

PONTIFICIA UNIVERSIDAD CATÓLICA DEL PERÚ

FACULTAD DE CIENCIAS E INGENIERÍA



**PONTIFICIA
UNIVERSIDAD
CATÓLICA
DEL PERÚ**

**IDENTIFICACION Y VALIDACION DE BIOMARCADORES
ÓMICOS INVOLUCRADOS EN EL PRONÓSTICO DE
RECURRENCIA BIOQUÍMICA DE CÁNCER DE PRÓSTATA**

**Tesis para optar por el Título de Licenciado en Química,
que presenta el bachiller:**

MIGUEL ARTURO ESPINOZA PORTOCARRERO

ASESOR(A): Dra. Fanny Lys Casado Peña

Lima, setiembre de 2018



Para mi madre, mis hermanos, mi abuelo Juan y mis tías.

AGRADECIMIENTOS

A mi pequeña y bella familia, por brindarme su apoyo, cariño y comprensión durante toda la carrera universitaria.

A mi madre, por siempre haberme dado palabras de aliento para nunca rendirme y haber movido cielo y tierra por apoyarme en mi desarrollo profesional.

A mi asesora Fanny Casado, por su guía y comprensión durante toda esta etapa universitaria.

A AUNA-Oncosalud por haberme brindado la oportunidad de trabajar con miembros de la Unidad de Investigación Básica y Traslacional, en particular a los biólogos Joseph Pinto y Jhajaira Araujo por su acompañamiento, paciencia y apoyo en esta etapa.

A los profesores Alfredo Ibañez y Giancarlo Sal y Rosas, por su apoyo y supervisión en la realización de la tesis.

A todos los docentes de la sección de Química por haberme instruido en los conocimientos necesarios y compartido sus experiencias para mi desarrollo profesional y personal.

A mis amigos, compañeros y seres queridos que me acompañaron en esta etapa tan exigente pero gratificante. Les agradezco los ánimos y consejos que me brindaron para cumplir mis metas académicas y personales. Quiero hacer una mención especial a mis amigos con quienes empecé la carrera de química en Facultad, ya que con ellos emprendí esta aventura y siempre estuvieron presentes en los momentos decisivos de mi desarrollo personal.

A la Pontificia Universidad Católica del Perú, por haberme brindado una formación interdisciplinaria y de excelencia, por los todas las experiencias y recuerdos.

RESUMEN

El cáncer de próstata es la segunda neoplasia más común en la población masculina del mundo. El comportamiento clínico del cáncer de próstata localizado es muy variable; mientras que algunos casos tienen un tipo agresivo de cáncer que resulta en metástasis y muerte del paciente, otros tendrán un tipo indolente que se puede curar con terapias o monitorear cuidadosamente. Existen múltiples sistemas de estratificación de riesgo de mortalidad que usan parámetros clínicos, como los niveles de PSA y la Puntuación de Gleason. No obstante, estos criterios no pueden predecir adecuadamente el riesgo de recurrencia bioquímica.

Los pacientes con cáncer de próstata no pueden ser dicotomizados con precisión en grupos de recurrencia bioquímica de bajo o alto riesgo mediante el uso de parámetros clínicos. Por este motivo se integró información genómica y clínica con el objetivo de identificar potenciales biomarcadores predictivos y generar firmas pronóstico que permitan una dicotomía más acertada de los pacientes. Se empleó una metodología de análisis estadístico predictivo de la recurrencia bioquímica utilizando genes relacionados con la regeneración de células madre de relevancia para la recurrencia bioquímica de cánceres, como la vía de señalización Wnt y la pluripotencia de células madre; y la contribución del parámetro clínico de la Puntuación de Gleason, de manera que se generó un firma pronóstico integrada.

La firma integrada fue validada en cohortes independientes de pacientes disponibles en repositorios internacionales para ser dicotomizados en grupos de riesgo que puedan asociarse con un pronóstico bueno o malo. De esta manera, se tendrá un mejor pronóstico de los pacientes y la asignación adecuada de terapias para su tratamiento. Queda claro que el desarrollo y validación de nuevos biomarcadores para evaluar el pronóstico de recurrencia en cáncer de próstata contribuiría a mejorar la salud en la mayoría de los países independientemente de sus características sociales, económicas, culturales y epidemiológicas.

