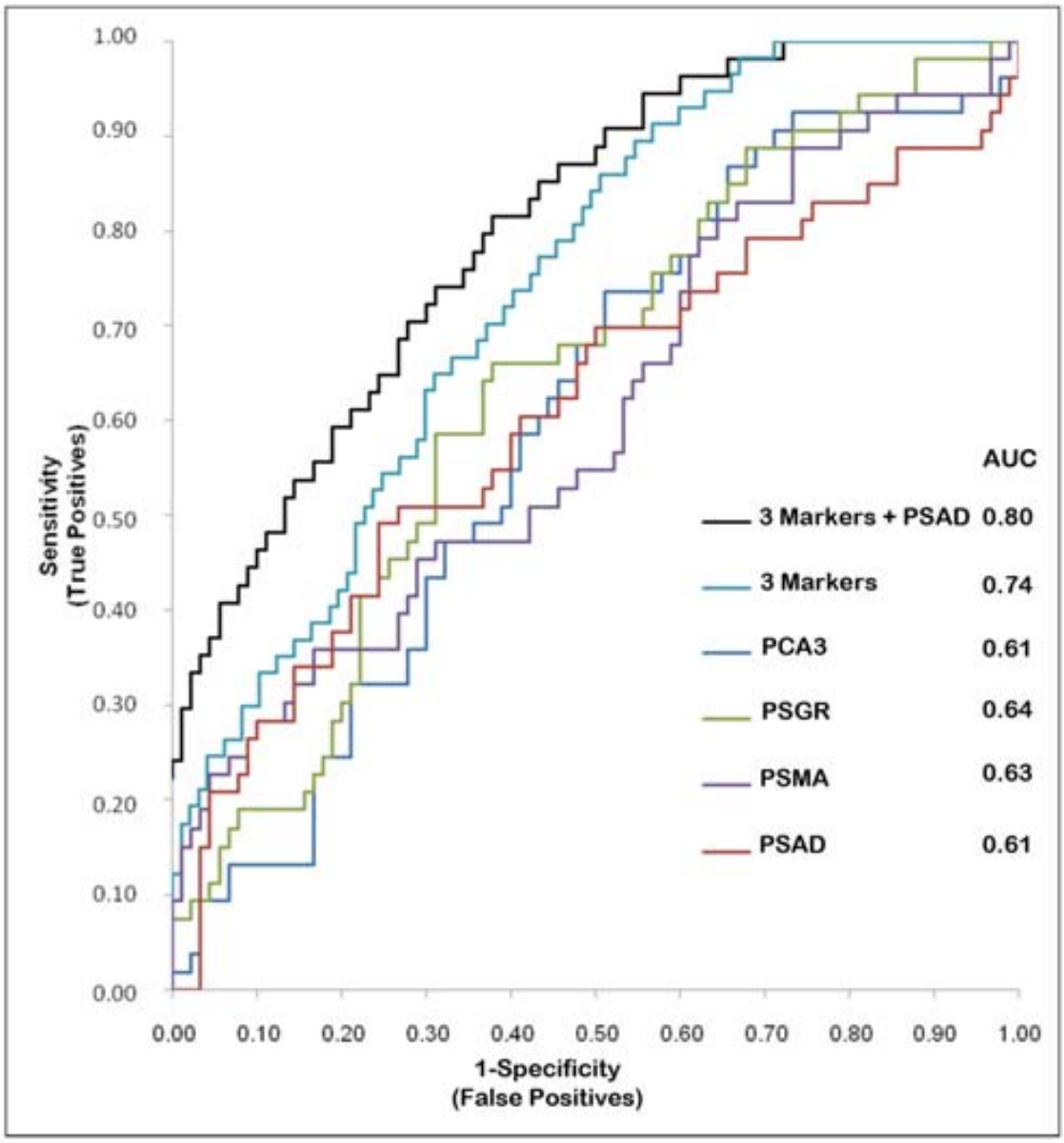
**3. Saved Biopsies:** We calculated the number of biopsies that could be avoided by using the combined marker model on the overall group and on the clinical risk group of special interest as avoided biopsies=true negatives+false negatives. With a sensitivity of 96%, 26% of the biopsies could be avoided in the overall group, including the complete range of PSA and repeated biopsies. In the analysis of the clinical risk group of special interest in the “gray zone” of PSA between 4 and 10 ng/mL and with no prior biopsy, 42% of the biopsies could be avoided with a sensitivity of 96% (Supplemental Figure 5).

**Supplemental Figure 3. Receiver-operating-characteristic (ROC) curves for the combined biomarker in men with prostate cancer detected and men with no cancer detected at biopsy (Overall group).** ROC-curve for PCA3 (dark blue), ROC-curve for PSGR (green) and ROC-curve for PSMA (purple), and PSAD (red) Combined 3 Gene-ROC (light blue). Combined 3 Gene+PSAD-ROC curve (black).

**Supplemental Figure 4. Receiver-operating-characteristic (ROC) curves comparing individual markers and a combined biomarker (PCA3, PSGR, PSMA + PSAD) in men with PSA 4-10 ng/mL and no previous biopsy**. ROC curve for PCA3 (dark blue), ROC curve for PSGR (green) and ROC curve for PSMA (purple), and PSAD (red) Combined 3 Gene-ROC (light blue). Combined 3 Gene+PSAD-ROC curve (black).

**Supplemental Figure 5. Biopsies that could be saved comparing all patients from the study with a clinical risk group of special interest (PSA in the range 4-10 ng/mL and no previous biopsies).** Biopsies saved (%) = (True negatives + False negatives) / All patients. Sensitivity = True positives / (True positives + False negatives).

**Supplemental Figure 3. Overall Group**



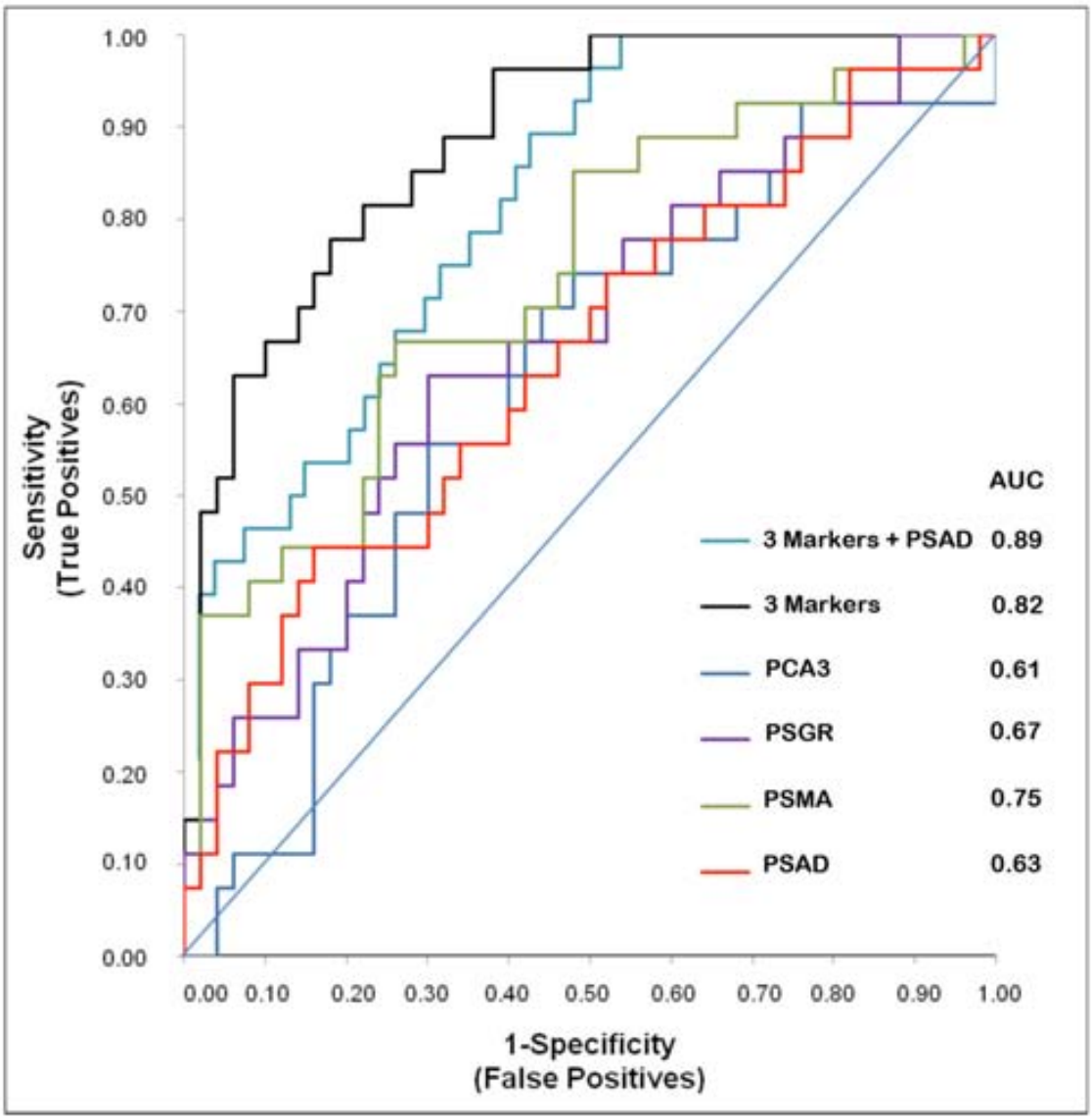

Area Under the Curve (AUC)

| Test Result Variable(s) | Area | Std. Error[a] | Asymptotic Sig.[b] | Asymptotic 95% Confidence Interval | |
|---|---|---|---|---|---|
| | | | | Lower Bound | Upper Bound |
| PCA3 | 0.605 | 0.046 | 0.031 | 0.514 | 0.696 |
| PSGR | 0.642 | 0.045 | 0.003 | 0.553 | 0.731 |
| PSMA | 0.625 | 0.047 | 0.010 | 0.533 | 0.718 |
| PSAD | 0.609 | 0.051 | 0.030 | 0.509 | 0.709 |

a. Under the nonparametric assumption.

b. Null hypothesis: true area = 0.5

**Supplemental Figure 4. Serum PSA (4-10 ng/mL) and no previous bipsoy**



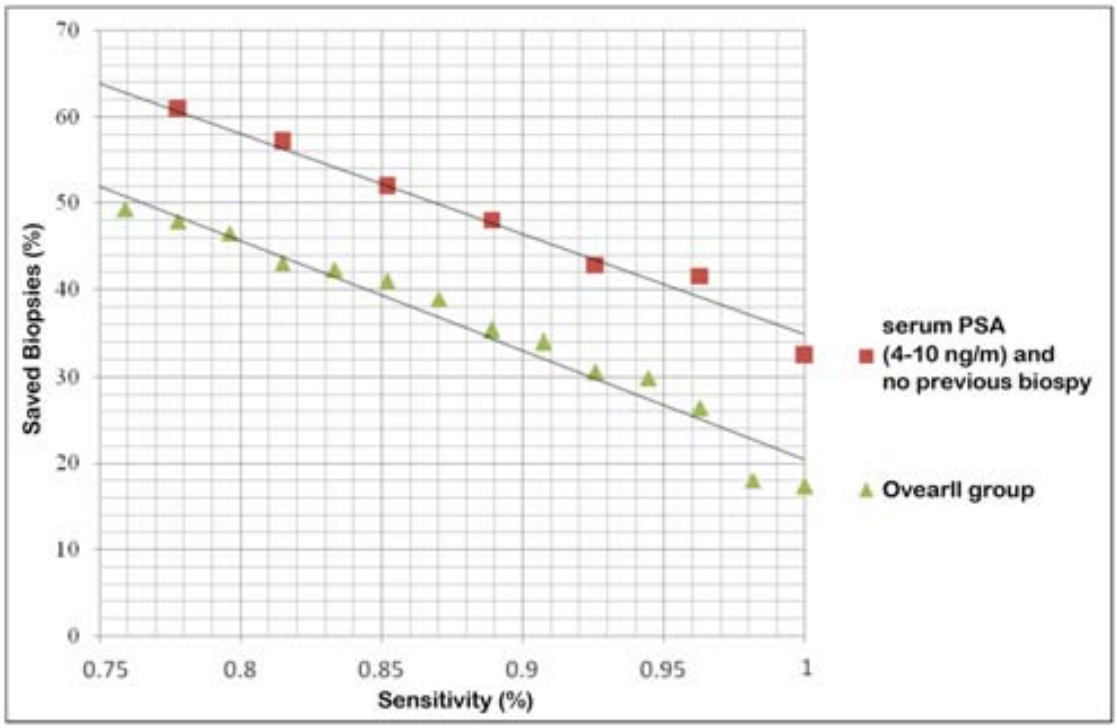

Area Under the Curve (AUC)

| Test Result Variable(s) | Area | Std. Error[a] | Asymptotic Sig.[b] | Asymptotic 95% Confidence Interval | |
|---|---|---|---|---|---|
| | | | | Lower Bound | Upper Bound |
| PCA3 | 0.605 | 0.065 | 0.117 | 0.477 | 0.733 |
| PSGR | 0.673 | 0.062 | 0.010 | 0.551 | 0.795 |
| PSMA | 0.746 | 0.058 | 0.000 | 0.633 | 0.860 |
| PSAD | 0.625 | 0.067 | 0.069 | 0.494 | 0.756 |

a. Under the nonparametric assumption

b. Null hypothesis: true area = 0.5

**Supplemental Figure 5. Biopsies Saved**



World J Urol
DOI 10.1007/s00345-010-0580-0

ORIGINAL ARTICLE

# Behavior of the PCA3 gene in the urine of men with high grade prostatic intraepithelial neoplasia

**Juan Morote · Marina Rigau · Marta Garcia · Carmen Mir · Carlos Ballesteros · Jacques Planas · Carles X. Raventós · José Placer · Inés M. de Torres · Jaume Reventós · Andreas Doll**

Received: 8 March 2010 / Accepted: 25 June 2010


**Abstract**
*Objective* An ideal marker for the early detection of prostate cancer (PCa) should also differentiate between men with isolated high grade prostatic intraepithelial neoplasia (HGPIN) and those with PCa. Prostate Cancer Gene 3 (PCA3) is a highly specific PCa gene and its score, in relation to the PSA gene in post-prostate massage urine (PMU-PCA3), seems to be useful in ruling out PCa, especially after a negative prostate biopsy. Because PCA3 is also expressed in the HGPIN lesion, the aim of this study was to determine the efficacy of PMU-PCA3 scores for ruling out PCa in men with previous HGPIN.
*Patients and methods* The PMU-PCA3 score was assessed by quantitative PCR (multiplex research assay) in 244 men subjected to prostate biopsy: 64 men with an isolated HGPIN (no cancer detected after two or more repeated biopsies), 83 men with PCa and 97 men with benign pathology findings (BP: no PCa, HGPIN or ASAP).
*Results* The median PMU-PCA3 score was 1.56 in men with BP, 2.01 in men with HGPIN ($p = 0.128$) and 9.06 in men with PCa ($p = 0.008$). The AUC in the ROC analysis was 0.705 in the subset of men with BP and PCa, while it decreased to 0.629 when only men with isolated HGPIN and PCa were included in the analysis. Fixing the sensitivity of the PMU-PCA3 score at 90%, its specificity was 79% in men with BP and 69% in men with isolated HGPIN.
*Conclusions* The efficacy of the PMU-PCA3 score to rule out PCa in men with HGPIN is lower than in men with BP.

**Keywords** PCA3 · Prostate cancer · High grade prostate intraepithelial neoplasia

## Introduction

An ideal marker for the early detection of prostate cancer (PCa) should also be able to differentiate between men with an isolated high grade prostatic intraepithelial neoplasia (HGPIN) and those with associated PCa [1]. Serum prostatic-specific antigen (PSA) levels have been widely used for diagnostic purposes for more than 25 years. However, unnecessary biopsies are frequently performed as a consequence of the high rate of false-positive results of PSA [2]. The low specificity of PSA has been the main reason for the development of intensive research projects to find new markers for PCa during the last decade.

Prostate Cancer Gen 3 (PCA3) was described by Bussemakers et al. [3] in 1999 as a PCa-specific mRNA that belongs to the class of non-coding RNAs whose biological role remains unknown. PCA3 was soon regarded as a potential diagnostic tool for the detection of PCa cells in tissue biopsies and bodily fluids [4]. The expression ratio between PCA3 mRNA and PSA mRNA, as the control transcript, in post-prostate massage urine (PMU-PCA3 score) was proposed as a useful tool, in addition to the

J. Morote, M. Rigau, J. Reventós and A. Doll contributed equally to this work.

J. Morote (✉) · C. Mir · C. Ballesteros · J. Planas · C. X. Raventós · J. Placer
Department of Urology, Vall d'Hebron University Hospital, Pg. Vall d'Hebron 119-129, 08035 Barcelona, Spain
e-mail: jmorote@vhebron.net

J. Morote · M. Rigau · M. Garcia · J. Reventós · A. Doll
Biomedical Research Unit, Research Institute Vall d'Hebron University Hospital, Barcelona, Spain

I. M. de Torres
Department of Pathology, Vall d'Hebron University Hospital, Barcelona, Spain

Published online: 07 July 2010

current methods utilized for the early detection of PCa [5–7]. Recently, the APTIMA® assay (Gen-Probe, San Diego, CA, USA; PROGENSA™ for European countries) resolved the problem of urine mRNA conservation and also provided a standardized technique [8]. After establishing the usefulness of the PMU-PCA3 score in ruling out PCa in men undergoing repeat biopsies [9], the basis for its usage was established through recent studies [10–13].

The biopsy finding of an HGPIN is a frequent indication for repeating the prostate biopsy [1]. However, the utility of the PMU-PCA3 score for ruling out PCa in men with HGPIN has not been analyzed. Today we know that PCA3 is expressed in the HGPIN lesions surrounding PCa [14] and that the PMU-PCA3 score seems to be higher in men with HGPIN [12]. Therefore, discovering the behavior of the PMU-PCA3 score in men with HGPIN seems to be a priority. In this study, we analyzed whether the PMU-PCA3 score was able to rule out PCa in men with HGPIN.

## Materials and methods

### Study design and patients

This is a case–control study where the PMU-PCA3 score was determined before prostate biopsy in 244 men. The study group was represented by 64 men with isolated HGPIN. The condition of isolated HGPIN was defined as no PCa detection after one to four repeat biopsies (1 in 32 men, 2 in 18 men, 3 in 11 men and 4 in 3 men). The control group was represented by 180 men: 83 men with PCa and 97 men with benign pathology findings (BP: no PCa, HGPIN or ASAP). Ten of the 83 men with PCa had previous HGPIN. The men with HGPIN or ASAP in the first biopsy were not included in this study. The median age of the overall group was 64 years (39–85), and the median serum PSA level was 6.4 ng/mL (1.5–189). Prostate massage was performed by systematically applying severe digital pressure to the prostate from the base to the apex and from the lateral to the median line of each lobe. Biopsies were performed using an end-fire ultrasound transducer (Falcon 2101, B-K Medical Inc.) and an automatic 18 gauge needle (Bard Inc.). The minimum number of cores removed in every procedure was 10, plus 1–8 additional cores removed according to the prostate volume and age. Written informed consent was obtained from all the study participants. The protocol was approved by the Institutional Ethical Research Committee.

### Sample preparation and procedure

Samples of PMU (~50 mL first capture) were collected in urine collection cups, kept on ice, transported to the laboratory and processed within 30 min. The urine samples were centrifuged at 2,500×*g* for 10 min at 4°C, and then the pellets were washed twice with cold PBS 1×. Finally, the pellets were stored with 1:5 RNA Later (Ambion, Foster City, CA) at −80°C until RNA extraction. Urine RNA was extracted with the QIAamp® Viral RNA Mini Kit (Qiagen, Hilden, Germany). Single-stranded cDNA synthesis was carried out using the SuperScript III Reverse Transcriptase (Invitrogen, Carlsbad, CA) and stored at −80°C until preamplification with the TaqMan Preamp Master Mix Kit (Applied Biosystems, Foster City, CA).

### Quantitative PCR analysis

PCA3 and the control transcript PSA, all from the TaqMan® Gene Expression Assay (Applied Biosystems, Foster City, CA), had to be analyzed by quantitative real time PCR (qPCR). Reactions were carried out in triplicate on an ABI-Prism-7900 qPCR machine, and only those results with a standard deviation <0.38 value were accepted (as recommended by the manufacturer). Threshold levels were set into the exponential phase of the qPCR. The data analysis was carried out using the ABI-Prism-7900SDS Software V2.3 (Applied Biosystems, Foster City, CA) with the same baseline and threshold set for each plate to generate threshold cycle ($C_t$) values for all of the genes in each sample. Since only a relatively small number of prostate cells are to be found in urine, we performed a cDNA preamplification step before the qPCR. The score was calculated as ctPCA3/ctPSA'1000 [8].

### Statistical analysis

All PMU-PCA3 scores were carried out on log-transformed data, which were applied to stabilize the variances. Quantitative variables were expressed in medians and semi-interquartile range (SIR) and qualitative variables in rates. Univariate analysis was carried out using the Mann–Whitney *U* and the Kruskall–Wallis test to compare quantitative variables according to two or more groups. Receiver-operating-characteristic (ROC) curves and the area under the curve (AUC) were used to assess the efficacy of the PMU-PCA3 score and to compare groups. Statistical analyses were performed using the statistical package SPSS v.15.

## Results

The medians (SIR) of the PMU-PCA3 scores were 1.56 (3.62) in men with BP, 2.01 (11.7) in men with IHGPIN and 9.06 (22.35) in men with PCa. A significant difference was observed among the three scores, $p < 0.001$. A significant difference was observed between the PMU-PCA3

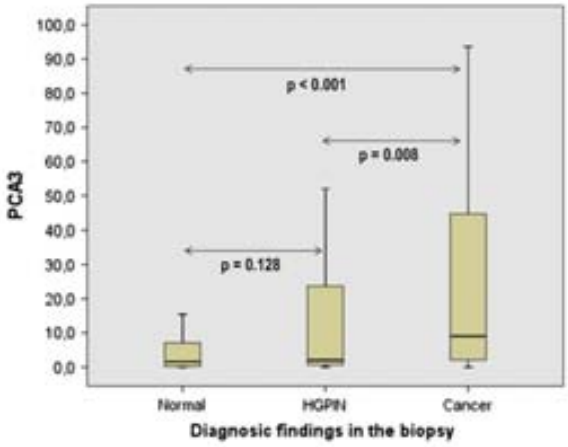



**Fig. 1** PMU-PCA3 score according to the biopsy diagnosis. *p* values represented the comparative analysis between every two groups

score of men with IHGPIN and PCa, $p = 0.008$; however, no differences were observed between the PMU-PCA3 score of men with HGPIN and BP, $p = 0.128$ (Fig. 1).

ROC curves for the two different populations were generated. The AUC was 0.629 in the subset of men with IHGPIN and PCa, while it was 0.705 in the subset of men with BP (without HGPIN) and PCa (Fig. 2). Setting the sensitivity for the PMU-PCA3 score at 90%, the specificity was 79% in men with BP, while it decreased to 67% in men with IHGPIN.

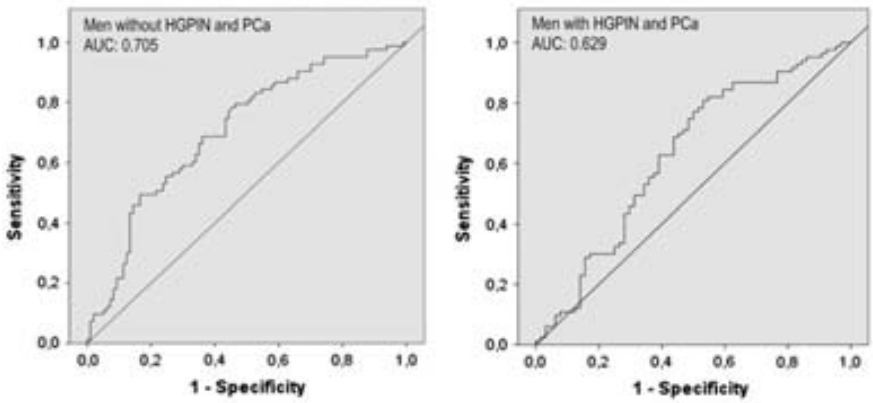
## Discussion

The finding of an HGPIN lesion in a prostate biopsy is a frequent cause for repetition of the procedure. Although some predictors of PCa in men with HGPIN have been analyzed, the urological community still wonders if and when a new biopsy may be required. We have recently described how, in contrast to isolated HGPIN lesions, the PTOV1 protein is over-expressed in HGPIN lesions associated with PCa. Therefore, we have suggested that PTOV1 immunostaining in HGPIN lesions could identify those men requiring a repeat biopsy, due to the high probability of associated PCa [1]. At present, we are involved in a validation study to confirm these results.

The PMU-PCA3 score has emerged as a tool for selecting those men with negative prostate biopsy, who require a repeat procedure. However, no study has been conducted to analyze the behavior of the PMU-PCA3 score in men with HGPIN. Today we know that PCA3 is expressed in the HGPIN lesions surrounding PCa. Popa et al. [14] recently conducted a study by means of in situ hybridization in radical prostatectomy specimens, in which PCA3 was also expressed in HGPIN lesions. Unfortunately, nothing is known about the PCA3 expression in isolated HGPIN lesions. Therefore, the efficacy of the PMU-PCA3 score in ruling out PCa in men with HGPIN should be defined. Till date, all the PMU-PCA3 score studies have considered a diagnosis of HGPIN as a negative result for PCa. Unfortunately, the recent biopsy nomogram, based on the worldwide data of 809 men subjected to prostate biopsy in a multi-institutional study using the Aptima platform to determine the PMU-PCA3 score, did not include HGPIN as a predictive variable [10].

We decided to review the data from our genomic and proteomic research project conducted on PMU, in order to better determine the behavior of the PMU-PCA3 score in men with HGPIN. We designed a case–control study with a study group represented by those men with a high probability of having an isolated HGPIN, since PCa had not been detected after one or more repeat biopsies. As a control group, we selected men with PCa (true positives) and men with benign pathology findings (true negatives). We observed that the PMU-PCA3 score of men with PCa was significantly higher than that observed in men with BP or IHGPIN, while the PMU-PCA3 score was only slightly

**Fig. 2** Efficacy of PMU-PCA3 score to predict PCa by ROC analysis in the subset of men with and without previously diagnosed HGPIN



higher in men with IHGPIN than it was in those men with BP. After the above observation, we wondered whether the efficacy of the PMU-PCA3 score for ruling out PCa in men with IHGPIN was similar to that observed in men with benign pathology. The efficacy of the PMU-PCA3 score was significantly lower in the subset of men having a previous HGPIN. The specificity of the PMU-PCA3 score decreased from 79 to 67% at 90% sensitivity. Therefore, we believe that previous HGPIN should be taken into account, in order to establish the usefulness of the PMU-PCA3 score as a tool for avoiding repeat biopsies. Perhaps a specific cut-off level for the PMU-PCA3 score should established for men with previous HGPIN, or perhaps a previous HGPIN should be taken into account as a predictive variable in the biopsy nomogram.

Our study had two main limitations. First, negative repeat biopsies cannot exclude a small probability of PCa in men with previous HGPIN. Second, the technique used to assess the PMU-PCA3 score in this study was not the same as the one used in the Aptima platform. For our PMU genomic project, our laboratory has developed a multiplex PCR technique for the analysis of PMU-multiple gene scores. Nevertheless, the present study has demonstrated the presence of cells expressing PCA3 in the PMU of men with isolated HGPIN.

In summary, the main message of this study is that the finding of an HGPIN in a prostate biopsy should be taken into consideration, in order to establish the clinical usefulness of the PMU-PCA3 score as a tool for avoiding unnecessary repeat biopsies.

## Conclusions

The efficacy of the PMU-PCA3 score in ruling out PCa in men with previous HGPIN seems to be lower than that observed in men with a previous negative biopsy. This finding should be taken into account, in order to establish appropriate thresholds for the PMU-PCA3 score. Moreover, biopsy nomograms, based on the PCA3 score, should take previous HGPIN into consideration as a predictive variable.

**Acknowledgments** This work was supported by Spanish Ministry of Health (RTICC RD06/0020/0058). Fundación para la Investigación en Urología (FIU) 125/2006. Recerca i de Societat de la Informacio de la Generalitat de Catalunya (Grants SGR00231, 00391), the Instituto de Salud Carlos III (PS09/00496, Red de Genomica del Cancer y Genotipado de tumores C03/10) Marina Rigau is the recipient of a predoctoral fellowship from the Fundación Francisco Cobos.

**Conflict of interest statement** The authors declare that they have no conflict of interest.

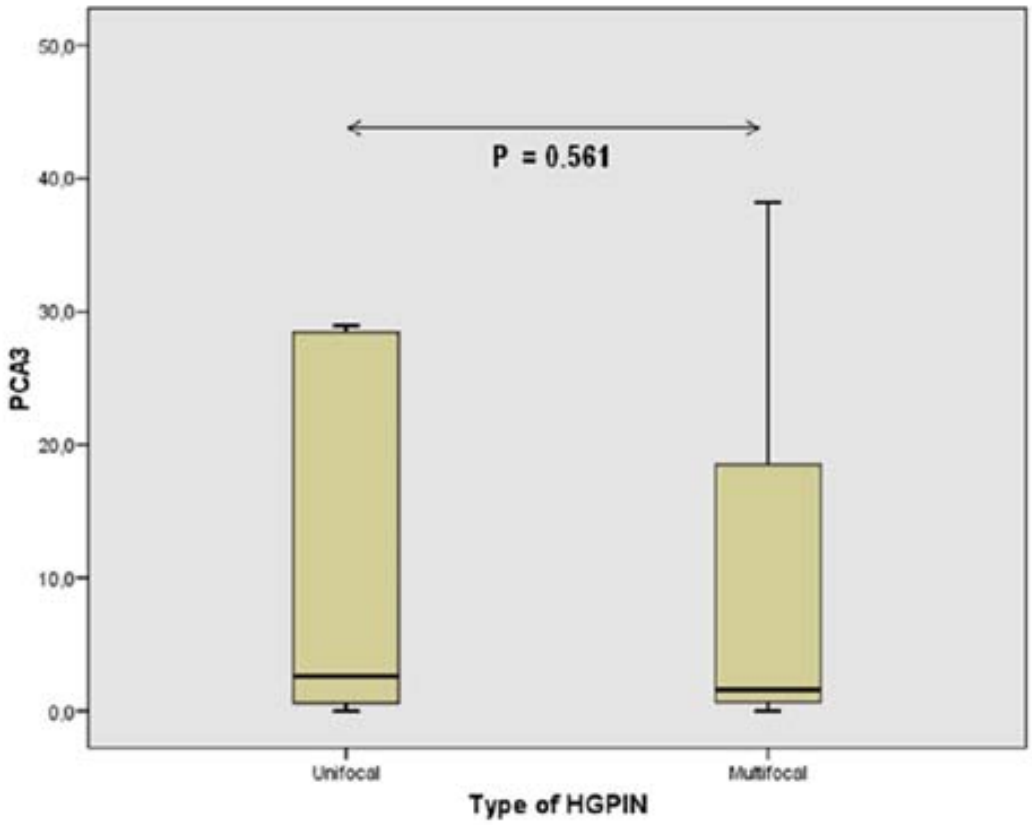
## References

1. Morote J, Fernández S, Alaña L et al (2008) PTOV1 expression predicts prostate cancer in men with isolated high-grade prostatic intraepithelial neoplasia in needle biopsy. Clin Cancer Res 14:2617–2622
2. Freedland SJ, Partin AW (2006) Prostate-specific antigen: update 2006. Urology 67:458–460
3. Bussemakers MJ, van Bokhoven A, Verhaegh GW et al (1999) DD3: a new prostate-specific gene, highly overexpressed in prostate cancer. Cancer Res 59:5975–5979
4. de Kok JB, Verhaegh GW, Roelofs RW et al (2002) DD3(PCA3), a very sensitive and specific marker to detect prostate tumors. Cancer Res 62:2695–2698
5. Hessels D, Klein Gunnewiek JM, van Oort I et al (2003) DD3(PCA3)-based molecular urine analysis for the diagnosis of prostate cancer. Eur Urol 44:8–15
6. Fradet Y, Saad F, Aprikian A et al (2004) UPM3, a new molecular urine test for the detection of prostate cancer. Urology 64:311–315
7. Tinzl M, Marberger M, Horvath S, Chypre C (2004) DD3PCA3 RNA analysis in urine—a new perspective for detecting prostate cancer. Eur Urol 46:182–186
8. Groskopf J, Aubin SM, Deras IL et al (2006) APTIMA PCA3 molecular urine test: development of a method to aid in the diagnosis of prostate cancer. Clin Chem 52:1089–1095
9. Marks LS, Fradet Y, Deras IL et al (2007) PCA3 molecular urine assay for prostate cancer in men undergoing repeat biopsy. Urology 69:532–535
10. Chun FK, de la Taille A, van Poppel H et al (2009) Prostate Cancer Gene 3 (PCA3): development and internal validation of a novel biopsy nomogram. Eur Urol 59:659–668
11. Deras IL, Aubin SM, Blase A et al (2008) PCA3: a molecular urine assay for predicting prostate biopsy outcome. J Urol 179:1587–1592
12. Haese A, de la Taille A, van Poppel H et al (2008) Clinical utility of the PCA3 urine assay in European men scheduled for repeat biopsy. Eur Urol 54:1081–1088
13. van Gils MP, Hessels D, van Hooij O et al (2007) The time-resolved fluorescence-based PCA3 test on urinary sediments after digital rectal examination; a Dutch multicenter validation of the diagnostic performance. Clin Cancer Res 13:939–943
14. Popa I, Fradet Y, Beaudry G, Hovington H, Tetu B (2007) Identification of PCA3 (DD3) in prostatic carcinoma by in situ hybridization. Mod Pathol 20:1121–1127

**SUPPLEMENTAL DATA "Behaviour of PCA3 gene in the urine of men with high grade prostatic intraepithelial neoplasia"**

The medians and (semi-inter-quartile-ranges) of the PCA3 scores were 2.63 (14.10) in men with previously diagnosed unifocal HGPIN and 1.59 (9.85) in men with previously diagnosed multifocal HGPIN, p=0.561.

**Supplemental Figure1. Urine expression of PCA3 according the grade of HGPIN.** Post-PM urine PCA3 scores between unifocal and multifocal HGPIN. P value represented the comparative analysis between those groups.



## The Discovery and Qualification of a Panel of Urine Biomarkers for Prostate Cancer Diagnosis

The use of biomarkers is a main focus in the detection of presymptomatic prostate cancer (PCa). However, currently available cancer biomarkers have the disadvantage of lacking specificity and/or sensitivity. Proteins secreted by cancer cells, also referred to as "cancer cell secretomes", are a promising source for biomarker discovery. A great advantage is that these cancer-secreted proteins and/or their fragments always enter body fluids, such as blood or urine, and can be measured via non-invasive assays. Since the protein products of PCa cells, as well as of normal prostate cells, can be detected in urine, their use as a proximal body fluid in the detection of PCa is very attractive. In fact, they may represent the best compromise between a minimally invasive diagnostic technique and the possibility of obtaining enough material for a correct diagnosis. In the present study we used DIGE proteomic analysis on 30 age-matched, post-prostate massage urine specimens, in order to identify the differentially expressed proteins in patients with PCa. 24 potential biomarkers were identified, the majority of which were secreted proteins associated with several well-known, functional cancer pathways, such as NFκB, PDGFBβ, or β-catenin. Qualification of 15 of the 24 identified biomarker candidates was then undertaken by relative quantification using an SRM-based assay on 50 urine samples obtained after PM. 19 of the 50 samples (38%) presented PCa. Using logistic regression analysis, 7 peptides corresponding to 5 different proteins were selected. A multiplex ROC curve using those 7 peptides showed an AUC value of 0.93. Fixing the sensitivity at 95%, we achieved a specificity of 78%. Our data demonstrated that proteomic analysis is able to reveal novel biomarkers/diagnostic profiles for PCa in urine. This constitutes an important step towards advancing the accurate diagnosis of PCa, which currently represents a major setback in our ability to cure patients who suffer from the disease.

## INTRODUCTION

Prostate cancer (PCa) has emerged as the most common neoplasia among European and American men (24.1% of all cases), followed by lung cancer (15.5%) and colorectal cancer (13%) (1). In spite of the fact that the introduction of the prostate specific antigen (PSA) test in the late 80's of the last century has led to a spectacular increase in the detection of PCa, the rate of mortality has not significantly decreased in the last few decades (2). Currently, prostatic biopsy (PB) is the only diagnostic method for the detection of PCa, and its indication is based on elevated serum levels of PSA and/or on alterations in the digital rectal exam (DRE). The main limitation of serum PSA as a tumor marker is its lack of specificity, which results in a high negative biopsy rate. As a consequence of the current screening parameters, around 2/3 of the approximately 1,300,000 biopsies made yearly in the United States (390,000 in Europe) are negative and, therefore, unnecessary (3), (4). Novel methods are needed to reduce the number of unnecessary biopsies, while still maintaining a high rate of cancer detection. These new methods will make the diagnosis of PCa more effective and be able to predict its behavior with better reliability than the present criteria that define risk.

Proteins and peptides "leaked" from tissues into clinically accessible body fluids, such as blood or urine, have led to the possibility of diagnosing the disease at an early, presymptomatic stage, as well as being able to monitor its responses to therapy through testing these body fluids for the presence of disease-related biomarkers (5). The term, "secretome," refers to the rich, complex set of molecules secreted or shed from the surface of living cells (6). Cancer cell secretomes are responsible for the "cross-talk" among cancer cells (7). Understanding this language could greatly increase our knowledge regarding the molecular mechanisms of neoplasia. Likewise, since the secreted products from both normal prostate epithelial cells and PCa cells can be detected in the urine of men, their use in the detection of PCa as proximal body fluids, which have been in contact with only one or a few tissues, is very attractive. These molecules could serve as the best compromise between a minimally invasive technique and the possibility of obtaining enough material for a correct diagnosis.

Urine has evolved as one of the most attractive body fluids in clinical proteomics with a potentially rapid application in the clinics (8). Extensive analysis by proteomics has revealed that normal urine consists of more than 1,500 proteins (9). Approximately 30% of the proteins found in normal human urine are plasma proteins, while the other 70% are proteins which are excreted from mechanisms other than plasma ultra-filtration or which have different origins, such as the bladder or prostate (directly from prostatic

secretions that leave the prostate ducts or from proteins originating in the prostate that leave the vasculature and are not later ultra-filtered by the kidneys) (10-12). Exploration of the urine proteome has certain advantages over other body fluids, such as the plasma proteome. It has become an attractive target for clinical research, due to the availability of samples in large quantities, the non-invasive nature of collection methods, the higher thermodynamic stability of the samples and the possibility of repeat sampling (8). Moreover, urine samples represent more proximal fluids to PCa than plasma, since proteins of prostatic origin are secreted directly into the urine without passing through kidney ultra-filtration (13).

The main practical challenges in the detection and quantification of proteins in urine by proteomic techniques are the high levels of salt and the wide dynamic range of protein concentrations (14). The very abundant proteins and their degradation products hamper the reliable detection of the low-abundance components (15), which are frequent targets in biomarker development. Furthermore, since urine contains different, highly abundant proteins and the dynamic range is not as high as in plasma, commercial plasma pretreatment depletion kits are not useful. Also, the protein content in urine is about a factor 1,000 less compared to plasma (9), and in consequence, it requires approximately a 1000-fold concentration. Nevertheless, approaches to biomarker discovery in urine have been hindered by concerns for reproducibility and an inadequate standardization of proteomics protocols, due to intra and inter-individual variability (16). Reproducible procedures for the preparation of protein samples isolated from human urine are essential for meaningful proteomic analyses (17).