TEMA DE TESIS PARA OPTAR EL TÍTULO DE LICENCIADO EN QUIMICA

TÍTULO : Identificación y validación de biomarcadores ómicos involucrados en el pronóstico de recurrencia bioquímica de cáncer de próstata

ÁREA : Bioquímica

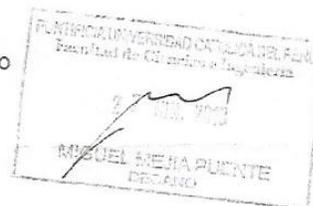
ASESOR : Dr. Fanny Lys Casado Peña

ALUMNO : Miguel Arturo Espinoza Portocarrero

CÓDIGO : 20112236

FECHA : 25 de junio del 2018

NUMERO : # 220



DESCRIPCIÓN Y OBJETIVOS:

El cáncer de próstata es la segunda neoplasia más común en la población masculina del mundo. El comportamiento clínico del cáncer de próstata localizado es muy variable; mientras que algunos casos tienen un tipo agresivo de cáncer que resulta en metástasis y muerte del paciente, otros tendrán un tipo indolente que se puede curar con terapias o monitorear cuidadosamente. Existen múltiples sistemas de estratificación de riesgo de mortalidad que usan parámetros clínicos, como los niveles de la proteína PSA y el índice de Gleason. Sin embargo, estos criterios no pueden predecir adecuadamente el riesgo de recurrencia bioquímica. Los avances en genómica y técnicas moleculares permiten generar perfiles moleculares para subtipar varios tipos de cáncer y seleccionar intervenciones terapéuticas más específicas, pero no están disponibles en la práctica clínica cotidiana. Dado que los pacientes con cáncer de próstata no pueden ser dicotomizados con precisión en grupos de recurrencia bioquímica de bajo o alto riesgo simplemente mediante el uso de características clínicas, se propone la integración de información genómica y clínica con el objetivo de generar firmas pronósticas que permitan la selección de terapias idóneas. El presente estudio, empleará técnicas de análisis estadístico predictivo de la recurrencia bioquímica utilizando genes relacionados con la regeneración de células madre de relevancia para la recurrencia bioquímica de cánceres, como la vía de señalización Wnt y la pluripotencia de células madre; y la contribución del parámetro clínico del índice de Gleason para generar una firma pronóstica integrada. La firma integrada será validada en cohortes independientes de pacientes disponibles en repositorios internacionales para ser dicotomizados en grupos de riesgo que pueden asociarse con un pronóstico bueno o malo.

Máximo: 100 páginas

FIRMA DEL ASESOR: Fanny Lys Casado Peña

FIRMA DEL COORDINADOR: [Signature]

[Signature]

ÍNDICE GENERAL

1. INTRODUCCIÓN	1
2. MARCO TEÓRICO	2
2.1. Cáncer.....	2
2.2. Signos representativos del cáncer.....	3
2.2.1. Autosuficiencia en señales para el crecimiento.....	3
2.2.2. Insensibilidad a las señales inhibitoras del crecimiento.....	4
2.2.3. Evasión de la muerte celular programada.....	5
2.2.4. Potencial replicativo ilimitado.....	5
2.2.5. Angiogénesis sostenida.....	6
2.2.6. Invasión de tejidos y metástasis.....	7
2.3. Ciencias ómicas.....	8
2.4. Cáncer de Próstata.....	10
2.4.1. Próstata.....	10
2.4.2. Incidencia y mortalidad.....	12
2.4.3. Métodos de Monitoreo.....	13
2.5. Biomarcadores.....	15
2.6. Vías de señalización bioquímicas vinculadas al cáncer.....	16
2.6.1. Vía de señalización Wnt.....	17
2.6.2. Células Madre de Cáncer y su efecto en la recurrencia del cáncer.....	20
2.6.3. Células madre de cáncer.....	20
2.6.3.1. Recurrencia del cáncer.....	21
2.7. Imagenología.....	23
3. OBJETIVOS	27
3.1. Objetivo General.....	27