In the past few years, proteomic studies on body fluids have been developed (13, 18, 19). Advances in liquid chromatography (LC) and two dimensional gel electrophoresis (2D-GE), in combination with mass spectrometry (MS), have significantly facilitated the challenging detection of secreted proteins (20). Several research teams have worked on profiling the healthy human urine proteome using electrophoresis or LC, followed by MS (11, 21, 22). In addition, the combination of 2D-GE followed by matrix-assisted laser desorption/ionization MS tandem time-of-flight (MALDI-TOF-TOF MS) has represented one of the most utilized strategies for protein-biomarker identification in differential proteomic analyses (23). This technique is especially adequate for the direct comparison of protein expression and has been used to find patterns of expression associated with cancer in several tissues, such as liver, bladder, lung, esophagus, breast and prostate. Moreover, 2D-GE is the most accessible technique, allowing the study of large molecules and enabling an estimate of the actual molecular weight of large molecules. The identification of biomarkers by means of mass spectrometry methods is easy to

perform from 2D-GE spots and has been used for the description of the urinary proteome (24). Although 2D-GE has been used in laboratory practice since 1979 (25), there have only been a limited number of studies that have focused on the potential power of using urine as a proximal body fluid. The main reason for the low number of successful 2D-GE studies is the low reproducibility rate between the various gels and the laboratories (26). One possibility for overcoming this weak point could be the use of 2D fluorescence-based Differential Gel Electrophoresis (2D-DIGE) technology, which, using fluorescent dyes and internal standards, provides better reproducibility and more accurate quantification (27). The combination of 2D-DIGE followed by MS has been used for the discovery of new biomarkers in urine for different diseases (28-31). There are several studies on PCa that utilize 2D-DIGE for the protein profiling of markers associated with lymph node metastasis (32), as well as the identification of serum markers for the progression of PCa (33) and the elucidation of intracellular signaling pathways that regulate the radio-resistance of human PCa cells (34). In the case of PCa, only a few studies exist, which have identified new proteomic markers in urine using MS (35-37). However, to our knowledge, no study based on 2D-DIGE analysis of urine on PCa has yet been published.

Proteomics profiling is a promising approach for discovering the biomarkers utilized in secretome analysis (6, 7). The evaluation of candidate protein profiles is a crucial phase of the new biomarker development pipeline. There are generally four different phases in the discovery of new biomarkers: the discovery phase, the qualification phase, the verification phase, and the validation phase (38). Discovery stage can employ different model systems or a variety of human samples, such as body fluids, and usually comprises a binary comparison between the diseased and controlled states (38). The product of the discovery phase is a list of "candidate biomarkers" with a high false discovery rate of differentially expressed proteins, at least for lower abundant proteins (38). The next step, "qualification," is used to confirm that the differential candidate can be seen using an alternative target method and to confirm differential expression in simplified comparisons between the diseases and normal states (38). In the verification stage, the analysis is extended to a larger number of samples. Proteins that have successfully passed the first 3 stages are considered biomarkers of high value, which are translated into large-scale, clinical validation studies (39).

For the qualification and verification phases, selected reaction monitoring (SRM) coupled with stable isotope dilution (SID) MS represents an effective strategy for a targeted, quantitative proteomics approach (40). SRM enables the rapid screening of hundreds of analytes in one LC-MS/MS run and allows the detection and accurate quantification of

predetermined protein sets in biological samples (40). For each protein, a set of "best representing" peptides is determined on the basis of their uniqueness (proteotypic peptides (PTP)) and performance in LC–MS analysis (41). This technique allows the quantitative analysis of peptides with a linear response over near five orders of magnitude down to a low ng/mL concentration (40). In addition, the ability to multiplex SRM experiments increasingly expedites the development of clinical biomarker studies (14). Normally, in biomarker studies it is critical to make comparisons between samples taken from healthy and diseased patients, in order to identify the peptides (proteins) that show specific differences in expression between these two groups. When SRM is not coupled with SID a relative quantification comparing each targeted peptide in healthy and disease samples to detect and qualify potential biomarkers can be easily performed. The determination of relative changes in protein concentrations is often the first stage of a proteomics study. Such analyses rely on measuring peptide ions in individual samples based on their absolute signal intensity after proper normalization, often referred to as label-free method (42).

The hypothesis of this study is that PCa-secreted proteins or their fragments enter directly into urine and can be measured via non-invasive assays. Thus, urine may reflect a broad variety of pathological conditions of the prostate and may represent a more reliable source of biomarkers than serum or other, more distal body fluids. To date, only a limited number of studies have analyzed cancer secretomes, and few have used urine; however, the results regarding biomarker discovery are very promising.

## EXPERIMENTAL PROCEDURES

### Patients and urine collection

The biological samples collected from selected patients were the first 50 mL of urine gathered in the first micturition produced after PM. PMs were performed by systematically applying severe pressure to the prostate from the base to apex and from the lateral to the median line of each lobe. It is known that through the practice of this technique, the desquamation of a large quantity of products of tumor origin is induced. These products were present in the urine samples and were able to be gathered in the micturition. All urine samples were obtained from the Department of Urology of the Vall d'Hebron Hospital in Barcelona, Spain, and were taken from patients subjected to PB. Patients with other known tumors and/or previous PCa therapies were excluded from the study. PB achieved the definitive diagnosis of all patients. Biopsies were performed using an end-fire ultrasound transducer (Falcon 2101, B-K Medical, Inc.) and an 18

Gauge disposable automatic biopsy needle (Bard, Inc.). The minimum number of cores removed in every procedure was 10, and between 1 and 8 additional cores were removed, according to the Vienna monogram (43). Clinical and pathological data of all the samples included in the present study are shown in Table 1. Once the samples had been collected in the receptacles provided for that purpose, the receptacles were closed and stored on ice. Samples were then transported directly to the laboratory for processing within 30 min of their collection. Urine samples were centrifuged at 2500g at 4ºC for 10 min. The supernatants were separated and protease inhibitors (Sigma-Aldrich, Saint Louis, MO, USA) were added before storing the samples at –80ºC.

**Discovery phase**

***1. Urine protein extraction for the 2D-DIGE experiment in urine total protein extract:*** We selected 12 urine samples obtained after PM, i.e., 6 histologically confirmed PCa and 6 control samples collected from age-matched patients having a typical background of benign prostate hyperplasia (BPH), atrophy and chronic inflammation (Table 1). Urinary supernatants were precipitated at 4ºC overnight using acetonitrile (1:5). After precipitation, proteins were pelleted by centrifugation at 4,000g for 40 min at 10ºC. The pellets were resuspended in 2mL rehydration buffer (RB), 7M urea, 2M thiourea, and 4% (w/v) CHAPS. The protein solution was centrifuged at 21,000g for 10 min at 10ºC prior to desalting for removal of non-soluble material. The protein solution was then placed on an Ultrafree 15 centrifugal filter with a 5-kDa cutoff membrane (Millipore, Billerica, MA, USA) and diluted with 13mL of RB. The protein solution was desalted and concentrated at 2,000 g at 10ºC until the volume reached 200 $\mu$L, as described elsewhere (44). Protein concentration was determined using the BioRad RC DC Protein Assay (BioRad, California, USA).

***2. Urine protein extraction and enrichment for 2D-DIGE experiment*** **in ProteoMiner enriched urine protein extracts*:*** We selected 18 urine samples obtained after PM, 9 histologically confirmed PCa and 9 control samples (Table 1). 40 mL of each sample was filtered through a 0.45$\mu$m PVDF membrane filter, Millipore (Millipore, Billerica, MA, USA). The urine samples were concentrated to a final volume of 200µL with Amicon 3kDa cutoff ultra-filtration units (Millipore, Billerica, MA, USA). Then, to remove any inorganic salts the sample is dialyzed using Sodium phosphate 25mM pH 7 (1:10) (45). Protein concentration was determined using the BioRad RC DC Protein Assay (BioRad, California, USA). Then, the most abundant proteins were removed using ProteoMiner (BioRad, California, USA), according to the manufacturer's protocol. In brief, two micro Bio-Spin columns (BioRad, California, USA) were loaded with a ratio of 1mg protein

10µL resin. Because there are a limited number of binding sites for each protein, the high-abundance proteins (HAP) quickly reached bead capacity, and the excess HAPs are passed through the column. After sample binding and washing, bound proteins were eluted from the column with 200µL of elution buffer (5% acetic acid, 8M urea, and 2% CHAPS). Samples were stored at -20ºC until use.

*Two Dimensional Differential in Gel Electrophoresis (2D-DIGE):* Following extraction, any other interfering components were removed by a modified TCA-acetone precipitation (2D-CleanUp kit, Amersham Biosciences, Piscataway, NJ, USA) and were dissolved in DIGE lysis buffer (7M urea, 2M thiourea, 30mM Tris-HCl pH 8.4% (w/v) CHAPS). Protein concentration was determined again using the BioRad RC DC Protein Assay (BioRad, California, USA) to ensure that there was a sufficient quantity for the performance of a DIGE experiment (lower limit 75 $\mu$g). Finally, the pH was adjusted to 8.5 for DIGE labeling.

A pool, which consisted of equal amounts of each of the samples analyzed in the DIGE experiment, was prepared to be used as an internal standard for quantitative comparisons (46). Urine samples were labeled with Cy3 or Cy5 cyanine dyes, while the internal standard pooled samples were labeled with Cy2 dye. To avoid any possible bias introduced by labeling efficiency, half of the samples from each group were labeled with Cy3 dye and the other half with Cy5 dye. After 30 min. of incubation on ice in the dark, the reaction was quenched with 10mM lysine and additionally incubated for 10 min. Samples were finally combined, according to the experimental design, at 50$\mu$g of protein per Cy dye per gel, and diluted 2-fold with IEF sample buffer (7M urea, 2M thiourea, 4% (w/v) CHAPS, 2 % 1.4-Dithioerythritol (DTT), 2% pharmalytes pH 4-7, and 0.002% bromophenol blue). The 2D-DIGE was performed using GE-Healthcare reagents and equipment (GE Healthcare, Chalfont St. Giles, U.K.). First-dimension IEF was performed on IPG strips (24cm; linear gradient pH 4-7) using an Ettan IPGphor system (GE Healthcare, Chalfont St. Giles, U.K.). Initially, strips were incubated overnight in 450µL of RB with 1% pharmalytes pH 3-7, 100mM DeStreak, and 0.002% bromophenol blue. Then, samples were applied via cup loading near the acidic end of the strips. After focusing for a total of 67 kV x h, strips were equilibrated first for 15 min. in 6mL of reducing solution (6M urea, 100mM Tris-HCl pH 8, 30% (v/v) glycerol, 2% (w/v) SDS, 5mg/mL DTT, and 0.002% bromophenol blue) and then in 6mL of alkylating solution (6M urea, 100mM Tris-HCl pH 8, 30% (v/v) glycerol, 2% (w/v) SDS, 22.5mg/mL iodoacetamide (IAA), and 0.002% bromophenol blue) for 15 min on a rocking platform. Second-dimension SDS-PAGE was run by overlaying the strips on 12.5% isocratic Laemmli gels (24 × 20 cm), cast in low-fluorescence glass plates on an Ettan DALTsix

system (GE Healthcare, Chalfont St. Giles, U.K.). Gels were run at 20°C, at a constant power of 2.5 W/gel for 30 min, followed by 17 W/gel until the bromophenol blue tracking front reached the end of the gel. Fluorescence images of the gels were acquired on a Typhoon 9400 scanner (GE Healthcare, Chalfont St. Giles, U.K.). Cy2, Cy3, and Cy5 images were scanned at 488nm/520nm, 532nm/580, and 633nm/670nm excitation/emission wavelengths, respectively, at a 100$\mu$m resolution.

*Image and statistical analysis from 2D-DIGE experiments:* Image analysis and statistical quantification of relative protein abundance were performed using Progenesis Samespots v2.0 software (NonLinear Dynamics, Newcastle, U.K.). Protein features were selected by presenting *t* test values <0.05, when comparing PCa samples and benign samples, together with a fold-change greater than ±1.4.

*Multivariate Data Analysis of DIGE experiments:* To visualize differences between the complex proteome datasets, we used the multivariate analysis tool principal component analysis (PCA) as an explorative tool. A list of spots with their normalized spot volumes per gel was mean-centered. Subsequently, the dataset was examined with PCA. The dataset can be visualized as a cloud of points, where each point represents a sample in a multidimensional space. The coordinates of these points are represented by the spot intensities (dimensions). PCA reduces the large number of dimensions of a dataset into a smaller number of dimensions in such a way that most of the variance of the dataset is described by the first principal components (47, 48).

*Protein identification:* Protein spots of interest were excised from the gel using an automated Spot Picker (GE Healthcare, Uppsala, Sweden). Identification was first attempted on spots selected from one of the gels of the experiment after a total fluorescent staining using Flamingo (BioRad, California, USA). The gel contained a mixture of all of the protein samples analyzed, since the internal standard pool was loaded onto all of the gels. In-gel trypsin digestion was performed as described (49), using autolysis stabilized trypsin Gold (Promega, Madison, USA). Tryptic digests were purified using Zip Tip microtiter plates (Millipore, Bedford, MA). MALDI-TOF-MS/MS analysis of tryptic peptides was performed on an Ultraflex TOF-TOF Instrument (Bruker, Bremen, Germany). Samples were prepared using $\alpha$-cyano-4-hydroxy-cinnamic acid (CHCA) as a matrix on anchor-chip targets (Bruker Daltonics, Billerica, USA). Calibration was performed in the external mode using a peptide calibration standard kit (Bruker Daltonics, Billerica, USA). The spectra were processed using Flex Analysis 3.0 software (Bruker Daltonics, Billerica, USA). Peak lists were generated using the signals in the *m/z* 800-4,000 region with a signal-to-noise threshold of greater than 3. The SNAP algorithm included in the software was used to select the monoisotopic peaks from the isotopic

distributions observed. After removing *m/z* values corresponding to usually observed matrix cluster ions, an internal statistical calibration was applied. Peaks corresponding to frequently seen keratin and trypsin autolysis peptides were then removed. The resulting final peak list was used for identification of the proteins by peptide mass fingerprint. The Mascot 2.2 program (Matrix Science Ltd., London, U.K.) was used to search the Swiss-Prot 55.4 database, limiting the search to human proteins (19,630 sequences). Search parameters were as follows: trypsin cleavages excluding N-terminal to P, 1 or 2 missed cleavages allowed, cysteine carbamidomethylation set as fixed modification, methionine oxidation as variable modification, mass tolerance less than 50 ppm, monoisotopic mass values. In cases of identified fragments of proteins, the searched protein masses were set to the observed masses, according to gel position. The criteria for positive identification were a significant Mascot probability score (score > 55, $p < 0.05$). Alternatively, proteins were identified by ion trap MS, as described (50).

***3. Signaling pathways analysis***

Functional pathway and network analyses were generated through the use of Ingenuity Pathway Analysis (IPA) (version 2.0, Ingenuity® Systems, Mountain View, CA, USA). IPA identified those canonical pathways, biological processes and gene interaction networks that were most significant to the proteins selected from the 2D-DIGE analysis. Each protein designation was mapped in the Ingenuity Pathways Knowledge Base. Proteins that met the expression ratio cut-off of 1.5, a p-value cut-off of 0.05 for differential expression and were associated with a canonical pathway in the Ingenuity Pathways Knowledge Base were considered for the analysis.

**Qualification phase**

***4. Relative quantification SRM based assay***

The study population consisted of 50 men, 19 (38%) were positive for PCa and 31 (62%) were negative (benign controls without cancer) (Table 1). No type of selection was made, so that this group would have a similar prevalence to that which would normally be encountered in such a group of patients.

*Urine protein extraction and digestion:* First, urine supernatant samples were filtered through a 0.45 $\mu$m PVDF membrane filter (Millipore, Bedford, MA). 20 mL of each urine sample were concentrated to a final volume of 200 µL with Amicon 3kDa cutoff ultra-filtration units (Millipore, Bedford, MA). Then, the samples were diluted ten-fold with 2M urea 50 mM Ammonium Bicarbonate (BA) and concentrated again by ultra-filtration to a final volume of 200 µL in the same 3 kDa ultra-filtration unit. Thus, any salts in the

original samples were diluted. Protein concentration was determined using the BioRad RC DC Protein Assay (BioRad, California, USA). To ensure that the individual samples reflected the system adequately (each sample represented the same proportion of the system), we based the following experiment on the total amount of protein sample as determined by protein quantification assay. 20 µg of each sample were used to perform the following steps. We added 2pmols of Alcohol Dehydrogenase (ADH1) from baker's yeast (Sigma-Aldrich, Saint Louis, MO, USA) to each sample for use as a first internal standard. Then, the samples were reduced with 12.5mM DTT in 50mM BA at RT for 1 hour and after, they were alkylated with 40mM IAA in 50mM BA at RT for 30 min. IAA was quenched with 50mM N-acetyl-cysteine (NAC) in 50mM BA at RT for 30 min. Volume samples were adjusted to 1M Urea with 100mM BA. Samples were then digested with trypsin (Promega, Madison, WI) using a ratio of 1:10 (w/w). Trypsin digestion was carried out at 37ºC overnight. Digestion was halted by adding 1µL of 0.25% formic acid (FA) in water. 5pmols of Bovine Serum Albumin (BSA) peptides (Sigma-Aldrich, Saint Louis, MO, USA) were added to each sample for use as a second internal standard. The digested samples were diluted in 0.1 % FA to a final concentration of 0.5µg/µL. Samples were preserved at -20ºC until analyzed by LC-MS.

*SRM methods setup:* ProteinPilot software 3.0 (Applied Biosystems) was used to set up the SRM methods for the target peptides. One to three proteotypic peptides (peptides uniquely associated with the protein of interest) were selected for each targeted protein, and three ion fragments were selected for each of the parent masses. The uniqueness of the proteotypic peptides chosen was verified by running an identity search of the amino acid sequences against the UniProt database using the BLAST algorithm (51). In addition, the peptides selected were chosen following several criteria: amino acid sequences preferably not containing chemically reactive residues (i.e., Methionine or Tryptophan); peptides without missed tryptic cleavages; length of the peptide between 8 and 20 amino acids. In total, 41 peptides corresponding to 15 human proteins, 1 yeast and 1 bovine protein were selected. Yeast peptides were used to monitor the digestion process, and bovine peptides were used as a normalization tool. Standard curves for internal standards were performed, in order to find out the necessary amount to add to each sample in order to obtain a good signal (Supplemental data 2).

*LC-MS setup:* One µg of each urine sample digest was analyzed in triplicate on a Tempo Nano LC System (AB SCIEX, LLC, Dublin, CA, USA) in tandem with a Hybrid triple Quadrupole-Linear Ion Trap Mass Spectrometer 4000 Q TRAP® LC/MS/MS (AB SCIEX, LLC, Dublin, CA, USA). Peptide mixtures were initially concentrated on a 300 um id, 5mm ZORBAX (Agilent Technologies, Germany) precolumn and were subsequently

loaded onto a 75 um id, 15 cm PepMap nanoseparation column (LC Packings, Netherlands). An ACN gradient (40% of B in 60 min where B is 98% ACN, 0.1%FA; flow rate 300 nL/min) was used to elute the peptides through an emitter (New Objectives, Woburn, MA, USA) onto the electrospray ion source (ESI). A spray voltage of 2800 keV was used with a heated ion transfer tube set at a temperature of 150°C. The mass spectrometer was operated using the SRM or Information Dependent Acquisition (IDA) mode. For SRM acquisitions, the first quadrupole (Q1) was set at low (1.5 FWHM) mass resolution and the third quadrupole (Q3) at unit (0.7 FWHM) mass resolution. In the IDA acquisition mode, Q1 and Q3 were set at low (1.5 FWHM) mass resolution. Scheduled SRM was used to achieve monitoring of a large number of transitions while keeping a high enough dwell time for each transition. For scheduled SRM an acquisition time window of 3 min was set around the observed mean elution time for each precursor peptide. IDA MSMS scans were acquired for 4.2s on an m/Z range from 250 to 15000. Nitrogen was used as the collision gas at 100-105 psig. Collision energies were calculated according the precursor m/z and charge.

*Data analysis:* Analyst software (AB SCIEX, LLC, Dublin, CA, USA) was used to process the SRM data obtained. Internal standards added to each sample (ADH1 and BSA) helped us to monitor whether the sample processing was adequate and also helped to normalize the data. The sum of the peak areas of each transition-peptide was used to calculate the relative amounts of the given peptides [Sum Area (peptide-Protein)/Sum Area (peptide-Internal standard)].

***5. Statistical Analysis***

Logistic regression analysis was performed using SPSS software, in order to determine the best urine biomarker candidates. The characterization of candidate biomarkers was accomplished by comparing their log mean values (univariate analysis). The t-test with Welch correction was used when the distribution of the data was normal, and a Mann-Whitney test was used in other cases (Prism4, V.4a).

## RESULTS

### Discovery phase

***1. 2D-DIGE analysis and identification of differentially expressed proteins in urine total protein extracts***

We compared the urine supernatants obtained after PM from 6 PCa patients to 6 from benign patients using 2D-DIGE technology. 2D-DIGE permitted the direct comparison of all of the components from all of the patients, as the protein gels were normalized with a

pool of proteins from all of the samples included in the study. The proteins were resolved on a 12.5% acrylamide gel after an IEF on non-linear gradient strips of pH 4-7, and the statistical analysis was performed using the Progenesis SameSpots v2.0 software. To avoid any possible bias derived from labeling efficiency, half of the samples from each group were labeled with Cy3 dye and the other half with Cy5 dye. A third fluorescent dye, Cy2, was used to label the internal standard sample. The three samples (PCa, benign and the internal standard pool) were then run on each gel. We selected for further analysis those spots, which presented fold changes greater than ±1.4 with a p-value ≤0.05 (ANOVA). We selected the statistically significant spots existing in the 2D-DIGE analyses and submitted these to MALDI-TOF-MS/MS identification. We identified 17 spots, 10 over-expressed and 7 under-expressed (Table 2). This set of selected, differentially expressed spots was then used for correlation analysis and hierarchical clustering, based on Pearson's Product Moment correlation coefficients (52), in order to classify the detected spots depending on the behavior among the two types of samples. Principal Component Analysis (PCA) showed the same different behaviors between these two groups (Figure 1).

***2. 2D-DIGE analysis and identification of differentially expressed proteins* in ProteoMiner enriched urine protein extracts**

We selected urine supernatants obtained after PM from 9 PCa patients and 9 benign patients. We had previously enriched the urine-extracted proteins using ProteoMiner (BioRad, California, USA). Previous studies in our lab supported the utilization of ProteoMiner technology (Supplemental data 1). The 2D-DIGE experiment was performed identically to the 2D-DIGE analysis described above. We identified 24 spots. 13 of these proteins were over-expressed and 11 proteins under-expressed, when comparing the PCa samples to the control samples (Table 2). The differentially expressed spots were used for correlation analysis and hierarchical clustering, in order to classify the detected spots depending on the behavior between the two types of samples. PCA showed the same different behaviors between these two groups (Figure 1).

***3. Protein summary and ingenuity pathways analysis of all of the proteins identified in the 2D-DIGE experiments***

In summary, when we compared the supernatants from the urine of PCa patients to that of benign age-matched individuals, we were able to identify 24 potential biomarkers, 15 of which were under-expressed (62%) and 9 of which were over-expressed (38%) (Table 2). 62% of the identified proteins were localized in the extracellular space, and 12% were proteins localized in the plasma membrane. IPA showed that the majority of these

proteins were secreted components of several well-known, functional cancer and inflammation networks, such as NFKB, PDGFBβ and PSA (Figure 2).

**Qualification phase**

***4. Relative quantification by SRM-based assay***

Qualification of 15 of the 24 identified biomarker candidates was carried out by relative quantification using an SRM-based assay on 50 urine samples obtained after PM. 19 of the 50 samples (38%) presented PCa (Table 1). Triplicate values were obtained for each sample analyzed. In order to correct differences of injection or instrument performance between the different runs the signal of BSA peptide (HLVDEPQNLIK), added as internal standard, was used to normalize the intensity of the measured areas [Area Target peptide/Area HLVDEPQNLIK], A proven method for the validation of peptide transitions is to acquire the MS/MS spectra, which are then subjected to sequence database searches to confirm that the detected signals are derived from the targeted peptide (53). We obtained MS/MS spectra and the protein identification result for every sample.

***5. Statistical analysis of the Relative Quantification data***

Logistic regression analysis was used to select the candidate protein biomarkers to be precisely quantified in the second validation step. 7 peptides corresponding to 5 different proteins were selected. The Area Under the curve of each selected peptide was represented by Receiver Operating Characteristic (ROC) curves. A new protein marker was generated using a combination of these 7 peptides that showed an AUC value of 0.932. Fixing the sensitivity at 95%, we achieved a specificity of 78% (Figure 3a). One of the selected protein biomarkers was PSA (or KLK3), which was a significant predictor for PCa (p-value < 0.0001) (Figure 3b).

**DISCUSSION**

In the past few years the proteomics field has moved onward to clinical applications, particularly for biomarker discovery and the diagnosis and prognosis of human diseases (54). The need for developing more effective cancer biomarkers and therapeutic methods has led to the study of cancer cell secretomes, as a means of identifying and characterizing diagnostic and prognostic markers and identifying potential drug and therapeutic targets. In particular, the contribution of proteomics to the understanding of the pathogenesis and diagnosis of diseases from the genitourinary tract, such as PCa, has been considerable (55). PCa cells secrete different products, which can be found in urine and, as PM leads to an enrichment of prostatic fluid in the first urine catch after this

procedure, urine has been identified as an ideal clinical sample for biomarker discovery. Moreover, urine is readily available in almost all patients, and its collection is very simple and non-invasive.

The two main obstacles encountered in the discovery of biomarkers in urine, as well as in plasma, are the following: the high dynamic range, which results in only the high-abundance proteins being displayed on any typical 2D-GE map, and the low reproducibility of 2D-GE. Urine, as any other body fluid, contains some highly abundant proteins, such as albumin or IgG. Therefore, it is necessary to enrich the samples (decreasing the level of abundant proteins), in order to get a better representation of the low-abundance proteins, which can be the most interesting in biomarker discovery. Nowadays, there is still great debate regarding the performance of pretreatments that have been used on plasma samples, as well as on urine samples for biomarker discovery (56). Although many depletion protocols exist for the treatment of plasma samples, the field is not nearly as clear for urine samples (45). Since there is only a small overlap between the high–abundance plasma proteins, the use of commercial plasma depletion columns can be not efficient. These commercial kits will remove the proteins that are bound to high-abundance proteins, such albumin, and these are inadvertently lost with the use of commercial immunodepletion products. For this reason, in our study we used a concentration/equalization technique called ProteoMiner (45). Instead of simplifying the complex mixture into fractions or partitioning away the most abundant species, it captures all of the species present in the solution up to the saturation of the solid phase ligand library. This automatically results in a dramatic dilution of the most abundant species with a concurrent concentration of the diluted and rare ones. Here, the concept was applied to the analysis of those extremely diluted urinary proteins that normally escape regular detection methods, offering us a view on the normally hidden urine proteome. Another advantage of this technique was that, in contrast to most of the common approaches, it required less time for urine preparation and involved fewer sample-losing steps prior to the 2D mapping (57-59). In the present study, we evaluated whether the ProteoMiner could be used on urine samples as an enrichment technique, in order to obtain better quality and spots resolution for a 2D-DIGE analysis (Supplemental data 1).

Although the analysis of urine samples has become increasingly popular in recent years, due to the non-invasive nature of sample collection and its relatively reduced complexity compared to serum, only a limited number of studies have focused on PCa. However, to our knowledge no study has been performed for PCa using the combination of urine samples and 2D-DIGE. A pilot study for PCa using Capillary Electrophoresis MS (CE-

MS) was able to define a potential urinary polypeptide pattern with 92% sensitivity and 96% specificity (60). However, these data could not be validated on a larger cohort. Later on, the same group went on to describe a refinement of the PCa-specific biomarker pattern using samples from 51 PCa and 35 BPH patients. The result was a model that included 12 potential biomarkers with a correct classification in 89% of the PCa and 51% of the BPH patients in a second blind cohort of 213 samples. Inclusion of age and free PSA increased the sensitivity and specificity to 91% and 69%, respectively (37). M'Koma and collaborators also performed a large-scale urinary proteomic analysis among BPH, HGPIN and PCa. Using MALDI-TOF, 71.2% specificity and 67.4% sensitivity were reported in discriminating between PCa and BPH (61).

The recent explosion in proteomic technologies centering on MS has provided great opportunities for researchers to use them as "bridging technologies" for clinical proteomic investigation of diseases relevant to changes in body fluids (5). To efficiently translate proteomic technologies to the clinics, it is necessary to contemplate four stages in the biomarker development pipeline: discovery, qualification, verification, and validation. This manuscript has attempted to follow this biomarker development pipeline in the discovery phase for identifying potential urine biomarkers for PCa by using 2D-DIGE coupled with MALDI-TOF-MS/MS. These "candidate biomarkers" were verified using SRM-based approach.

Along with protein identification, protein quantification is another component of the biomarker discovery process. This step determines the changes in protein expression between the disease and control states. The utilization of targeted proteomic techniques, such as LC-MS/MS, should be ideal for both verification and validation studies in a high throughput fashion. Targeted proteomic approaches, such as the SRM-based assay, are emerging in proteomics as ideal tools for complementing shotgun qualitative studies (62). SRM may be used to achieve the precise quantification of a specific group of proteins. A predefined precursor and one of its fragments are selected by two mass filters of one triple quadrupole MS instrument and are monitored over time for precise quantification. A series of transitions (precursor/fragment ion pairs), in combination with the retention time of the targeted peptide, can constitute a definitive assay. One of the advantages of this technique is the possibility of quantifying a great number of peptides after only one single LC-MS/MS experiment. SRM can bridge the gap between biomarker discovery, usually performed on a few samples, and validation by antibody-based approaches, which are costly and slow to develop (53).

In the present study, 24 potential biomarkers (15 down- and 9 up-regulated) were found differentially expressed in a significant manner in the urinary specimens of patients with

PCa compared to the age-matched controls confirmed by PB, the current gold standard in PCa diagnostics. 15 of the 24 biomarkers were qualified first by the SRM relative quantification based assay on an independent set of 50 samples (38% PCa samples). The reasoning behind this was the reduction of the candidate protein number for a second verification step using absolute quantification. Relative quantification can be based on the signal intensities of the specific SRM transitions. In an MS experiment, the intensity of the signal can be plotted over time as the peptide elutes from the chromatographic column (extracted ion chromatogram, XIC). The area under this curve, for a given peptide under identical experimental conditions, is linearly related to its amount (63). The magnitude of the fluctuations depends on the peptide sequence, the background and several experimental factors that cannot be precisely controlled (e.g., fluctuations in ionization efficiency over time and matrix effects that lead to ion suppression due to enhancement from co-eluting analytes). Several peptides (2 or 3) from each protein are quantified, in order to avoid false conclusions due to these effects (40). After a logistic regression analysis of the data obtained from the Relative quantification SRM assay, we obtained a panel of 7 peptides within 5 different proteins (Figure 3a) that were able to distinguish between PCa samples and benign control samples with a sensitivity of 95% and a specificity of 78%.

In conclusion, our data demonstrated that proteomic analyses are able to reveal novel biomarkers/ diagnostic profiles for PCa in urine. This constitutes an important step towards advancing its accurate diagnosis, which currently represents a setback in our ability to cure patients of PCa. There are two basic considerations, which must be applied in all biomarker clinical studies. These are the fact that the number of independent variables should always be kept to a minimum, certainly less than the number of samples investigated. Also, any such approach must always be confirmed by a blind validation set. Thus, the next step for urinary proteomics will be to move onward towards the verification of these biomarkers in a much larger sample size and to translate these findings to a much easier format for the development of diagnostic kits or tests that may be directly used in clinical practice (54). The results obtained from these studies will have a rapid application in the clinics that can potentially influence decisions, which will, in turn, improve the health system, as well as the clinical, managerial and/or public practices for health outcomes in PCa.

## BIBLIOGRAPHY

1. Ferlay, J., Parkin, D. M., and Steliarova-Foucher, E. (2010) Estimates of cancer incidence and mortality in Europe in 2008. *Eur J Cancer* 46, Page.

2. Jemal, A., Siegel, R., Ward, E., Hao, Y., Xu, J., and Thun, M. J. (2009) Cancer statistics, 2009. *CA Cancer J Clin*, Page.

3. Catalona, W. J., Smith, D. S., Ratliff, T. L., Dodds, K. M., Coplen, D. E., Yuan, J. J., Petros, J. A., and Andriole, G. L. (1991) Measurement of prostate-specific antigen in serum as a screening test for prostate cancer. *N Engl J Med* 324, Page.

4. Makinen, T., Tammela, T. L., Stenman, U. H., Maattanen, L., Aro, J., Juusela, H., Martikainen, P., Hakama, M., and Auvinen, A. (2004) Second round results of the Finnish population-based prostate cancer screening trial. *Clin Cancer Res* 10, Page.

5. Boja, E., Hiltke, T., Rivers, R., Kinsinger, C., Rahbar, A., Mesri, M., and Rodriguez, H. (2010) Evolution of Clinical Proteomics and its Role in Medicine. *J Proteome Res*, Page.

6. Makridakis, M., and Vlahou, A. (2010) Secretome proteomics for discovery of cancer biomarkers. *J Proteomics*, Page.

7. Pavlou, M. P., and Diamandis, E. P. (2010) The cancer cell secretome: A good source for discovering biomarkers? *J Proteomics*, Page.

8. Decramer, S., Gonzalez de Peredo, A., Breuil, B., Mischak, H., Monsarrat, B., Bascands, J. L., and Schanstra, J. P. (2008) Urine in clinical proteomics. *Mol Cell Proteomics* 7, Page.

9. Adachi, J., Kumar, C., Zhang, Y., Olsen, J. V., and Mann, M. (2006) The human urinary proteome contains more than 1500 proteins, including a large proportion of membrane proteins. *Genome Biol* 7, Page.

10. Thongboonkerd, V., McLeish, K. R., Arthur, J. M., and Klein, J. B. (2002) Proteomic analysis of normal human urinary proteins isolated by acetone precipitation or ultracentrifugation. *Kidney Int* 62, Page.

11. Pieper, R., Gatlin, C. L., McGrath, A. M., Makusky, A. J., Mondal, M., Seonarain, M., Field, E., Schatz, C. R., Estock, M. A., Ahmed, N., Anderson, N. G., and Steiner, S. (2004) Characterization of the human urinary proteome: a method for high-resolution display of urinary proteins on two-dimensional electrophoresis gels with a yield of nearly 1400 distinct protein spots. *Proteomics* 4, Page.

12. Oh, J., Pyo, J. H., Jo, E. H., Hwang, S. I., Kang, S. C., Jung, J. H., Park, E. K., Kim, S. Y., Choi, J. Y., and Lim, J. (2004) Establishment of a near-standard two-dimensional human urine proteomic map. *Proteomics* 4, Page.

13. Hu, S., Loo, J. A., and Wong, D. T. (2006) Human body fluid proteome analysis. *Proteomics* 6, Page.

14. Selevsek, N., Matondo, M., Carbayo, M. S., Aebersold, R., and Domon, B. Systematic quantification of peptides/proteins in urine using selected reaction monitoring. *Proteomics* 11, Page.

15. States, D. J., Omenn, G. S., Blackwell, T. W., Fermin, D., Eng, J., Speicher, D. W., and Hanash, S. M. (2006) Challenges in deriving high-confidence protein identifications from data gathered by a HUPO plasma proteome collaborative study. *Nat Biotechnol* 24, Page.

16. Afkarian, M., Bhasin, M., Dillon, S. T., Guerrero, M. C., Nelson, R. G., Knowler, W. C., Thadhani, R., and Libermann, T. A. (2010) Optimizing a proteomics platform for urine biomarker discovery. *Mol Cell Proteomics*, Page.

17. Court, M., Selevsek, N., Matondo, M., Allory, Y., Garin, J., Masselon, C. D., and Domon, B. Toward a standardized urine proteome analysis methodology. *Proteomics* 11, Page.

18. Petricoin, E. F., Ardekani, A. M., Hitt, B. A., Levine, P. J., Fusaro, V. A., Steinberg, S. M., Mills, G. B., Simone, C., Fishman, D. A., Kohn, E. C., and Liotta, L. A. (2002) Use of proteomic patterns in serum to identify ovarian cancer. *Lancet* 359, Page.

19. Schiffer, E., Mischak, H., and Novak, J. (2006) High resolution proteome/peptidome analysis of body fluids by capillary electrophoresis coupled with MS. *Proteomics* 6, Page.

20. Kulasingam, V., and Diamandis, E. P. (2008) Strategies for discovering novel cancer biomarkers through utilization of emerging technologies. *Nat Clin Pract Oncol* 5, Page.

21. Sun, W., Li, F., Wu, S., Wang, X., Zheng, D., Wang, J., and Gao, Y. (2005) Human urine proteome analysis by three separation approaches. *Proteomics* 5, Page.

22. Wang, X., Lv, H., Zhang, G., Sun, W., Zhou, D., Jiao, G., and Yu, Y. (2008) Development and validation of a ultra performance LC-ESI/MS method for analysis of metabolic phenotypes of healthy men in day and night urine samples. *J Sep Sci* 31, Page.

23. Hanash, S. M. (2000) Biomedical applications of two-dimensional electrophoresis using immobilized pH gradients: current status. *Electrophoresis* 21, Page.

24. Thongboonkerd, V. (2007) Practical points in urinary proteomics. *J Proteome Res* 6, Page.

25. Anderson, N. G., Anderson, N. L., and Tollaksen, S. L. (1979) Proteins of human urine. I. Concentration and analysis by two-dimensional electrophoresis. *Clin Chem* 25, Page.

26. Gygi, S. P., Corthals, G. L., Zhang, Y., Rochon, Y., and Aebersold, R. (2000) Evaluation of two-dimensional gel electrophoresis-based proteome analysis technology. *Proc Natl Acad Sci U S A* 97, Page.

27. Hoorn, E. J., Hoffert, J. D., and Knepper, M. A. (2006) The application of DIGE-based proteomics to renal physiology. *Nephron Physiol* 104, Page.

28. Banon-Maneus, E., Diekmann, F., Carrascal, M., Quintana, L. F., Moya-Rull, D., Bescos, M., Ramirez-Bajo, M. J., Rovira, J., Gutierrez-Dalmau, A., Sole-Gonzalez, A., Abian, J., and Campistol, J. M. (2010) Two-dimensional difference gel electrophoresis urinary proteomic profile in the search of nonimmune chronic allograft dysfunction biomarkers. *Transplantation* 89, Page.

29. Taneja, S., Sen, S., Gupta, V. K., Aggarwal, R., and Jameel, S. (2009) Plasma and urine biomarkers in acute viral hepatitis E. *Proteome Sci* 7, Page.

30. Jiang, F. Z., Cui, Q. L., and Jiang, J. H. (2009) [A case report of maple syrup urine disease]. *Zhongguo Dang Dai Er Ke Za Zhi* 11, Page.

31. Orenes-Pinero, E., Corton, M., Gonzalez-Peramato, P., Algaba, F., Casal, I., Serrano, A., and Sanchez-Carbayo, M. (2007) Searching urinary tumor markers for bladder cancer using a two-dimensional differential gel electrophoresis (2D-DIGE) approach. *J Proteome Res* 6, Page.

32. Pang, J., Liu, W. P., Liu, X. P., Li, L. Y., Fang, Y. Q., Sun, Q. P., Liu, S. J., Li, M. T., Su, Z. L., and Gao, X. (2010) Profiling protein markers associated with lymph node metastasis in prostate cancer by DIGE-based proteomics analysis. *J Proteome Res* 9, Page.

33. Byrne, J. C., Downes, M. R., O'Donoghue, N., O'Keane, C., O'Neill, A., Fan, Y., Fitzpatrick, J. M., Dunn, M., and Watson, R. W. (2009) 2D-DIGE as a strategy to identify serum markers for the progression of prostate cancer. *J Proteome Res* 8, Page.

34. Skvortsova, I., Skvortsov, S., Stasyk, T., Raju, U., Popper, B. A., Schiestl, B., von Guggenberg, E., Neher, A., Bonn, G. K., Huber, L. A., and Lukas, P. (2008) Intracellular signaling pathways regulating radioresistance of human prostate carcinoma cells. *Proteomics* 8, Page.

35. Rehman, I., Azzouzi, A. R., Catto, J. W., Allen, S., Cross, S. S., Feeley, K., Meuth, M., and Hamdy, F. C. (2004) Proteomic analysis of voided urine after prostatic massage from patients with prostate cancer: a pilot study. *Urology* 64, Page.

36. M'Koma A, E., Blum, D. L., Norris, J. L., Koyama, T., Billheimer, D., Motley, S., Ghiassi, M., Ferdowsi, N., Bhowmick, I., Chang, S. S., Fowke, J. H., Caprioli, R. M., and Bhowmick, N. A. (2007) Detection of pre-neoplastic and neoplastic prostate disease by MADI profiling of urine. *Biochem Biophys Res Commun* 353, Page.

37. Theodorescu, D., Schiffer, E., Bauer, H. W., Douwes, F., Eichhorn, F., Polley, R., Schmidt, T., Schofer, W., Zurbig, P., Good, D. M., Coon, J. J., and Mischak, H. (2008) Discovery and validation of urinary biomarkers for prostate cancer. *Proteomics Clin Appl* 2, Page.

38. Rifai, N., Gillette, M. A., and Carr, S. A. (2006) Protein biomarker discovery and validation: the long and uncertain path to clinical utility. *Nat Biotechnol* 24, Page.