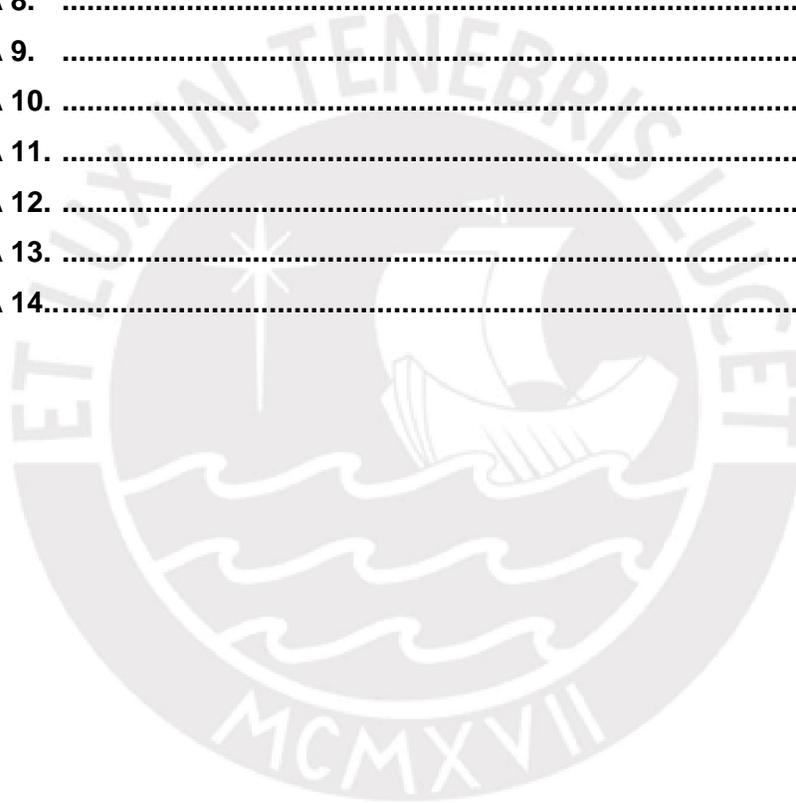
3.2. Objetivos Específicos	27
4. PARTE EXPERIMENTAL	28
4.1. Metodología	28
4.1.1. Obtención y validación de firma genética para pronóstico de recurrencia en cáncer de próstata	28
4.1.1.1. Generación de la lista maestra de genes	28
4.1.1.2. Selección de Cohortes	28
4.1.1.3. Selección de genes y pacientes	29
4.1.1.4. Normalización de los niveles de expresión de genes	29
4.1.1.5. Obtención de la firma genética para pronóstico de recurrencia en cáncer de próstata	30
4.1.1.6. Estimación de la sobrevida libre de recurrencia	31
4.1.1.7. Validación de la firma pronóstico.....	31
4.1.1.8. Integración de la Puntuación de Gleason para las cohortes estudiadas	32
4.2. Resultados	34
4.2.1. Obtención y validación de firma genética para pronóstico de recurrencia en cáncer de próstata	34
4.2.1.1. Selección de genes y pacientes	34
4.2.1.2. Obtención de la firma de pronóstico de genes	35
4.2.1.3. Estimación de la sobrevida libre de recurrencia en base a la firma pronóstico	36
4.2.1.4. Integración de la Puntuación de Gleason.....	39
5. DISCUSIÓN DE RESULTADOS	42
5.1. Firma pronóstico y estimación mediante el método de Kaplan Meier.....	42
5.2. Genes involucrados en las cascadas bioquímicas	45
5.2.1. TP53	45
5.2.2. FZD3 y FZD 5	46

5.2.3.	PLCB2	48
5.2.4	Vinculación de los genes identificados.....	49
6.	CONCLUSIONES	51
7.	BIBLIOGRAFÍA	53



ÍNDICE DE FIGURAS

FIGURA 1.	2
FIGURA 2.	10
FIGURA 3.	14
FIGURA 4.	17
FIGURA 5.	19
FIGURA 6.	24
FIGURA 7.	26
FIGURA 8.	34
FIGURA 9.	37
FIGURA 10.	38
FIGURA 11.	40
FIGURA 12.	41
FIGURA 13.	41
FIGURA 14.	49



ÍNDICE DE TABLAS

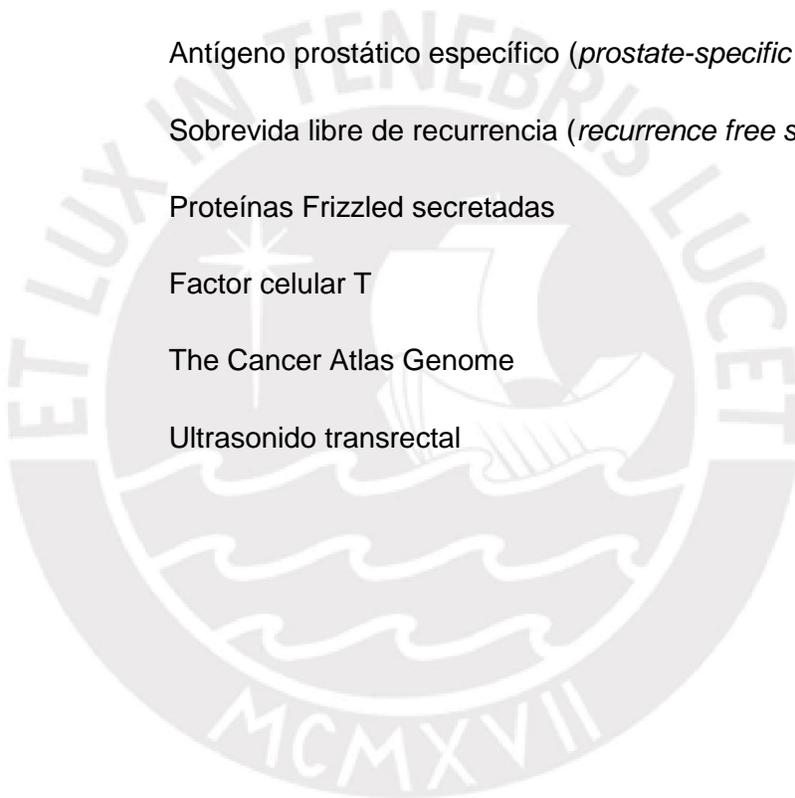
TABLA 1.	29
TABLA 2.	32
TABLA 3.	33
TABLA 4.	35
TABLA 5.	35
TABLA 6.	36
TABLA 7.	37
TABLA 8.	39
TABLA 9.	40