39. Rodriguez, H., Rivers, R., Kinsinger, C., Mesri, M., Hiltke, T., Rahbar, A., and Boja, E. Reconstructing the pipeline by introducing multiplexed multiple reaction monitoring mass spectrometry for cancer biomarker verification: an NCI-CPTC initiative perspective. *Proteomics Clin Appl* 4, Page.

40. Gallien, S., Duriez, E., and Domon, B. Selected reaction monitoring applied to proteomics. *J Mass Spectrom* 46, Page.

41. Kiyonami, R., Schoen, A., Prakash, A., Peterman, S., Zabrouskov, V., Picotti, P., Aebersold, R., Huhmer, A., and Domon, B. (2010) Increased selectivity, analytical precision, and throughput in targeted proteomics. *Mol Cell Proteomics*, Page.

42. Choi, S., Kim, J., Yea, K., Suh, P. G., and Ryu, S. H. Targeted label-free quantitative analysis of secretory proteins from adipocytes in response to oxidative stress. *Anal Biochem* 401, Page.

43. Remzi, M., Fong, Y. K., Dobrovits, M., Anagnostou, T., Seitz, C., Waldert, M., Harik, M., Marihart, S., Marberger, M., and Djavan, B. (2005) The Vienna nomogram: validation of a novel biopsy strategy defining the optimal number of cores based on patient age and total prostate volume. *J Urol* 174, Page.

44. Khan, A., and Packer, N. H. (2006) Simple urinary sample preparation for proteomic analysis. *J Proteome Res* 5, Page.

45. Castagna, A., Cecconi, D., Sennels, L., Rappsilber, J., Guerrier, L., Fortis, F., Boschetti, E., Lomas, L., and Righetti, P. G. (2005) Exploring the hidden human urinary proteome via ligand library beads. *J Proteome Res* 4, Page.

46. Alban, A., David, S. O., Bjorkesten, L., Andersson, C., Sloge, E., Lewis, S., and Currie, I. (2003) A novel experimental design for comparative two-dimensional gel analysis: two-dimensional difference gel electrophoresis incorporating a pooled internal standard. *Proteomics* 3, Page.

47. Kleno, T. G., Leonardsen, L. R., Kjeldal, H. O., Laursen, S. M., Jensen, O. N., and Baunsgaard, D. (2004) Mechanisms of hydrazine toxicity in rat liver investigated by proteomics and multivariate data analysis. *Proteomics* 4, Page.

48. Karp, N. A., Griffin, J. L., and Lilley, K. S. (2005) Application of partial least squares discriminant analysis to two-dimensional difference gel studies in expression proteomics. *Proteomics* 5, Page.

49. Shevchenko, A., Wilm, M., Vorm, O., and Mann, M. (1996) Mass spectrometric sequencing of proteins silver-stained polyacrylamide gels. *Anal Chem* 68, Page.

50. Esselens, C. W., Malapeira, J., Colome, N., Moss, M., Canals, F., and Arribas, J. (2008) Metastasis-associated C4.4A, a GPI-anchored protein cleaved by ADAM10 and ADAM17. *Biol Chem* 389, Page.

51. http://blast.ncbi.nlm.nih.gov/Blast.cgi.

52. Deighton, R. F., Short, D. M., McGregor, R. J., Gow, A. J., Whittle, I. R., and McCulloch, J. (2009) The utility of functional interaction and cluster analysis in CNS proteomics. *J Neurosci Methods* 180, Page.

53. Lange, V., Picotti, P., Domon, B., and Aebersold, R. (2008) Selected reaction monitoring for quantitative proteomics: a tutorial. *Mol Syst Biol* 4, Page.

54. Thongboonkerd, V. (2008) Urinary proteomics: towards biomarker discovery, diagnostics and prognostics. *Mol Biosyst* 4, Page.

55. Theodorescu, D., and Mischak, H. (2007) Mass spectrometry based proteomics in urine biomarker discovery. *World J Urol* 25, Page.

56. Di Girolamo, F., Boschetti, E., Chung, M. C., Guadagni, F., and Righetti, P. G. "Proteomineering" or not? The debate on biomarker discovery in sera continues. *J Proteomics* 74, Page.

57. Boschetti, E., and Giorgio Righetti, P. (2008) Hexapeptide combinatorial ligand libraries: the march for the detection of the low-abundance proteome continues. *Biotechniques* 44, Page.

58. Boschetti, E., and Righetti, P. G. (2008) The ProteoMiner in the proteomic arena: A non-depleting tool for discovering low-abundance species. *J Proteomics*, Page.

59. Righetti, P. G., and Boschetti, E. (2008) The ProteoMiner and the FortyNiners: Searching for gold nuggets in the proteomic arena. *Mass spectrometry reviews*, Page.

60. Theodorescu, D., Fliser, D., Wittke, S., Mischak, H., Krebs, R., Walden, M., Ross, M., Eltze, E., Bettendorf, O., Wulfing, C., and Semjonow, A. (2005) Pilot study of capillary electrophoresis coupled to mass spectrometry as a tool to define potential prostate cancer biomarkers in urine. *Electrophoresis* 26, Page.

61. M'Koma, A. E., Blum, D. L., Norris, J. L., Koyama, T., Billheimer, D., Motley, S., Ghiassi, M., Ferdowsi, N., Bhowmick, I., Chang, S. S., Fowke, J. H., Caprioli, R. M., and Bhowmick, N. A. (2007) Detection of pre-neoplastic and neoplastic prostate disease by MALDI profiling of urine. *Biochem Biophys Res Commun* 353, Page.

62. Kiyonami, R., Schoen, A., Prakash, A., Peterman, S., Zabrouskov, V., Picotti, P., Aebersold, R., Huhmer, A., and Domon, B. Increased selectivity, analytical precision, and throughput in targeted proteomics. *Mol Cell Proteomics* 10, Page.

63. Andersen, J. S., Lam, Y. W., Leung, A. K., Ong, S. E., Lyon, C. E., Lamond, A. I., and Mann, M. (2005) Nucleolar proteome dynamics. *Nature* 433, Page.

## FIGURE LEGENDS

**Table 1: Clinical and Pathological data of patients**

**Table 2: Proteins identified in 2D-DIGE analyses**

**Figure 1. 2D-DIGE analyses.** (A) Example of a representative gel of a 2D-DIGE experiment on total protein urine samples. (E) Example of a representative gel of a 2D-DIGE experiment on ProteoMiner enriched urine samples. (B) and (F) Example of differential spots in PCa samples against Benign ones, in the first and second DIGE experiments, respectively. (C and D) Principal Component Analysis, where blue points correspond to benign samples and red points PCa ones from the DIGE experiment in A) and B), respectively. (D and H) Correlation analysis enables the clustering of proteins according to their expression profile among control urine samples and PCa urine samples in the two DIGE experiments.

**Figure 2: Biological information of the spots identified.** (A) Identified Protein spots relative expression. (B) Cellar localization of identified proteins (C) IPA network interaction of over-expressed protein spots (red figures) and under-expressed protein spots (green figures).

**Figure 3. Relative qualification data Analysis.** (A) ROC curves of the 7 selected peptides (yellow, black, green, blue, pink, grey and red lines) and Mutli-ROC curve of the MultiMarker (brown line). (B) Example of urine based biomarker for PCa (box-and-whisker plot). Relative level of PSA peptide (LSEPAELTDAVK) and the corresponding behavior in the 2D-DIGE comparison analysis.

**Supplemental Figure 1: ProteoMiner set up experiments**. ProteoMiner technique allows enrichment of Urine supernatant samples. (A) Silver staining gel of urine supernatant sample using first DIGE protein extraction protocol. (B) Silver staining gel of urine supernatant sample using PoteoMiner enrichment protocol. (C and D) examples of spots increased when use the ProteoMiner enrichment protocol, right panel, comparing to non-ProteoMiner protocol, left panel. (E) DIGE fluorescent image gel of two urine supernatant protein extracted samples using ProteoMiner technique. All spots are overlapped.

**Supplemental Figure 2: SRM-control peptides Standard Curves.** (A) BSA from Bovine peptides monitored using SRM assay and (B) ADH from Yeast peptides monitored using SRM assay in a 4000 qTRAP from Applied Biosystems.

**Table 1: Clinical and Pathological data of patients**

| | Protein Discovery | | | | Protein Qualification | |
|---|---|---|---|---|---|---|
| | 1st 2D-DIGE | | 2nd 2D-DIGE | | Relative Quantification | |
| | All patients average | (min, max) | All patients average | (min, max) | All patients average | (min, max) |
| No of patients | 12 | | 18 | | 50 | |
| Age (yr) | 66.8 | (56, 80) | 66.6 | (53, 77) | 65.5 | (47, 79) |
| PSA level (range) | 21.6 | (4.8, 82.2) | 38.15 | (3.8, 44.3) | 11.65 | (3.8, 166.2) |
| PSA free | 2.41 | (0.59, 8.57) | 2.07 | (0.49, 4.69) | 2.03 | (0.43, 43,65) |
| Ratio free PSA/tPSA | 0.16 | (0.08, 0.29) | 0.19 | (0.08, 0.41) | 0.15 | (0.06, 0.41) |
| Prostate cancer | 6 (50%) | | 9 (50%) | | 19 (38%) | |
| Gleason < 7 | 3 (50%) | | 1 (11%) | | 1 (5%) | |
| Gleason = 7 | 1 (17%) | | 5 (56%) | | 11 (58%) | |
| Gleason > 7 | 2 (33%) | | 3 (33%) | | 5 (26%) | |
| Benign | 6 (50%) | | 9 (50%) | | 31 (62%) | |

**Table 2: Proteins identified in 2D-DIGE analyses**

| DIGE | Spot | Fold NvsC | Anova NvsC | Theor. pI | Observ. pI | Theor. MW (Da) | Observed MW (Da) | Score | Pept (Seq. %) |
|---|---|---|---|---|---|---|---|---|---|
| 1 | 1 | -6,13 | 0,048 | 7,62 | 7 | 29293 | 30000 | 99 | 10 (43%) |
| 1 | 2 | -5,76 | 0,049 | 8,47 | 6 | 58277 | 22000 | 72 | 16 (30%) |
| 1 | 9 | -3,35 | 0,009 | 4,99 | 6 | 51578 | 20000 | 104 | 15 (31%) |
| 1 | 17 | -2,78 | 0,001 | 7,55 | 6,8 | 26077 | 23000 | 74 | 8 (38%) |
| 1 | 21 | -2,40 | 0,007 | 8,27 | 6,8 | 22324 | 22000 | 76 | 8 (39%) |
| 1 | 25 | -2,11 | 0,022 | 6,23 | 6,2 | 143654 | 20000 | 74 | 18 (16%) |
| 1 | 40 | -1,73 | 0,041 | 6,76 | 6,7 | 20146 | 20000 | 66 | 5 (29%) |
| 1 | 43 | 1,57 | 0,035 | 5,44 | 4,5 | 122983 | 120000 | 78 | 17 (14%) |
| 1 | 29 | 1,99 | 0,036 | 8,48 | 4,8 | 31647 | 35000 | 92 | 10 (38%) |
| 1 | 134 | 2,29 | 0,050 | 5,58 | 4,5 | 54809 | 75000 | 166 | 21 (41%) |
| 1 | 19 | 2,71 | 0,045 | 5,37 | 4 | 46878 | 50000 | 203 | 24 (51%) |
| 1 | 18 | 2,77 | 0,040 | 5,37 | 4 | 46878 | 50000 | 236 | 27 (55%) |
| 1 | 16 | 2,83 | 0,032 | 6,02 | 6 | 188569 | 75000 | 173 | 30 (17%) |
| 1 | 13 | 3,04 | 0,045 | 8,48 | 4,8 | 31647 | 37000 | 107 | 12 (43%) |
| 1 | 8 | 3,43 | 0,044 | 5,37 | 4,5 | 46878 | 50000 | 169 | 28 (55%) |
| 1 | 4 | 4,05 | 0,042 | 5,84 | 4,5 | 40678 | 50000 | 105 | 12 (42%) |
| 1 | 3 | 4,30 | 0,033 | 8,48 | 4,8 | 31647 | 37000 | 115 | 12 (52%) |
| 2 | 701 | -8,22 | 0,010 | 5,83 | 5,7 | 44880 | 40000 | 74 | 1 (4%) |
| 2 | 692 | -5,89 | 0,002 | 5,83 | 5,6 | 44880 | 40000 | 66 | 5 (31%) |
| 2 | e32 | -2,69 | 0,009 | 5,71 | 5,6 | 36900 | 35000 | 125 | 3 (10%) |
| 2 | e32 | -2,69 | 0,009 | 5,83 | 5,5 | 44880 | 40000 | 192 | 2 (8%) |
| 2 | 1020 | -2,68 | 0,036 | 5,87 | 5,4 | 53406 | 15000 | 163 | 2 (8%) |
| 2 | 672 | -2,50 | 0,046 | 5,28 | 4,9 | 45371 | 37000 | 169 | 4 (14%) |
| 2 | 998 | -2,49 | 0,048 | 6,75 | 5,9 | 16102 | 15000 | 50 | 1 (8%) |
| 2 | 998 | -2,49 | 0,048 | 5,83 | 5,9 | 44880 | 15000 | 42 | 1 (3%) |
| 2 | e10 | -2,28 | 0,024 | 7,66 | 7 | 36489 | 35000 | 173 | 3 (4%) |
| 2 | e10 | -2,28 | 0,024 | 8,5 | 7 | 80014 | 15000 | 69 | 2 (3%) |
| 2 | e47 | -1,80 | 0,006 | 6,07 | 6 | 64262 | 60000 | 78 | 8 (17%) |
| 2 | 275 | 1,44 | 0,046 | 6,02 | 6,1 | 188569 | 100000 | 65 | 8 (4%) |
| 2 | 155 | 1,48 | 0,042 | 5,44 | 5 | 122983 | 150000 | 74 | 9 (12%) |
| 2 | 325 | 1,75 | 0,012 | 5,92 | 5 | 71317 | 75000 | 126 | 3 (6%) |
| 2 | 242 | 1,79 | 0,008 | 6,81 | 6,1 | 79280 | 95000 | 99 | 10 (19%) |
| 2 | 905 | 1,84 | 0,011 | 5,61 | 5,1 | 21129 | 20000 | 123 | 9 (41%) |
| 2 | 243 | 1,85 | 0,004 | 6,81 | 6 | 79280 | 95000 | 812 | 12 (23%) |
| 2 | 240 | 1,91 | 0,005 | 6,81 | 6 | 79280 | 95000 | 295 | 31 (48%) |
| 2 | 244 | 1,92 | 0,003 | 6,81 | 6 | 79280 | 95000 | 238 | 28 (45%) |
| 2 | 383 | 1,99 | 0,038 | 5,37 |  | 46878 |  | 184 | 17 (48%) |
| 2 | 285 | 2,13 | 0,017 | 6,02 | 6,2 | 188569 | 100000 | 221 | 3 (2%) |
| 2 | 403 | 2,45 | 0,006 | 5,44 | 5 | 122983 | 75000 | 73 | 8 (9%) |
| 2 | 44 | 2,90 | 0,019 | 5,44 | 5 | 122983 | 150000 | 156 | 15 (19%) |
| 2 | 43 | 2,95 | 0,018 | 5,44 | 5 | 122983 | 150000 | 145 | 16 (17%) |