ABREVIATURAS

ADN	Ácido desoxirribonucleico
APC	Adenomatous polyposis coli
BCL9	Linfoma de células B9
CBP	Proteína de unión CREB
CK1α	Caseína quinasa 1 α
CSCs	Células madre de cáncer (<i>cancer stem cells</i>)
CV	Coefficiente de variación
DKKs	Dickkopfs
DVL	Dishevelled
ECM	Matriz extracelular (<i>extracellular matrix</i>)
ERD	Examinación rectal digital
FZD	Proteína receptora Frizzled
GEO	Gene Expression Omnibus
GS	Puntuación de Gleason (<i>gleason score</i>)
GSK3 β:	Glucógeno sintasa quinasa 3 β
H&E:	Hematoxilina-eosina (<i>hematoxylin and eosin</i>)
HR:	Razón de riesgo (<i>hazard ratio</i>)
IMS:	Imagenología por espectrometría de masas (<i>imaging mass spectrometry</i>)
KEGG:	Enciclopedia de Kioto de genes y genomas (<i>Kyoto encyclopedia of genes and genomes</i>)

LEF:	Factor potenciador linfoide
Log 2:	Logaritmo en base 2
MALDI:	Ionización/desorción laser asistida por una matriz (<i>matrix-assisted laser desorption/ionization</i>)
ARNm:	Ácido ribonucleico mensajero
MSKCC:	Centro de cáncer conmemorativo Sloan-Kettering (<i>memorial Sloan-Kettering cancer center</i>)
PSA:	Antígeno prostático específico (<i>prostate-specific antigen</i>)
RFS:	Sobrevida libre de recurrencia (<i>recurrence free survival</i>)
SFRPs:	Proteínas Frizzled secretadas
TCF:	Factor celular T
TCGA:	The Cancer Atlas Genome
USTR:	Ultrasonido transrectal



1. INTRODUCCIÓN

Desde la segunda mitad del siglo XX, se ha observado un incremento en el tiempo de vida de las personas y, por ende, las enfermedades cuyo riesgo incrementan con la edad. A su vez la edad avanzada incrementa el riesgo de padecer enfermedades degenerativas, tales como el cáncer de próstata, que si se diagnostican a tiempo y se aplica la terapia apropiada pueden ser tratadas con excelente prognosis. Sin embargo, en el Perú existe una alta tasa de mortalidad y de recurrencia luego de haber terminado los regímenes terapéuticos.

El cáncer de próstata es uno de los que presenta mayor incidencia a nivel mundial, pero su tasa de mortalidad varía de acuerdo con el país y de los cuales el Perú es el que presenta mayor tasa (cerca del 50%). El desarrollo y validación de nuevos biomarcadores para evaluar el pronóstico de recurrencia en cáncer de próstata contribuiría a mejorar la salud en la mayoría de los países, independientemente de sus características sociales, económicas, culturales y epidemiológicas. En el enfoque económico, una detección temprana de neoplasias y selección apropiada de la terapia que maximice los tiempos sin recurrencia después de tratamientos reducirá los costos asociados con su tratamiento en etapas más avanzadas de la enfermedad o en situaciones de recurrencia. Estos costos de tratamiento van más allá del dinero empleado en medicinas e infraestructura hospitalaria, ya que también involucran la pérdida de productividad tanto de los pacientes como familiares.

Por esta razón, surge la necesidad de desarrollar nuevas metodologías que permitan al médico y al paciente decidir sobre las mejores opciones terapéuticas a seguir de manera oportuna. La integración de información genómica e histopatológica proporcionaría una clasificación adecuada de los distintos subtipos de cáncer de próstata, lo que conducirá a una mejor precisión en el diagnóstico y pronóstico de los pacientes. Esta nueva herramienta, que involucra información molecular, busca es sistematizar la heterogeneidad intratumoral del cáncer con el propósito de encontrar y desarrollar una medicina más personalizada y así reducir las tasas de morbilidad y mortalidad.

2. MARCO TEÓRICO

2.1. Cáncer

El cáncer es una enfermedad que involucra cambios dinámicos en el genoma humano. Estos cambios se conocen comúnmente como daño genético o mutaciones, y se refieren a alteraciones en el ADN de las células. En otras palabras, el cáncer es una enfermedad genética adquirida en la que un único grupo de células y su progenie acumulan cambios hereditarios que dan como resultado el fenotipo celular del cáncer.¹

Los fundamentos de este conocimiento fueron establecidos principalmente por el descubrimiento de mutaciones que producen oncogenes con una ganancia de función dominante y genes supresores de tumores con una pérdida de función recesiva.² Los oncogenes promueven el crecimiento celular y los genes supresores de tumores lo inhiben. Una mutación de un oncogén conduce al cáncer porque las instrucciones que controlan la división celular han sido dañadas. Por otro lado, los genes supresores de tumores mutados pueden causar una inestabilidad genómica debido a que el sistema de reparación del ADN está comprometido.³ Cada una de estas mutaciones genéticas son eventos exclusivos que implican cómo una célula normal puede convertirse en tumorigénica.

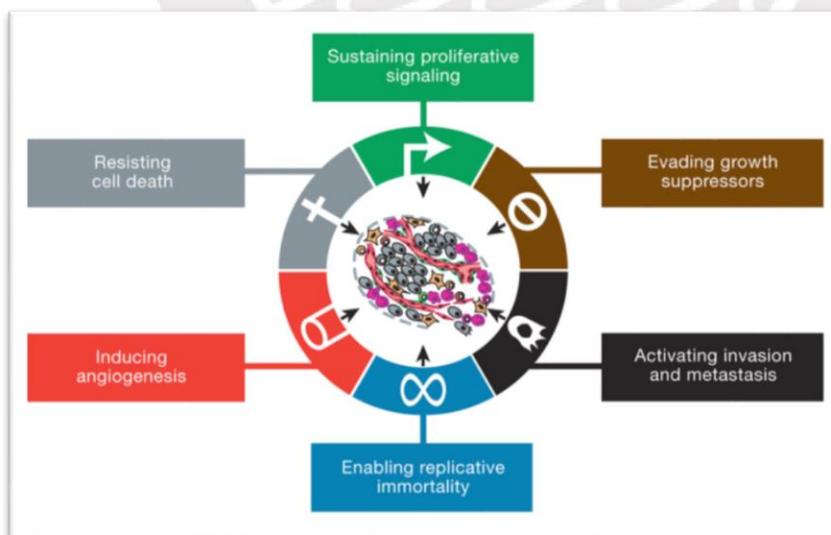


Figura 1. Los seis signos representativos para el desarrollo del cáncer sustentan el origen multifactorial de la enfermedad.²

Sin embargo, a partir del 2000, se ha venido acumulando evidencia que indica que la tumorigénesis en los seres humanos es un proceso de varios pasos y cada paso refleja alteraciones genéticas que impulsan el cambio de las células humanas normales hacia neoplasia maligna.² Se propusieron seis alteraciones esenciales en la fisiología de las células que promueven colectivamente el crecimiento maligno: Autosuficiencia en las señales de crecimiento, insensibilidad a las señales inhibitoras del crecimiento, evasión de la muerte celular programada (apoptosis), potencial replicativo ilimitado, angiogénesis sostenida y, invasión activada y metástasis (**Figura 1**). Cada una de estas capacidades adquiridas durante el desarrollo del tumor, hoy en día conocido como signos distintivos del cáncer, representa una manera exitosa de romper el mecanismo de defensa contra el cáncer de células y tejidos, y es la principal razón del enorme catálogo de genotipos de cáncer.⁴

2.2. Signos representativos del cáncer

A continuación, se presentarán cada de uno de los seis signos distintivos del cáncer para entender, de manera general, como surge el cáncer en los seres humanos.

2.2.1. Autosuficiencia en señales para el crecimiento

Los tejidos normales controlan cuidadosamente la producción y liberación de señales promotoras de crecimiento que instruyen en la entrada y progresión de la división y crecimiento celular. Estas señales mitogénicas de crecimiento (mitógenos) hacen que las células pasen a un estado proliferativo activo.² Su transmisión hacia el interior de la célula se ejecuta mediante el acoplamiento de moléculas de señalización, conocidas como factores de crecimiento, a un receptor en la membrana celular; esta etapa desencadena una cascada bioquímica dentro de la célula que resulta en la división celular. De esta manera, ningún tipo de célula normal puede proliferar en la ausencia de estas señales estimuladoras.

El propósito de controlar la división celular es mantener la arquitectura y funcionamiento del tejido normal. En el caso de las células cancerígenas, ellas mismas son las que generan sus propias señales de crecimiento, por lo que reducen su dependencia de estimulación de su microambiente tisular normal.⁵ Como consecuencia de esto, el

mecanismo homeostático, que normalmente regula el adecuado comportamiento de las distintas células en un tejido, se ve alterado.