**Figure 1: 2D-DIGE analyses.**

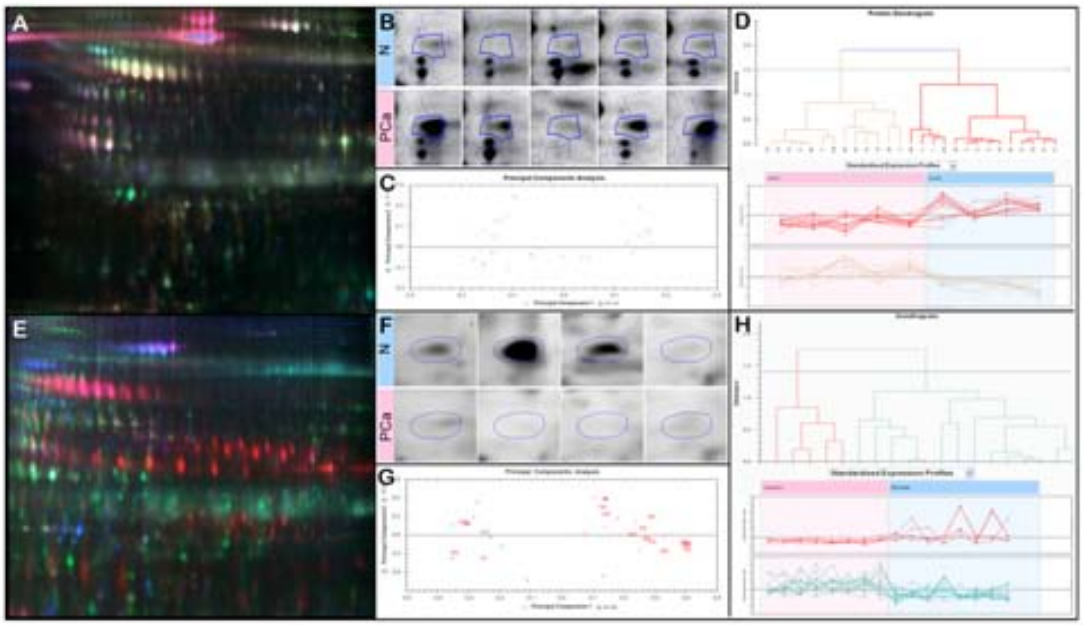



**Figure 2: Biological information of the spots identified.**

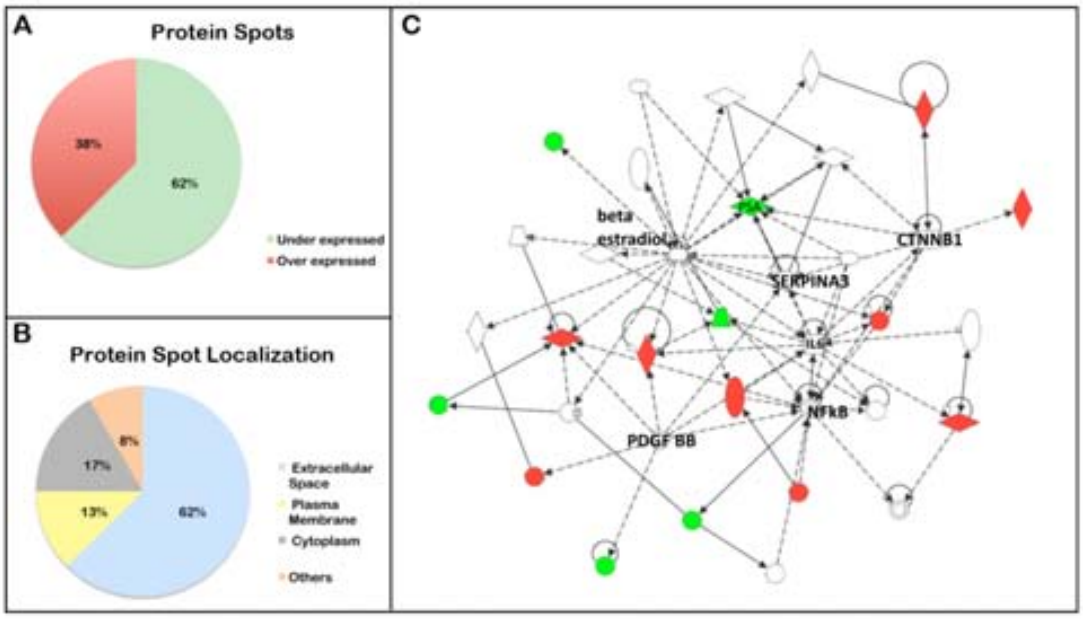
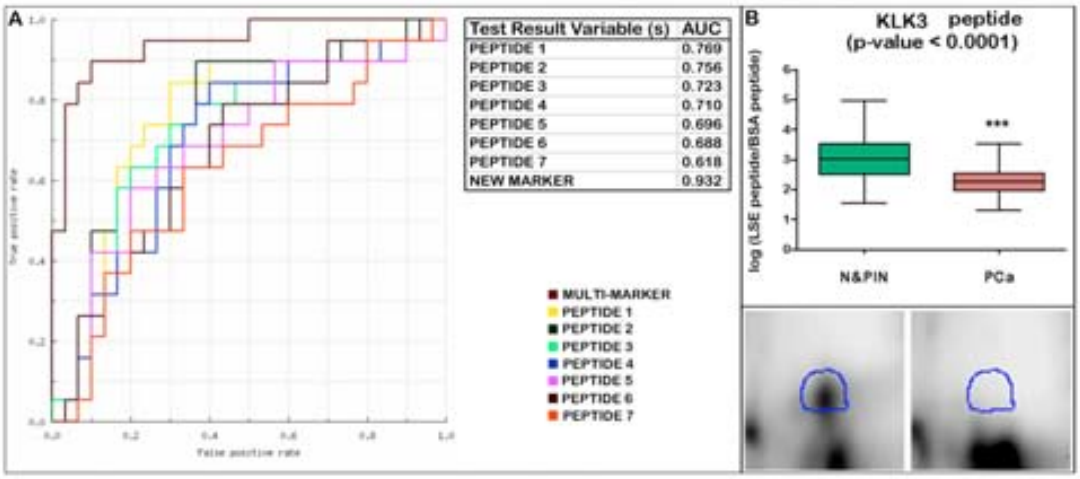
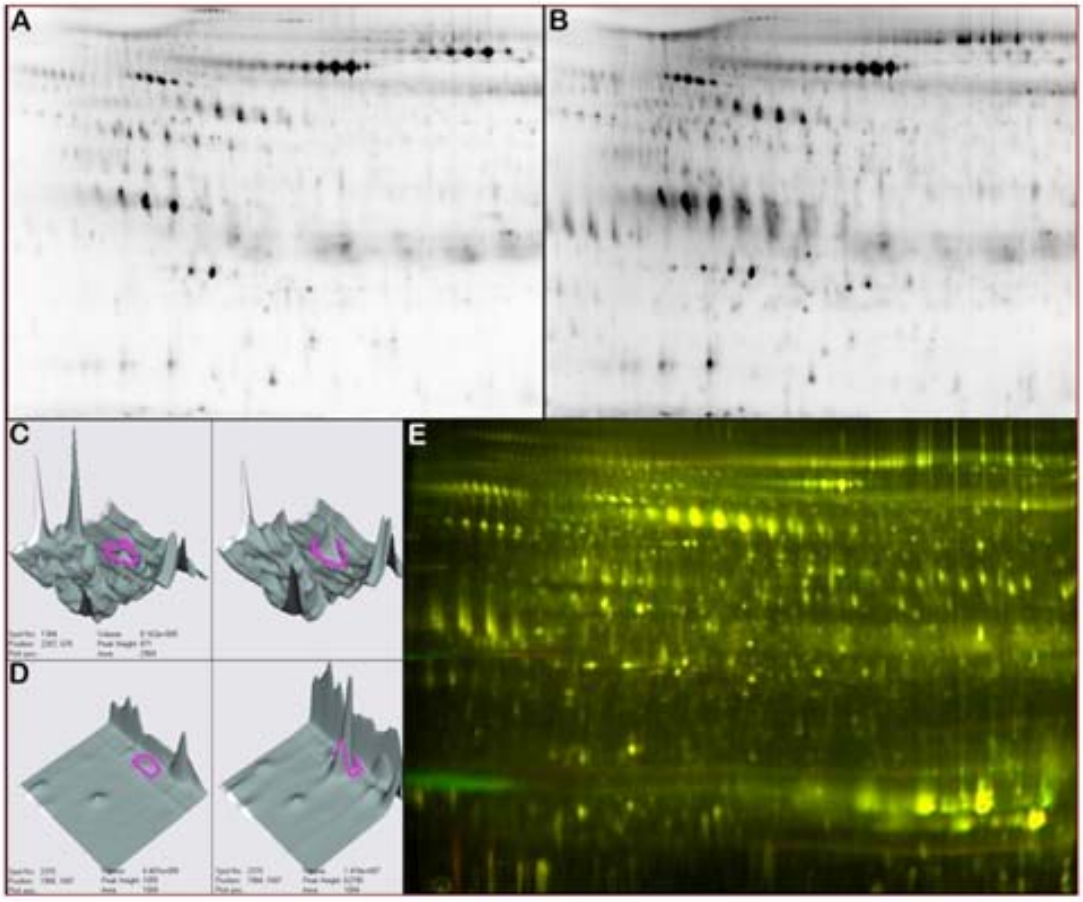


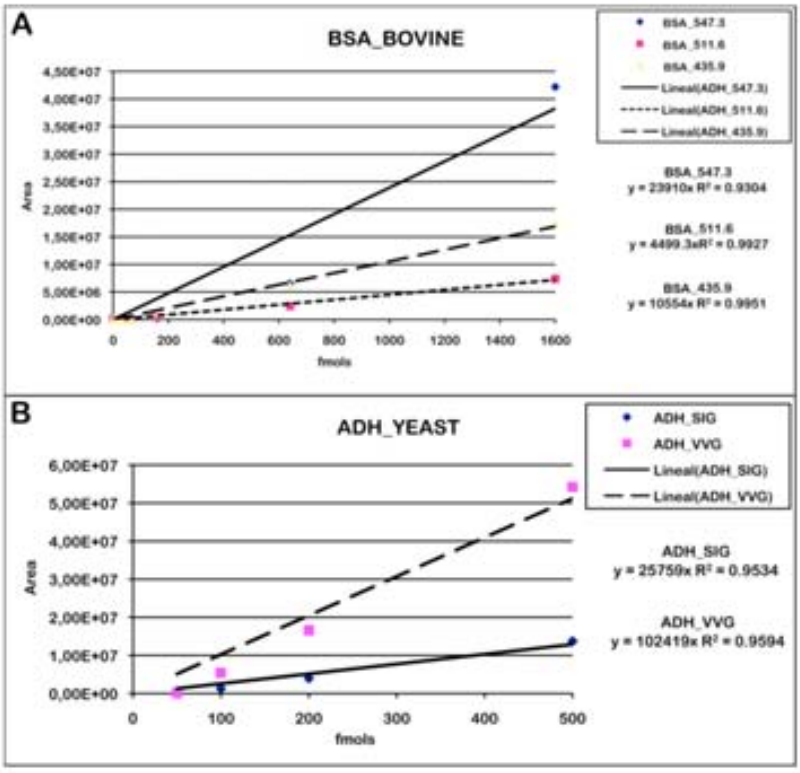
**Figure 3: Relative qualification data Analysis**




**Supplemental Figure 1: ProteoMiner set up experiments**



**Supplemental Figure 2: SRM-control peptides Standard Curves**



## Qualification and Verification of Prostate Cancer Candidate Biomarkers in Urine using Selected Reaction Monitoring Approach

Protein products of PCa cells, as well as of normal prostate cells, can be detected in urine, therefore their use as a proximal body fluid in the detection of PCa is very attractive. Urine has been defined as a liquid biopsy of the urogenital tract, and can provide information about the cancer secretome, which normally is lost in biopsy material. The qualification and verification of candidate biomarkers is a critical stage in the great biomarker discovery pipeline. Credentialed biomarkers that have successfully passed through this stage are considered verified biomarkers, which are of high value for translation into large-scale, clinical validation studies.

The evaluation of biomarkers in body fluids necessitates the development of robust methods to quantify proteins in body fluids, using large sets of samples. Selected Reaction Monitoring (SRM) is emerging as a technology that ideally complements the discovery capabilities of shotgun strategies through its unique potential for the reliable qualification and quantification of low abundance analytes in complex biological samples, such as urine samples. However, approaches to biomarker discovery using urine have been hindered by concerns about reproducibility and an inadequate standardization of proteomics protocols. Reproducible procedures for the preparation of protein samples isolated from human urine are essential for meaningful proteomic analyses.

In the present study, we performed the qualification of a set of 42 candidate biomarkers for PCa diagnosis on a set of 107 post-PM urine samples using SRM-based absolute quantification. The technique was applied to detect and quantify those proteins in urine samples. Before that, urine sample preparation and analytical procedures were optimized for SRM methodology. We standardized preparation of the urine protein samples for SRM analysis by using 9 different protocols. We “spiked” our samples with exogenous yeast and bovine proteins to monitor and optimize protein extraction and digestion conditions. Each protocol was repeated for 5 separate days by 2 different individuals.

**INTRODUCTION:**

Although prostate-specific-antigen (PSA) serum measurement in combination with a digital rectal examination (DRE) and, transrectal ultrasound-guided biopsy is currently the gold standard of care for prostate cancer (PCa) screening in Europe, PSA and DRE lacks significantly in specificity (33%) and biopsy lacks ideal sensitivity (20% false negatives). Therefore additional biomarkers are needed to supplement or potentially replace the currently used diagnostic techniques (1). Since the secreted products from both normal prostate epithelial cells and PCa cells can be detected in the urine of men, their use in the detection of PCa as proximal body fluids, which have been in contact with only one or a few tissues, is very attractive. They may represent the best compromise between a minimally invasive diagnostic technique and the possibility of obtaining enough material for a correct diagnosis.

Urinary proteomics becomes one of the most interesting sub-disciplines in the clinical proteomics area for biomarker discovery (2-4). Urine is produced by the kidneys and allows the human body to eliminate waste products from blood. Therefore, urine may contain information not only from the kidney and the urinary tract but also from distant organs via plasma obtained by glomerular filtration (5). Proteins and peptides "leaked" from tissues into clinically accessible body fluids, such as blood or urine, have led to the possibility of diagnosing the disease at an early, presymptomatic stage, as well as being able to monitor its responses to therapy through testing these body fluids for the presence of disease-related biomarkers (6). The term, "secretome," refers to the rich, complex set of molecules secreted or shed from the surface of living cells (7). Cancer cell secretomes are responsible for the "cross-talk" among cancer cells (8). Understanding this language could greatly increase our knowledge regarding the molecular mechanisms of neoplasia.

The urinary proteome has been studied by almost all proteomics technologies, but MS-based urinary protein and peptide profiling has emerged as most suitable for clinical applications. Nevertheless, biomarker discovery approaches in urine have been hindered by concerns for reproducibility and inadequate standardization of proteomics protocols (9). Reproducible procedures for the preparation of protein samples isolated from human urine are essential for meaningful proteomic analyses. Recent studies, demonstrated that organic solvent precipitation followed by in-solution digestion applied to the proteomic discovery and verification phases was showing an excellent consistency (10, 11).

The evaluation of the protein candidates is a crucial phase of the new biomarker development pipeline (6). The evaluation phase include the qualification of the biomarker candidate and the verification phase of its (6). Proteins successfully passing these stages are considered biomarkers, which are of high value to be translated into large-scale clinical validation studies (12). For the qualification and the verification phases, target proteomics such as, liquid chromatography in tandem with mass spectrometry (MS) using Selective Reaction Monitoring (SRM) represents an effective strategy (13, 14). SRM enables rapid screening of hundreds of analytes in one LC-MS/MS run and allows the detection and accurate quantification of predetermined protein sets in biological samples (15). For each protein, a set of 'best representing' peptides is determined on the basis of their uniqueness (proteotypic peptides or PTP) and performance in LC-MS analysis (16). When SRM is coupled with stable isotope dilution (SID) (15, 17), the technique allows an accurate quantitative analysis of peptides with a linear response over near five orders of magnitude down to a low ng/mL concentration (15). A common standard (a mixture of the samples to be analyzed) is labeled with heavy, stable isotopes and spiked into the individual samples. Quantification is then no longer based on the absolute signal intensity, but rather on the relative intensity of the analyte signals compared to that of the isotopically labeled internal standard. Each peptide is quantified relative to the matching, heavy labeled peptide. The isotope labeling should introduce a sufficiently large mass difference for both of the precursor ions, in order to prevent cross-talk with light transitions (13). In addition, the ability of multiplexing SRM experiments increasingly expedite the development of clinical biomarker studies (18). Thus, SRM can easily bridge the gap between biomarker discovery, usually performed on few samples, and validation by antibody-based approaches, which are costly and slow to develop (13).

In the longer term, the utility of SRM within the field of routine clinical measurements of diagnostic disease biomarkers relies on technological developments (6) but more importantly in analytical performance because of the sample complexity (15). Nonetheless, the prospects for the adoption of SRM assay as an essential element in the quantitative proteomics toolbox for high throughput screening of patients samples in the verification stage are promising (6).

The aim of this study is to determine a proteomic profile in urine supernatants that would be able to distinguish between PCa and benign samples. Previously, by 2D Differential in gel electrophoresis (2D-DIGE), coupled with Matrix-Assisted Laser Desorption Ionization - Time Of Flight MS (MALDI-TOF-MS), we identified a proteomic profile of 24 potential

biomarkers (15 down- and 9 over-expressed) for the detection of PCa in urine. Later on, those candidate biomarkers were qualified using a relatively quantification with SRM - based assay on a cohort of 50 urine supernatant samples. After logistic regression analysis, we obtained a panel of 7 peptides within 5 different proteins that were able to distinguish between PCa and benign samples. A multiplex ROC curve using those peptides showed an AUC value of 0.93 with 95% of sensitivity and 78% of specificity (Thesis annex 3, unpublished results). Here we want to conduct a second qualification of those candidate biomakers together with other promising biomarkers from PCa secretome data mining (figure 1 - Thesis annex 3). In addition, due to the large number of samples with pathological conditions, and due to the fact that many of the proteins included in this study are known to be related with prognosis we would also be able to define a perfil of prognostic biomarkers. Moreover, in order to obtain reliable data we defined a previous objective, which is to standardize a urine protein preparation protocol to be used in LC-MS/MS SRM assay. The final goal is the establishment of a simple diagnostic test, which can be used in hospitals and outpatient routines (similar to the ELISA, either multiplex or strips, reactive to the 3-5 molecules most representative of the differential profile).

## EXPERIMENTAL PROCEDURES

### 1. Urine sample preparation protocol standardization for SRM analysis:

#### *1a. Urine sample protein extraction standardization:*

Normal urine collection and storage: The experiment was performed using urine samples from 10 healthy donors (5 male and 5 female) from Vall d'Hebron Hospital, Barcelona, Spain. These healthy volunteers were 26 and 35 years old, and had no alcohol and sport 24 h prior urine collection. 50 mL of second voids of the day collected mid-stream in sterile urine collection tubs and cooled at 4°C. A complete protease inhibitor cocktail (Roche Diagnosis) was added immediately upon collection. The urines were pooled together, mixed and aliquot in 50 mL. The urines were centrifuged at 2,500 g at 4°C for 10 minutes. The supernatant were then separated and after the samples were stored at –80°C.

Before Urine protein extraction: Protein amount was determined with Micro Pyrogallol Red Method (Sigma-Aldrich, Saint Louis, MO, USA). 250 $\mu$g of total urine protein in 5 mL were used for each protein extraction protocol. 50 $\mu$g of two different yeast proteins,

Alcohol dehydrogenase 1 (ADH1) and Enolase1 (ENO1) (Sigma-Aldrich, Saint Louis, MO, USA) were added to each urine sample.

Urine Protein Extraction Protocols:

*(A) Organic Solvent Precipitation Protocol:* Urine samples were precipitated with 100% stock solutions of organic solvent (A.1-Methanol (CHROMASOLV® Plus for HPLC, Sigma-Aldrich, Saint Louis, MO, USA), A.2-Ethanol (for Molecular Biology, Merck, Darmstadt, Germany), A.3-Acetonitril (G CHROMASOLV® Plus for HPLC, Sigma-Aldrich, Saint Louis, MO, USA) and A.4-trichloroacteic acid (TCA) (Fluka, Steinheim, Germany)) at a ratio 1:5 (v/v). A.1, A.2 and A.3 were incubated at 2 different temperature conditions, room temperature (RT) and -20ºC for 16 h. A.4 was incubated only at RT for 16 h. After the precipitation urines were centrifuged at 14,000 g for 30 min at 4ºC. The pellet was washed once with the solvent used for initial precipitation, air-dried and resuspended in 500 $\mu$L of 8 M urea and 0.1 M Ammonium Bicarbonate (BA). Protein concentration was determined using Bradford Assay (Sigma-Aldrich, Saint Louis, MO, USA) with a standard curve derived from Bovine Serum Albumina (BSA). The samples were stored at -20ºC until the digestion step.

*(B) Triclor Acetic Acid (TCA) – ACETONE precipitation protocol:* Triclor Acetic Acid (TCA) (Fluka, Steinheim, Germany) was added at a ration 1:4 (v/v) in urine, vortex vigorously 1 min. and incubated at 4°C for 2 h. The samples were centrifuged twice at 14,000g for 15 min. at 4°C and the supernatant were discarded. 2.5 mL (-20°C) Acetone (CHROMASOLV® Sigma-Aldrich, Saint Louis, MO, USA), was added to the sample and was vortex vigorously during 1 min. Finally the pellet was air-dried and resuspended in 500 $\mu$L of 8 M urea and 0.1 M BA. Protein concentration was determined using Bradford Assay, as explained before. The samples were stored at -20ºC until the digestion step.

*(C) Ultrafitration protocol:* 2 M Urea and 0.05 M BA was added at a ratio 1:5 (v/v) in urine and mixed. Ultrafree 15 centrifugal filters with a 3 kDa-size exclusion membrane (Millipore, Billerica, MA, USA) were used for the ultrafiltration, following the manufacture's instruction. Each sample was added to the reservoir of the device, centrifuged at 4,000 rpm at 4ºC since the volume was reached 500 $\mu$L. Protein concentration was determined using Bradford Assay, as explained before. The samples were stored at -20ºC until the digestion step.

1 Dimensional in Gel Electrophoresis (1D-GE): Protein extracted sample were loaded with Loadding Buffer 1 X in a PreCast Gel of 4-12% Acrylamide (NUPAGE ® Novex ® Bis-Tris Mini Gels, Invitrogen). The gels were stained with silver staining during 30 min.

Protein digestion protocol: After the protein extraction 10 $\mu$g of protein of each protocol resulting sample was digested. The samples were reduced with 40 mM 1.4-Dithioerythritol (DTT) in 50 mM BA at 37ºC at 800 rpm during 30 min. After that the samples were alkylated with 125 mM iodoacetamide (IAA) in 50 mM BA at 37ºC and 800 rpm during 30 min. Samples were then digested with trypsin (Promega, Madison, WI, USA) using a ratio 1:10 (w/w) at 37ºC overnight. Digestion was stopped by adding formic acid (FA) 0.25% (Bioultra, Fluka, Steinheim, Germany) since obtain a pH 2-3. Samples were preserved at -20ºC until the reverse phase procedure.

Reverse C18-Clean up: Sep-Pak C18 reverse phase cartridges, 100 mg (Waters, Milford, MA, USA) were used to CleanUp the samples after the protein digestion and before the LC-MS analysis. The cartridges were wetted with 1 mL of Methanol, cleaned with 1 mL 80% Acetonitril in HPLC-grade water and equilibrated with 4 mL 0.1% FA in HPLC-grade water. The flow-through was discarded after each step. The samples were loaded into the cartridge and the flow-through was reloaded in a second step. The cartridges were washed with 6 mL 0.1% FA in HPLC-grade water. Finally, the peptides were eluted using 1 mL Acetonitrile 50% and FA 0.1% and concentrated in the speed vaccum (Thermo Fisher, San Jose, CA, USA). Finally the samples were dissolved in 50 $\mu$L FA 0.01% and stored until the LC-MS analysis.

Absolute Quantification Analysis by SRM-based assay: SRM methods for ADH1 and ENO1 yeast proteins were performed using Pinpoint software 1.0 (Thermo Fisher, San Jose, CA, USA). Three peptides for each protein were monitored. The corresponding HeavyPeptide™ Crude (Thermo Fischer Scientific, Ulm, Germany) for all these peptides was synthesized. 200 fmol of each ILP peptide were spiked into the sample. TSQ Vantage (Thermo Fisher, San Jose, CA, USA) was used to analyze (200 ng of urine samples with the correspondent added ILP). The samples were run in duplicate values.

Data analysis: The SRM data were processed using Pinpoint v1.0 Software. For to five SRM transitions, were monitored for both endogenous (light) and the internal standard (heavy) peptides. The area under each transition for each peptide was calculated by using Pinpoint. Data were processed by integrating the appropriate peaks in an extracted ion chromatogram for the light and the heavy peptide, followed by calculation of the ratio between [Area (exogenous peptide)/Area (ILP)] was used to quantify the amount of the exogenous peptide in each sample.