Las células cancerígenas adquieren autosuficiencia al perturbar cada una de las vías que restringen la respuesta proliferativa en las células normales. Una clase de mutaciones actúa originando un cortocircuito en el requisito obligatorio de las células somáticas de recibir señales mitogénicas externas. Tales mutaciones resultan en la producción autocrina, como en caso de los embriones, de un mitógeno que normalmente es limitante. Una segunda clase de mutaciones desreguladoras del crecimiento son aquellas que activan permanentemente las vías de señalización que responden a estas señales. Esto significa que los receptores no necesitarían de un ligando para activar las señales de crecimiento hacia el interior de las células, lo que resulta en un crecimiento descontrolado. Finalmente, la última vía está vinculada con la desregulación del mecanismo de retroalimentación negativa, el cual se encarga de regular el flujo de señales que atraviesan el circuito intracelular. Este mecanismo cuenta con factores inhibidores del crecimiento, tales como el factor- β y los interferones, lo cuales ejercen un potente efecto antiproliferativo. Algún defecto en este mecanismo aumentaría la señalización proliferativa.⁵

2.2.2. Insensibilidad a las señales inhibitoras del crecimiento

El ciclo celular es un proceso por el que las células deben atravesar para proliferar. En un tejido sano, las células que lo conforman tienden a realizar tres actividades: dividirse, reposar o morir. De esta manera, para que el tejido mantenga su tamaño debe existir una tasa entre las células que mueren y las que se dividen.

Para la formación de los tumores, las células cancerígenas deben evadir procesos que regulan negativamente como la proliferación celular, los cuales dependen de la acción de genes supresores de tumor.⁵

Estos genes codifican proteínas supresoras de tumor, las cuales cumplen la función de limitar el crecimiento y proliferación celular. Las más destacadas son la proteína TP53 y la RB (asociada al Retinoblastoma); éstas operan como nodos centrales de control y

gobiernan la decisión sobre si las células pueden entrar al ciclo celular o, alternativamente, activar mecanismos apoptóticos⁵. Por su parte, las células cancerígenas sufren alteraciones en estos genes y codifican proteínas ineficientes que no van a poder diferenciar entre una célula sana o una cancerígena. De esta manera, las células cancerígenas, que presentan mutaciones, continúan con su proliferación y crecimiento lo cual resulta en la alteración de la estructura del tejido.

2.2.3. Evasión de la muerte celular programada

La apoptosis, una forma de muerte celular programada, es un mecanismo particular por el cual las células se programan para morir por el bienestar del organismo. El propósito es asegurarse de que las células que presentan el ADN dañado o alguna anomalía sean destruidas y no puedan multiplicarse. Esto permitirá el adecuado desarrollo del organismo mediante el mantenimiento de los tejidos de su cuerpo.

Para que las células sean sometidas a la apoptosis, se deben detectar algunas señales de estrés fisiológico, como elevadas señales de oncogenes o ADN dañado.⁵ Uno de los supresores de tumor involucrados es el TP53, el cual activa las vías de señalización, en presencia de estas señales, que dirigen la destrucción de la célula. En el caso del cáncer, estas células ignoran o evaden este tipo de señales, para así transmitir las mutaciones a las siguientes generaciones de células. Nuevamente, para poder realizar esto, las células deben tener una alteración en el gen que codifica a esta proteína; esto evita que la célula desencadene la apoptosis. De esta manera, las células cancerígenas no son destruidas y permite que las mutaciones permanezcan en el genoma, lo cual favorecerá la aparición de otras nuevas.

2.2.4. Potencial replicativo ilimitado

La mayoría de las células son capaces de atravesar por un número limitado de ciclos sucesivos de división. Una vez alcanzado el límite de divisiones, las células entran en un estado antiproliferativo (senectud) o, simplemente, se les induce a la apoptosis (crisis).⁵ Esta barrera está principalmente vinculada con los telómeros, que son repeticiones de hexanucleótidos que protegen la parte final de los cromosomas, ya que durante el proceso de replicación parte del ADN se pierde. El propósito es el de evitar que el ADN,

que codifica proteínas para los distintos procesos bioquímicas, se vea comprometido y, así, no resulte en la producción de proteínas deficientes.

La telomerasa es una enzima especializada que añade segmentos repetitivos de telómeros a los extremos del ADN. Su función es importante ya que mantiene la simetría de los cromosomas, pero en el caso del cáncer, se activa de manera inapropiada o cuando no se les requiere. En otras palabras, existe una regulación positiva de la expresión de esta enzima, lo cual resulta en el aumento del tamaño de los telómeros; este nivel de expresión está presente en muchas de las células “inmortales”, como es el caso del cáncer.⁵

2.2.5. Angiogénesis sostenida

Los tejidos normales del cuerpo humano tienen vasos sanguíneos que los atraviesan para distribuir el oxígeno del cuerpo, así como los distintos nutrientes que necesitan. Del mismo modo, las células cancerígenas también requieren de estos componentes para su crecimiento. En ambos casos, las células deben encontrarse a una distancia adecuada para que los nutrientes y el oxígeno se suministren lo suficiente como para su supervivencia. Conforme aumenta el tamaño de los tejidos, estos requerirán un suministro de sangre cada vez más cercano. El proceso por el cual se estimula el crecimiento de los vasos sanguíneos es conocido como angiogénesis.⁵

Durante el desarrollo de un embrión, la cicatrización de una herida o el ciclo menstrual de la mujer, la angiogénesis es un proceso normal dado que se manifiesta durante momentos específicos y en beneficio del organismo. Algunos de los genes que participan en este proceso, ya sea como inductores o inhibidores, son el factor de crecimiento endotelial vascular A (VEGF-A, por sus siglas en inglés) y la trombospondina-1 (TSP-1), respectivamente.⁵ La alteración de este interruptor angiogénico en las células cancerígenas incrementa la producción de factores inductores que promueven el crecimiento de vasos sanguíneos, mientras disminuye la cantidad de factores que inhiben este proceso. El resultado es que los tumores, conforme van desarrollándose, presentan una vasculatura adecuada que les permite el suministro continuo de nutrientes para su crecimiento.

2.2.6. Invasión de tejidos y metástasis

Un tumor en crecimiento se encuentra localizado en un órgano específico, pero su crecimiento descontrolado genera demanda de espacio y nutrientes por lo que existe una selección de células con la capacidad de invadir tejidos circundantes y/o más distantes. La diseminación de las células cancerígenas hacia otros órganos dentro del cuerpo se conoce como metástasis y se asocia a cómo va progresando la neoplasia hacia un grado de malignidad mayor.⁵ Este signo del cáncer se asocia con el desarrollo de alteraciones en la forma de las células, así como la unión a otras células y a la matriz extracelular (ECM, por sus siglas en inglés).

La ECM está compuesta por células en el tejido y éstas se encuentran unidas unas a otras para que el tejido mantenga una estructura adecuada. La E-cadherina es una molécula importante para el mantenimiento de la estructura del tejido ya que ayuda a ensamblar las células epiteliales mediante uniones adherentes. En el caso del cáncer, existe una disminución en la regulación de esta proteína o una inactivación mutacional que resulta en el desprendimiento de las células cancerígenas de la ECM, las cuales van a esparcirse por el torrente sanguíneo o el sistema linfático, y se ubicarán, posteriormente, en otros órganos o, incluso, en los huesos.⁵ La unión entre células de un tejido es una característica crítica para evitar la metástasis, debido a que mantiene al tumor localizado para que no llegue a afectar otros órganos del cuerpo.

En resumen, los signos representativos del cáncer proporcionan un marco teórico útil para la generación de hipótesis con respecto al diagnóstico, pronóstico y propuesta de blancos de tratamientos del cáncer.

2.3. Ciencias ómicas

Las ciencias ómicas involucran toda información adquirida por las distintas disciplinas como la genómica, la transcriptómica, la proteómica y la metabolómica para el entendimiento de las interacciones complejas dentro de sistemas biológicos⁶:

- La genómica estudia la estructura, función y expresión de todos los genes en un organismo.
- La transcriptómica estudia el ARN mensajero (ARNm) dentro de una célula u organismo.
- La proteómica estudia conjuntos de proteínas de manera sistemática ya sea en base a su estructura o función dentro de una célula, sistema u organismo.
- La metabolómica estudia los perfiles y patrones de metabolitos en un sistema (célula, tejido u organismo) bajo un conjunto de condiciones específicas.

Los diversos estudios sobre el genoma, transcriptoma, proteoma y metaboloma utilizan tecnológicas de alto rendimiento, como la secuenciación de ADN o la identificación de proteínas y metabolitos por espectrometría de masas, las cuales generan enormes cantidades de información bioquímica. Por este motivo, se requiere de herramientas computacionales y métodos estadísticos para el manejo y análisis de este conjunto de datos biológicos.⁷ Es en este escenario donde la bioinformática representa un rol esencial.

La bioinformática es un campo interdisciplinario que combina las ciencias de la computación y estadística con ciencias biológicas y biomédicas, como la bioquímica, genética, genómica, entre otras más.⁸ La principal función de esta disciplina es la de facilitar el manejo, análisis e interpretación del conjunto de datos biológicos, para transformarlo en información y conocimientos útiles. De manera más específica, la bioinformática abarca el desarrollo de bases de datos, para almacenar y recuperar datos; de algoritmos, para analizar y determinar las relaciones entre datos biológicos; y de herramientas estadísticas para identificar, interpretar y extraer datos.⁷

La integración de las ciencias ómicas será una herramienta necesaria y útil para que los múltiples datos ómicos se transformen en un modelo integral que describa con precisión los distintos procesos biológicos. Sin embargo, existen limitaciones frente a la información ómica disponible y lo que se busca, actualmente, es que el conjunto de datos se encuentre públicamente disponible para toda la comunidad científica. De esta manera, la mayor difusión y manejo de información ómica permitiría una aplicación más rápida del conocimiento integrativo de estas ciencias hacia la práctica clínica.⁹