***1b. Urine Sample Protein Digestion Protocol Standardization***

Time and Ratio Trypsin digestion setting: After protein extraction, a pool of 20 μg of urine protein with 2 pmol of ADH1 was prepared. The different samples were reduced with 40 mM DTT and alkylated with 125 mM IAA. Then the samples were digested in different time conditions (1h, 3h, 6h and 20 h) and different ratios of μg Trypsin:μg Urine protein (1:10, 1:20 and 1:50). Triplicate samples for each condition were digested. Digestion was halted by adding 1 μL of 0.25% FA in HPLC water. 5 pmols of Bovine Serum Albumin (BSA) peptides (Sigma-Aldrich, Saint Louis, MO, USA) were added to each sample to be use as an internal standard. The digested samples were diluted in 0.1 % FA to a final concentration of 0.5μg/μL. Samples were preserved at -20°C until analyzed by LC-MS.

Methods for 4 different peptides for ADH1 and 2 different peptides for BSA were monitored using an SRM-based assay. 2 μg of each sample were run in triplicate values on a Tempo Nano LC System (AB SCIEX, LLC, Dublin, CA, USA) in tandem with a Hybrid triple Quadrupole-Linear Ion Trap Mass Spectrometer 4000 Q TRAP® LC/MS/MS (AB SCIEX, LLC, Dublin, CA, USA). The area under each transition for each peptide was calculated by Analyist software System (AB SCIEX, LLC, Dublin, CA, USA). The ratio between ADH1 (peptide EALDFFARA) / BSA (peptide HLVDEPQNLIK) was used to relatively quantify the amount of ADH protein.

N-terminal alkylation of peptides: After the protein extraction, 20 μg of protein was digested using the following conditions: Samples were reduced with DTT in 50 mM BA at 37°C and 800 rpm during 30 min. and alkylated with IAA in 50 mM BA at 37°C and 800 rpm during 30 min. A quenching step was performed, only when was required, by adding N-Acetyl Cysteine (NAC) in 50mM BA at 37°C and 800 rpm during 30 min. The amount of DTT, IAA and NAC for the different conditions was: (A1) 40 mM DTT/125 mM IAA/156 mM NAC, (A2) 40 mM DTT/125 mM IAA, (B1) 12.5 mM DTT/ 40 mM IAA / 50 mM NAC and (B2) 12.5 mM DTT/ 40 mM IAA. Samples were then digested with trypsin (Promega, Madison, WI, USA) using a ratio of 1:10 (w/w), 37°C overnight and 800 rpm, finally, the digestion procedure was stopped by adding formic acid (FA) 0.25% since obtain a pH 2-3. Samples were preserved at -20°C until analyzed by LC-MS.

SRM methods for ADH1 and ENO1 yeast proteins were performed using Pinpoint software 1.0. Three peptides for each protein were monitored. The corresponding HeavyPeptide™ Crude (Thermo Fischer Scientific, Ulm, Germany) for all these peptides was synthesized. 200 fmol of each ILP peptide were spiked into the sample. TSQ

Vantage was used to analyze 50 ng of urine samples with the correspondent added ILP. The samples were run in duplicate values. Two independent samples from each protocol and triplicate values for each sample were run in a TSQ-Vantage using SRM-based assay. The area under each transition for each peptide was calculated by using Pinpoint. Data were processed by integrating the appropriate peaks in an extracted ion chromatogram for the light and the heavy peptide, followed by calculation of the ratio between [Area (exogenous peptide)/Area (ILP)] was used to quantify the amount of the exogenous peptide in each sample.

**2. Absolute Quantification Analysis:**

Patients and urine collection: This study obtained approval from the institutional review board. All urine samples were obtained from the Department of Urology of the Vall d'Hebron Hospital in Barcelona, Spain. Urine samples were collected from 107 consecutive men, who had been referred for prostate biopsy (PB), immediately post-prostate massage (PM). The indications for PB were an abnormal DRE and/or serum PSA levels higher than 4 ng/mL. Patients with other known tumors and/or previous PCa therapies were excluded from the study. Written informed consent was obtained from all patients. The study population consisted of 107 men, 36 (34%) were positive for PCa and 71 (66%) were negative (benign controls without cancer). The clinical and pathological information data for these 107 patients are shown in table 1.

*PM methodology:* PM was performed by systematically applying severe digital pressure to the prostate from the base to the apex and from the lateral to the median line of each lobe. It is known that through the practice of this technique, the desquamation of a large quantity of products of tumor origin is induced. These products were present in the urine samples and were able to be gathered in the micturition.

*PB methodology:* The biopsies were performed using an end-fire ultrasound transducer (Falcon 2101, B-K Medical, Inc.) and an automatic 18gauge needle (Bard, Inc.). The minimum number of cores removed in every procedure was 10, and between 1 and 8 additional cores were removed, according to the Vienna nomogram (19).

The biological samples collected from selected patients were the first 50mL of urine gathered in the first micturition produced after PM. Once the samples had been collected in the receptacles provided for that purpose, the receptacles were closed and stored on ice. Samples were then transported directly to the laboratory for processing within 30 min of their collection. Urine samples were centrifuged at 2500 g at 4°C for 10 min. The

supernatants were separated and protease inhibitors (Sigma-Aldrich, Saint Louis, MO, USA) were added before to store the samples at -80°C.

Urine protein extraction and digestion: 10 mL of urine supernatant was quantified using Micro Pyrogallol Red Method. 250 $\mu$g in 5 mL final volume were precipitated with 100 % stock solutions of acetonitrile at a ratio 1:5 (v/v). Samples were incubated over night at RT. After the precipitation step, samples were centrifuged at 14,000 g for 30 min at 4°C. The pellet was washed with acetonitrile, air-dried and resuspended in 500 $\mu$L of 8 M Urea and 0.1M BA. Protein concentration was determined using Bradford Assay with a standard curve derived from BSA. The samples were stored at -20°C until the digestion step. 20 $\mu$g of protein were digested. The samples were reduced with 12.5 mM DTT in 50 mM BA at 37°C and 800 rpm during 30 min and after, alkylated with 40 mM IAA in 50 mM BA at 37°C and 800 rpm during 30 min. IAA was quenched with 50 mM NAC in 50 mM BA at 37°C and 800 rpm during 30 min. Volume samples were adjusted since reached the concentration of 1 M urea in 100 mM BA. Samples were then digested with trypsin using a ratio of 1:10 (w/w). The trypsin digestion was carried out at 37°C overnight. Digestion procedure was stopped by adding formic acid (FA) 0.25 % since obtain a pH 2-3. Samples were preserved at -20°C until the following step.

Reverse C18-Clean up: Sep-Pak C18 reverse phase cartridges 100 mg were used to clean and desalt the samples after the protein digestion and before the LC-MS analysis. The samples were processed using the same procedure explained before. Finally, the samples were dissolved in 50 $\mu$L FA 0.01% and stored until the LC-MS analysis.

In silico analysis to select the best PTPs: An *in silico* digestion was performed for each targeted proteins (to know the theoretical peptides that belongs to these proteins). These peptides were matched within an *in silico* Human Proteome digestion database to select the PTP (uniqueness in the Human Proteome). One to three PTP were selected for each targeted protein, in according to the following criteria: i) It must be a tryptic peptide (which means that the proteolysis is performed by trypsin), ii) the amino acid sequence must not contain chemically reactive residues, i.e., cystein and methionine residues and a limited number of tryptophan, iii) the length of peptides must be between 8 and 20 amino acids in anticipation of the synthesis, iv) the peptides selected must present good ionization properties for the mass spectrometry experiments and v) the hydrophobicity based on the Krokhin model (20) is estimated for the LC-MS experiments.

Absolute quantification SRM-based assay: SRM methods were created using Pinpoint software 1.0 (data not shown). In total 98 peptides corresponding to 42 different proteins

were selected. The corresponding HeavyPeptide™ Crude for these peptides was synthesized. The peptides were selected from a combination of different sources as, Peptide Atlas, *in silico* digestion and previous data from LC-MS/MS experiments using a pool of urine sample in order to select those peptides that were more represented in the urine sample. 50 fmol/μL of each labeled peptide were spiked into the sample. TSQ Vantage was used to analyze 400 ng/μL of urine sample.

Data analysis: The SRM data were processed using Pinpoint v1.0 Software. Four to five SRM transitions, were monitored for both endogenous (light) and the internal standard (heavy) peptides. The area under each transition for each peptide was calculated by using Pinpoint. Data were processed by integrating the appropriate peaks in an extracted ion chromatogram for the light and heavy peptide, followed by calculation of the ratio of peaks areas multiplied by the absolute amount of internal standards (50 fmols).

## RESULTS

### 1. Urine sample preparation protocol standardization:

#### *1.a. Urine protein extraction protocol standardization:*

Total amount of protein recovery: A pool of urine samples was used to compare the extracted the protein content within 9 different protocols. Each protocol was repeated 5 different days and within 2 different persons (1 and 2) (10 replicates per urine protein protocol extraction) (Figure 1). The protein amount after the protein extraction protocol was measured by using Bradford assay and the ratio between experimental and theoretical amount of protein was calculated. Average values, standard deviation and coefficient of variation (CV) values for each protein sample extraction protocol were calculated between the five replicates. The highest amount of total protein recovery after protein extraction was obtained using ultrafiltration protocol. In contrast, the lowest amount was obtained using TCA protocol. Methanol and ethanol (both, RT and -20ºC temperature conditions) were giving a similar amount of protein recovery. However, using acetonitril, as organic solvent for precipitation, the protein recovery was lower. A big difference using different temperature conditions was observed. RT was giving the double amount of protein recovery than -20ºC condition (Table 2).

Protein distribution pattern analysis: One replicate from each protein extraction protocol was used to perform a 1D-GE to look for the protein distribution pattern of each sample. 10 μg of each sample were loaded in 1D-GE. The Electrophoresis was run at 150 mA

per gel and 180V during 1h. In general the protein distribution pattern obtained was similar for all the protocols analyzed, only TCA protocol was showing a different pattern (Figure 2).

Optimization of SRM experiment for yeast exogenous proteins and yeast ILP: Standard curves for exogenous and labelled peptides for both, ADH1 and ENO1 yeast proteins were run in duplicate. Three peptides for each protein were monitored using a SRM-based assay in a TSQ Vantage (Figure 3).

Yeast protein absolute quantification by SRM-based assay: Replicate 1$^{st}$, 2$^{nd}$, 3$^{rd}$ and 4$^{th}$ of one manipulator were used to quantify ADH1 and ENO1 yeast exogenous peptides. 250 fmols of ILP were spiked into each sample. TSQ vantage was used to absolutely quantify those peptides using a SRM-based assay. Acetonitril precipitation, both RT and -20ºC conditions, were giving a high amount of exogenous yeast protein (Figure 4).

***1.b. Urine protein digestion protocol standardization:***

ADH1 was quantified by SRM assay (peptide EALDFFARA) / BSA (peptide HLVDEPQNLIK) under different digestion conditions of time (1h, 3h, 6h and 20h) and ratio μg Trypsin:μg Urine protein (1:10, 1:20 and 1:50). 1:10 ratio Trypsin:protein and the overnight digestion condition were giving the best signal in terms of relative amount of peptide (Figure 5a).

Regarding the N-terminal peptide derivatization problem; after an SRM-based absolute quantification assay, using ADH1 and ENO1 proteins/ILP, the digestion condition of 12.5 mM DTT/ 40 mM IAA / 50 mM NAC was giving the highest amount of protein (Figure 5b).

**2. Absolute Quantification Analysis:**

Once the previously identified proteins had been defined (by SRM-relative quantification, see thesis annex 3) and the sample preparation protocol had been selected the study was moved on to the second qualification step, for which we would like to perform an absolute quantification SRM-based assay on those selected candidate biomarkers on a larger cohort of urine samples.

11 proteins belonging to the discovery phase (see thesis annex 3) were selected. Five out of those 11 proteins were selected from the relative quantification assay. From literature, data mining of PCa secretome, the number of protein candidates was increased to 31 proteins. In total, 42 proteins candidate biomarkers were absolutely

quantified using SRM-based assay in a cohort of 107 urine samples, 36 (34%) out of 107 presented PCa (Table 1).

For all proteins, their corresponding signature peptides were selected, based on their proteotypicity, i.e. specific to the protein. In addition, the existing proteomic information about predicted PTPs presenting adequate MS properties reported in Peptide Atlas (21) has been studied. One to three peptides were selected for each targeted protein. In total 98 peptides, corresponding to 42 different proteins, had been selected. The best 3 to 4 transitions for each peptide were selected to generate the SRM method (data not shown). The selection of the optimal SRM transitions is essential for the success on the analysis. In order to generate a robust SRM method the corresponding labeled/heavy peptides (C-terminal $^{15}N$-, and $^{13}C$-labeled arginine or lysine) were synthesized for each protein for these target analytes in urine samples. In order to select the amount of heavy peptide that we had to add to the samples, dilution series of each heavy peptide were prepared on 200 ng of urine digest. Each dilution series correspond to triplicate values of 7 different amounts of spiked heavy peptide from 0 to 150 fmol. Fifty fmol/$\mu$L of each labeled peptide were spiked into the each sample to be analyzed.

Duplicate values of SRM analysis were generated for both heavy and light peptide from each sample (Figure 6). Pinpoint v1.0 Software was used to process all these data. The amount of the endogenous (light) peptide was calculated using one point quantification [Area endogenous pick peptide/Area labeled pick peptide]. Revision and statistical analysis on the obtained data is still on going.

## DISCUSSION

The contribution of proteomics to the understanding of the pathogenesis and diagnosis of diseases from the genitourinary tract has been considerable (22). In the past few years the proteomics field has moved onward to clinical applications, particularly for biomarker discovery, diagnostic and prognostics of human diseases (23). Differential proteomics is a scientific discipline that identifies the proteins associated with a disease by means of their altered levels of expression (24) (25) between the control and the disease states. Indeed, biomarker discovery in human tissue and body fluids is clearly one of the areas with enormous potential. Clinical proteomics has much more promise in medicine with new types of proteomic technologies combined with advanced bioinformatics currently being used to identify molecular signatures of diseases (6).

However, without proper study design and implementation of robust analytical techniques, the efforts and expectations may very easily be hampered.

MS-based proteomics strategies have become an integral part of the biomarker development workflow, including a discovery phase and the subsequent qualification, verification, and validation of the candidates in body fluids (25). One of the major bottlenecks in the process of identifying a novel biomarker is the qualification and verification of candidate proteins (26). At this stage of any biomarker discovery pipeline, it must be proven that a specified characteristics of the candidate biomarker correlates with the disease scenario (26). Normally, the evaluation of biomarkers in body fluids necessitates the development of robust methods to quantify proteins in complex background (18). The use of target proteomics by SRM may be used to achieve the precise quantification of a specific group of proteins. One of the advantages of this technique is the possibility of quantifying a great number of peptides after only one single LC-MS/MS experiment. However, while SRM shows considerable promise for protein quantification, the technique is still in its early stages for clinical applications (6). To obtain statistically significant data, expanding the number of analyzed components requires an increase in the number of samples analyzed and, consequently, greater computing power (6). Every sample manipulation increases the possibility of introducing artifacts, reduces reproducibility, and may further increase the complexity of samples (22). Thus, reproducible procedures for the preparation of protein samples isolated from human urines are essential for meaningful proteomic validation analyses.

Recent studies in this field have concluded that organic solvent precipitation, followed by in-solution digestion provide the best performance for urine proteome analysis in the discovery and verification phases (9, 10). Normally the use of clinical sample is limiting the amount of starting material and thus, the use of at least 20 $\mu$g of protein was defined in order to obtain high confidence in the following proteomic analysis (9). After protein extraction and prior to MS analysis, proteins need to be broken up into peptides by enzymatic digestion. An effective digestion procedure includes a denaturation step, a reduction of disulfide linkage, an alkylation of sulfhydryl groups to prevent reformation of disulfide bonds, and finally a digestion. Incomplete digestion, due to a high concentration of matrix proteins or other interfering compounds, can introduce many points of variation and adversely affect the final quantitative accuracy (27).

Here, we followed the same strategy as these groups, in order to compare, not only the different protein extraction procedures, such organic solvent precipitation (using different

temperatures) and ultrafiltration procedures (which we used in previous experiments but also to optimize the trypsin-based protein digestion that would be used in the quantification SRM-based assays. We concluded that, the utilization of organic solvent precipitation, such as acetonitril precipitation at RT, was giving us the best exogenous protein (ADH1- and ENO1 yeast) recovery, in terms of absolute quantification by SRM analysis, and also give a good total amount of protein, in terms of Bradford quantification (Table 2 and Figure 4). Both points are important when using urine samples, which generally have a low amount of protein. Moreover, using a ratio 1:10 (protein: trypsin) was giving the best levels of quantification using SRM based assay (Figure 5a). Finally, since any quantification studies will be affected by variable signal splitting due to any peptide modification, it is important to avoid peptide N-terminal IAA overalkylation, that is taking place in the overnight digestion where excess of alkylation reagent remains in the medium (28). Therefore, incubation step with NAC, prior to enzymatic digestion, avoided N-terminal IAA over-alkylation and, thus, helped to obtain better quantification results (Figure 5b).

PCa cells are secreting different proteins, which can be found in urine. As prostate massage leads to an enrichment of prostatic fluid in the first urine catch after this, urine is one of the ideal clinical samples for biomarker discovery in PCa. Moreover, the urine is readily available in almost all patients, and its collection is very simple and non-invasive. Urinary proteomics thus becomes one of the most interesting sub-disciplines in the clinical proteomics area (3). Our aim was to qualify a list of candidate biomarkers for PCa using SRM assay. From previous studies (see thesis annex 3) we were able to define a list of 11 protein biomarker candidates that need further qualification. Those proteins together with other promising proteins from the literature were selected. Here we conduct a qualification phase of 42 proteins using single point quantification based SRM assay in 107 PM-urine samples. The absolute amount of the targeted protein can be deduced from the relative intensity of the light/heavy transitions. The relative quantity of a specific analyte is then determined from the relative signal intensity of the signal in the full spectrum (13). In order to better select the best candidates further analysis on the obtained data is required; (i) It's important to identify targeted peptides, for which the ion fragmentation pattern obtained for the heavy form is different from the obtained from the light form. (ii) It is important to verify if the amount of the heavy peptide that was spiked in each sample is corresponding to the amount of 50fmol, using the standard curve quantification. (iii) Moreover, thanks to this standard curve we will be able to assess the Limit of detection (LOD) and Limit of quantification (LOQ) values for each peptide used in

the analysis. Those analyses are still ongoing; thus, no final conclusions have been included in this study at present. However, taking together all these facts it is highly probable that we will be able to identify those peptides that are giving a good signal and therefore an accurate quantification result. Once we have been reviewed all these points we will be ready to perform statistical analysis on this data, such as feature selection to select the best peptides able to differentiate between disease and control state. Since most MS based proteomic studies suggest that a single biomarker does not provide sufficient specificity, several biomarkers are often combined. This requires a two-step process, whereby in the first step biomarkers are defined, and in the second step these are combined in a predictive model (22). At the end would like to obtain a panel of biomarkers that alone or in combination with the existing PCa biomarker (serum PSA) will help us to better define patients with PCa. This constitutes an important step towards advancing its accurate diagnosis, which currently represents a setback in our ability to cure patients of PCa.

**BIBLIOGRAPHY**

1. Ryan, C. J., and Small, E. J. (2005) Progress in detection and treatment of prostate cancer. *Curr Opin Oncol* 17, Page.
2. Pisitkun, T., Johnstone, R., and Knepper, M. A. (2006) Discovery of urinary biomarkers. *Mol Cell Proteomics* 5, Page.
3. Decramer, S., Gonzalez de Peredo, A., Breuil, B., Mischak, H., Monsarrat, B., Bascands, J. L., and Schanstra, J. P. (2008) Urine in clinical proteomics. *Mol Cell Proteomics* 7, Page.
4. Thongboonkerd, V. (2007) Recent progress in urinary proteomics. *Proteomics Clin Appl* 1, Page.
5. Thongboonkerd, V., and Malasit, P. (2005) Renal and urinary proteomics: current applications and challenges. *Proteomics* 5, Page.
6. Boja, E., Hiltke, T., Rivers, R., Kinsinger, C., Rahbar, A., Mesri, M., and Rodriguez, H. (2010) Evolution of Clinical Proteomics and its Role in Medicine. *J Proteome Res*, Page.
7. Makridakis, M., and Vlahou, A. (2010) Secretome proteomics for discovery of cancer biomarkers. *J Proteomics*, Page.
8. Pavlou, M. P., and Diamandis, E. P. (2010) The cancer cell secretome: A good source for discovering biomarkers? *J Proteomics*, Page.
9. Afkarian, M., Bhasin, M., Dillon, S. T., Guerrero, M. C., Nelson, R. G., Knowler, W. C., Thadhani, R., and Libermann, T. A. (2010) Optimizing a proteomics platform for urine biomarker discovery. *Mol Cell Proteomics*, Page.
10. Court, M., Selevsek, N., Matondo, M., Allory, Y., Garin, J., Masselon, C. D., and Domon, B. Toward a standardized urine proteome analysis methodology. *Proteomics* 11, Page.
11. Afkarian, M., Bhasin, M., Dillon, S. T., Guerrero, M. C., Nelson, R. G., Knowler, W. C., Thadhani, R., and Libermann, T. A. Optimizing a proteomics platform for urine biomarker discovery. *Mol Cell Proteomics* 9, Page.
12. Rodriguez, H., Rivers, R., Kinsinger, C., Mesri, M., Hiltke, T., Rahbar, A., and Boja, E. Reconstructing the pipeline by introducing multiplexed multiple reaction monitoring mass spectrometry for cancer biomarker verification: an NCI-CPTC initiative perspective. *Proteomics Clin Appl* 4, Page.
13. Lange, V., Picotti, P., Domon, B., and Aebersold, R. (2008) Selected reaction monitoring for quantitative proteomics: a tutorial. *Mol Syst Biol* 4, Page.
14. Ong, S. E., Foster, L. J., and Mann, M. (2003) Mass spectrometric-based approaches in quantitative proteomics. *Methods* 29, Page.
15. Gallien, S., Duriez, E., and Domon, B. Selected reaction monitoring applied to proteomics. *J Mass Spectrom* 46, Page.
16. Kiyonami, R., Schoen, A., Prakash, A., Peterman, S., Zabrouskov, V., Picotti, P., Aebersold, R., Huhmer, A., and Domon, B. (2010) Increased selectivity, analytical precision, and throughput in targeted proteomics. *Mol Cell Proteomics*, Page.

17. Bordeerat, N. K., Georgieva, N. I., Klapper, D. G., Collins, L. B., Cross, T. J., Borchers, C. H., Swenberg, J. A., and Boysen, G. (2009) Accurate quantitation of standard peptides used for quantitative proteomics. *Proteomics* 9, Page.

18. Selevsek, N., Matondo, M., Carbayo, M. S., Aebersold, R., and Domon, B. Systematic quantification of peptides/proteins in urine using selected reaction monitoring. *Proteomics* 11, Page.

19. Remzi, M., Fong, Y. K., Dobrovits, M., Anagnostou, T., Seitz, C., Waldert, M., Harik, M., Marihart, S., Marberger, M., and Djavan, B. (2005) The Vienna nomogram: validation of a novel biopsy strategy defining the optimal number of cores based on patient age and total prostate volume. *J Urol* 174, Page.

20. Krokhin, O. V., Craig, R., Spicer, V., Ens, W., Standing, K. G., Beavis, R. C., and Wilkins, J. A. (2004) An improved model for prediction of retention times of tryptic peptides in ion pair reversed-phase HPLC: its application to protein peptide mapping by off-line HPLC-MALDI MS. *Mol Cell Proteomics* 3, Page.

21. Deutsch, E. W., Lam, H., and Aebersold, R. (2008) PeptideAtlas: a resource for target selection for emerging targeted proteomics workflows. *EMBO Rep* 9, Page.

22. Theodorescu, D., and Mischak, H. (2007) Mass spectrometry based proteomics in urine biomarker discovery. *World J Urol* 25, Page.

23. Thongboonkerd, V. (2008) Urinary proteomics: towards biomarker discovery, diagnostics and prognostics. *Mol Biosyst* 4, Page.

24. Etzioni, R., Urban, N., Ramsey, S., McIntosh, M., Schwartz, S., Reid, B., Radich, J., Anderson, G., and Hartwell, L. (2003) The case for early detection. *Nat Rev Cancer* 3, Page.

25. Rifai, N., Gillette, M. A., and Carr, S. A. (2006) Protein biomarker discovery and validation: the long and uncertain path to clinical utility. *Nat Biotechnol* 24, Page.

26. Elschenbroich, S., and Kislinger, T. Targeted proteomics by selected reaction monitoring mass spectrometry: applications to systems biology and biomarker discovery. *Mol Biosyst* 7, Page.

27. Engmann, O., Campbell, J., Ward, M., Giese, K. P., and Thompson, A. J. Comparison of a protein-level and peptide-level labeling strategy for quantitative proteomics of synaptosomes using isobaric tags. *J Proteome Res* 9, Page.

28. Boja, E. S., and Fales, H. M. (2001) Overalkylation of a protein digest with iodoacetamide. *Anal Chem* 73, Page.

## FIGURE LEGENDS

**Table 1. Clinical and pathological information of the samples**

**Table 2. Total amount of protein recovery after urine sample preparation.** Average, standard deviation (STDEV) and coefficient of variance (CV) values were obtained from daily and manipulation replicates (handling).

**Figure 1. Urine protein extraction standardization protocol workflow.** Experiment workflow to compare the different protein extraction protocols using a pool of healthy human urine using SRM based assay.

**Figure 2. Protein pattern distribution analysis by 1D-GE.** Protein distribution pattern of the different urine protein sample protocols.

**Figure 3. Standard curves for exogenous and labeled peptides (ILP) from ADH1 and ENO1.** (A) Standard curve for three peptides of ADH1 (exogenous) in duplicate values. (B) Standard curve for three peptides of ENO1 (exogenous) in duplicate values. (C) Standard curves for three peptides of ADH1 (labelled) in duplicate values. (D) Standard curve for three peptides of ENO1 (labelled) in duplicate values.

**Figure 4. Sample preparation absolute quantification by SRM-based assay.** The different sample preparation protocols were performed 4 times (1st, 2nd, 3rd and 4$^{th}$ replicates). The sum of the absolute amount of the three peptides monitored, for each protein (ADH1- and ENO1-yeast) were used to compare the different protocols.

**Figure 5. Urine protein digestion standardization for SRM assay.** (A) SRM-bassed Relative quantification assay of ADH1 (peptide EALDFFARA) / BSA (peptide HLVDEPQNLIK) followed under different digestion conditions of of time (1h, 3h, 6h and 20h) and ratio μg Trypsine:μg Urine protein (1:10, 1:20 and 1:50). (B) N-terminal peptide derivatization; SRM-based Absolute quantification assay using ADH1 and ENO1 proteins/labeled peptides resulting of the following digestion settings (A1) 40mM DTT/125mM IAA/156 mM NAC, (A2) 40mM DTT/125mM IAA, (B1) 12.5 mM DTT/ 40 mM IAA / 50 mM NAC and (B2) 12.5 mM DTT/ 40 mM IAA.

**Figure 6. Example of SRM transitions for both, heavy and light peptide for a specific protein in PCa and benign urine sample.** KLK3 SRM transitions for a specific endogenous and the corresponding ILP peptide.

**Table 1: Clinical and pathological information of the samples**

| | All patients average | (min - max) |
|---|---|---|
| No of patients | 107 | |
| Age (yr) | 69 | (49 - 107) |
| PSA level (range) | 11 | (1 - 195) |
| PSA free | 2.03 | (0.36 - 50) |
| Ratio free PSA/tPSA | 0.28 | (0 - 10.58) |
| Prostate cancer | 36 (34%) | |
| Gleason < 7 | 12 (33%) | |
| Gleason = 7 | 10 (28%) | |
| Gleason > 7 | 14 (39%) | |
| Benign | 71 (66%) | |

**Table 2.Total amount of protein recovery after urine sample preparation.**

| Protocol | Sample | Handling | Averge (µg/mL) Technical Replicate | STDEV Technical Replicate | CV Technical Replicate | Average (µg/mL) Handling | STDEV Handling | CV Handling |
|---|---|---|---|---|---|---|---|---|
| Methanol RT | A1.1 | 1 | 84.383 | 15.901 | 18.844 | | | |
| | A1.1 | 2 | 83.214 | 30.870 | 37.097 | 83.799 | 23.158 | 27.635 |
| Methanol -20°C | A1.2 | 1 | 105.449 | 22.017 | 20.879 | | | |
| | A1.2 | 2 | 91.508 | 40.974 | 44.776 | 98.478 | 31.868 | 32.361 |
| Ethanol RT | A2.1 | 1 | 93.308 | 30.814 | 33.024 | | | |
| | A2.1 | 2 | 101.739 | 28.434 | 27.948 | 97.523 | 28.303 | 29.022 |
| Ethanol -20°C | A2.2 | 1 | 96.643 | 37.415 | 38.715 | | | |
| | A2.2 | 2 | 103.347 | 44.956 | 43.500 | 99.995 | 39.152 | 39.154 |
| ACN RT | A3.1 | 1 | 48.555 | 17.792 | 36.643 | | | |
| | A3.1 | 2 | 52.384 | 9.588 | 18.304 | 50.469 | 13.624 | 26.995 |
| ACN -20°C | A3.2 | 1 | 22.804 | 5.162 | 22.636 | | | |
| | A3.2 | 2 | 32.414 | 10.146 | 31.302 | 27.609 | 9.124 | 33.047 |
| TCA | A4 | 1 | 7.699 | 2.510 | 32.599 | | | |
| | A4 | 2 | 8.283 | 0.884 | 10.668 | 7.991 | 1.800 | 22.530 |
| TCA-Acetone | B | 1 | 38.152 | 13.914 | 36.471 | | | |
| | B | 2 | 34.516 | 23.773 | 68.877 | 36.334 | 18.138 | 49.919 |
| Ultrafiltration | C | 1 | 107.714 | 28.037 | 26.030 | | | |
| | C | 2 | 112.886 | 16.904 | 14.975 | 110.300 | 21.996 | 19.942 |

**Figure 1. Urine protein extraction standardization protocol workflow.**

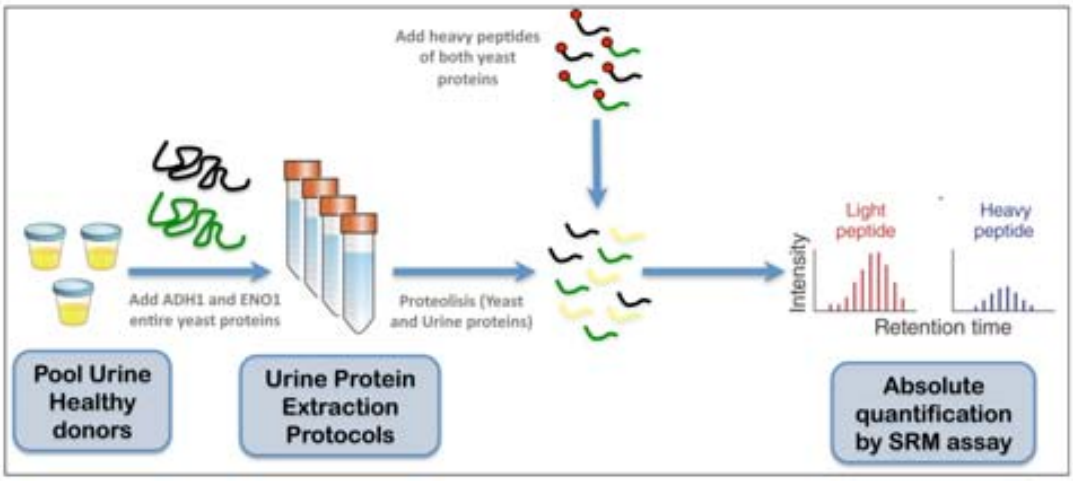
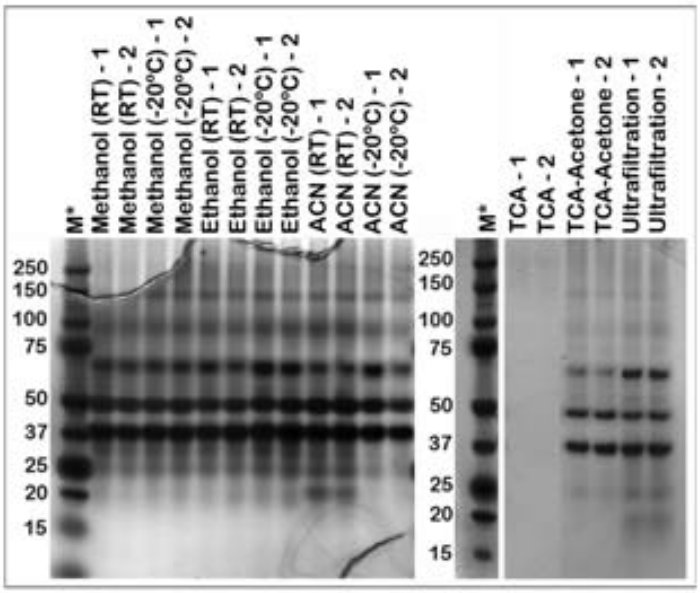



**Figure 2. Protein pattern distribution analysis by 1D-GE.**




* Protein marker; KDa bands

**Figure 3. Standard curves for exogenous and labelled peptides (ILP) from ADH1 and ENO1.**

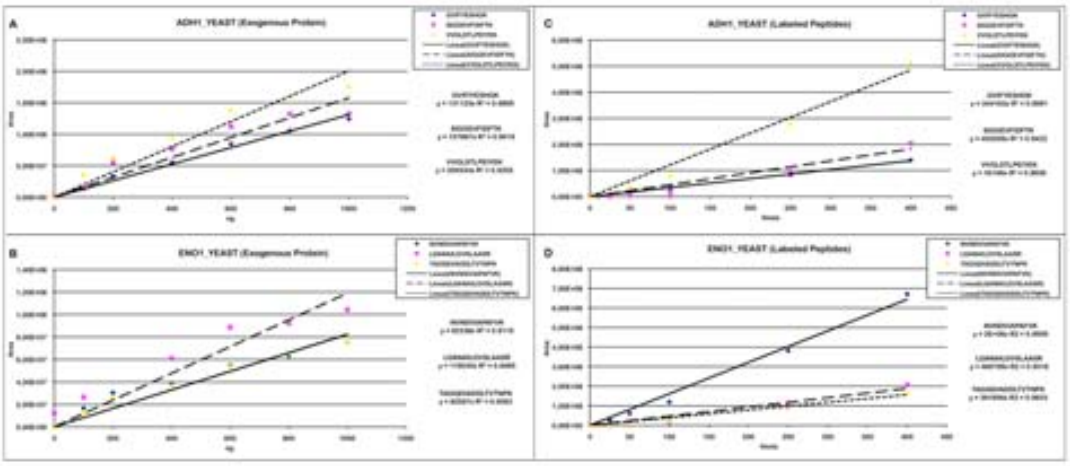


**Figure 4. Sample preparation absolute quantification by SRM-based assay.**

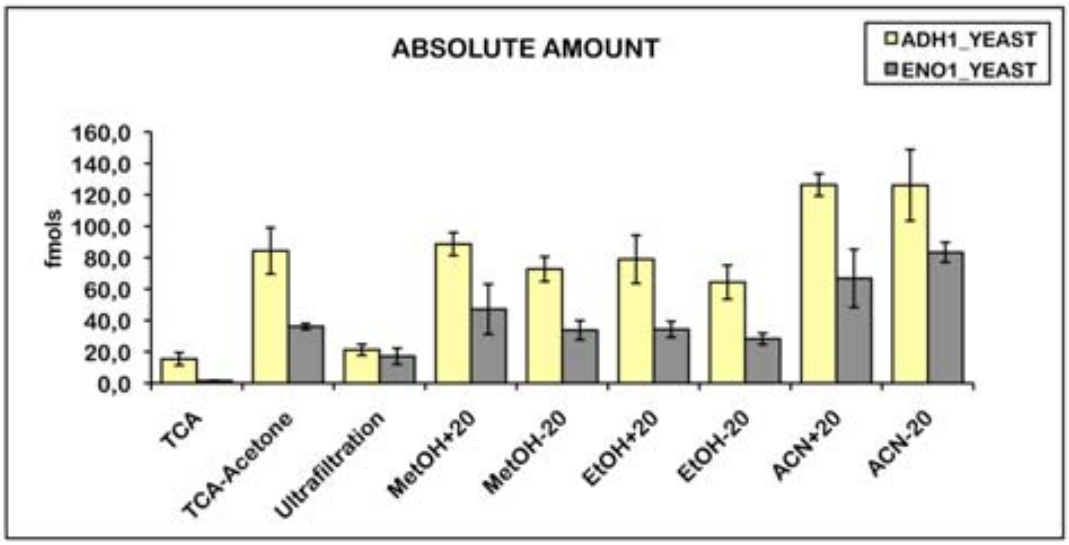


**Figure 5. Urine protein digestion standardization for SRM assay.**

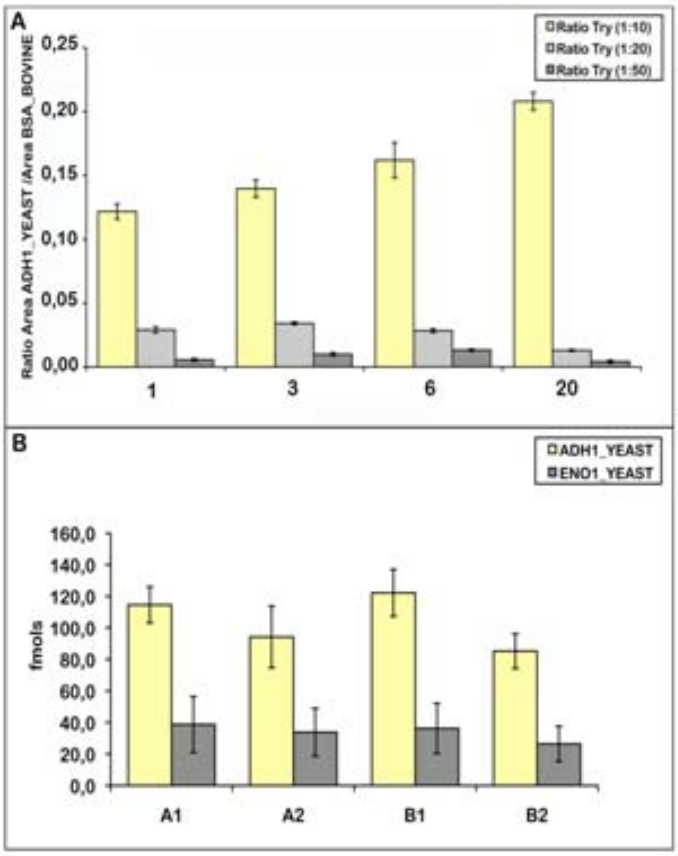
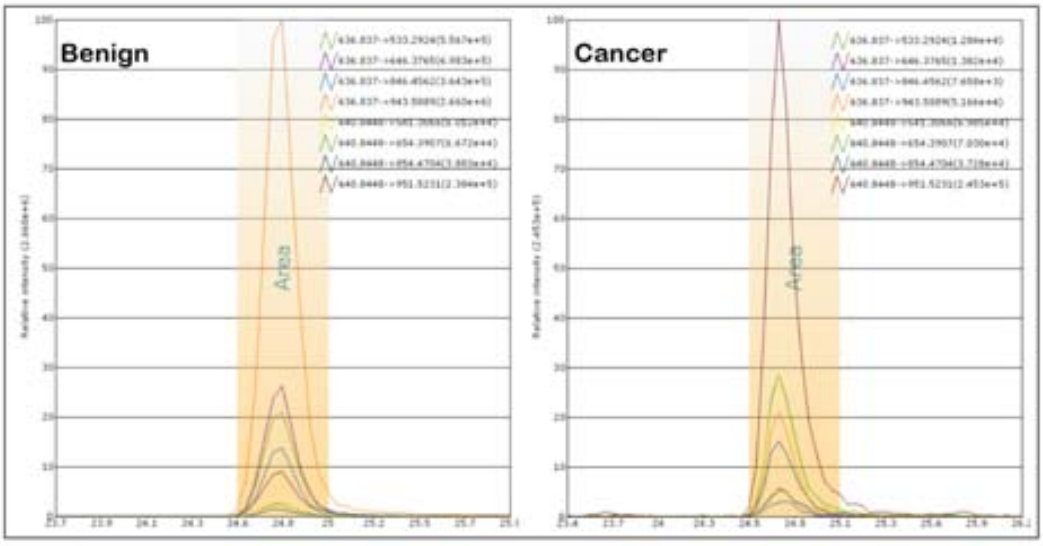



**Figure 6. Example of SRM transitions for both, heavy and light peptide for a specific protein in PCa and benign urine sample.**



Printed in Barcelona, June 2011

Cover design by Borja Exposito