2.4. Cáncer de Próstata

2.4.1. Próstata

La próstata es una glándula del tamaño de una nuez, ubicada debajo de la vejiga urinaria en los hombres y al frente del recto (**Figura 2**). Este órgano se divide en tres zonas: Periférica, central y transicional. La zona periférica compromete la mayor parte del volumen prostático (aproximadamente 65 %), mientras que la zona central es la segunda región más grande (25 % de su volumen). Por último, la zona de transición consiste de dos lóbulos que rodean al segmento uretral proximal lateralmente e interiormente.¹⁰

La principal función de este órgano es la de controlar el flujo urinario, pero a la vez produce algunas enzimas que juegan un papel en la función del fluido seminal. La testosterona y su metabolito más activo, dihidrotestosterona, sirven como nutriente para el tejido prostático y regulan la mitosis de las células.¹¹

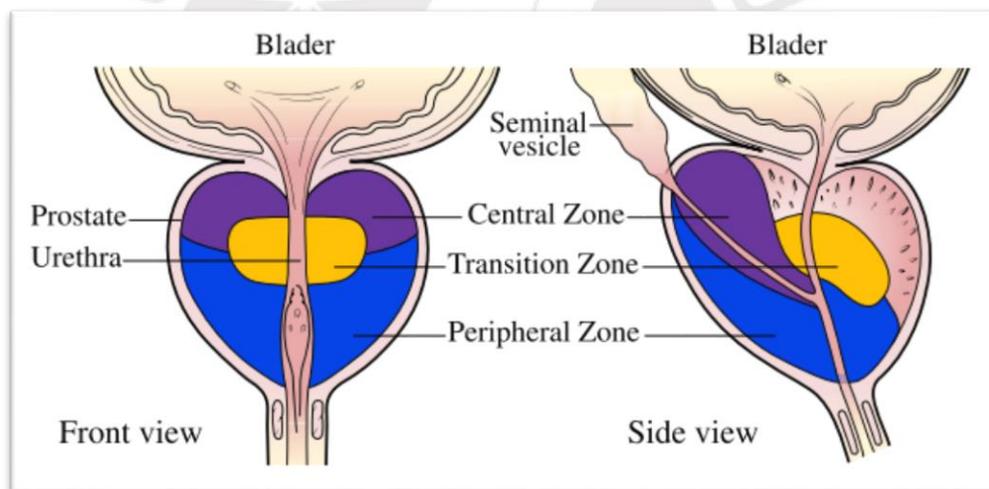


Figura 2: La anatomía de la próstata se divide en zonas periféricas, centrales y transicionales. Se presenta una vista frontal y lateral para ver las dimensiones de este órgano.¹⁰

El cáncer de próstata, también conocido como neoplasia maligna de próstata, generalmente se origina en el tejido glandular y se debe a que las células del tejido prostático empiezan a crecer sin control. A este tipo de cáncer se le conoce como

adenocarcinoma. Existen otros tipos menos comunes, como es el caso de sarcomas, carcinomas de células pequeñas, tumores neuroendocrinos y carcinomas de células transicionales. Sin embargo, el tipo de cáncer más común presente en la población masculina son los adenocarcinomas.

Algunos cánceres de próstata pueden crecer y esparcirse rápidamente, pero la gran mayoría lo hacen gradualmente. Los hombres jóvenes pueden presentar este tipo de cáncer, pero su desarrollo como enfermedad es muy lento, por lo cual recién llega a manifestar síntomas cuando alcanzan edades avanzadas. Estudios en autopsias muestran que muchos hombres adultos que presentan cáncer de próstata mueren por otras causas, mucho antes de que un síntoma se manifieste clínicamente.¹² Debido a esto, se expone al paciente a un sobrediagnóstico debido a que los métodos de detección también pueden detectar cánceres indolentes, y no resultar en síntomas o la muerte del paciente. Esto está vinculado con el sobretratamiento de los pacientes, ya que se les aplica tratamiento a quienes no se van a beneficiar o aquellos que presentan un tumor indolente, el cual no requeriría tratamiento. Por este motivo las técnicas de monitoreo para la detección temprana de cáncer de próstata deberían enfocarse en aquellas neoplasias que sí pueden progresar, y no en aquellas que son prácticamente inofensivas.

2.4.2. Incidencia y mortalidad

El cáncer de próstata es la segunda neoplasia más común en la población masculina a nivel mundial, únicamente superado por el cáncer de pulmón. Los valores de incidencia y mortalidad a nivel mundial para el año 2012, a raíz de esta enfermedad, fueron aproximadamente 1.1 millón y 307 000, respectivamente.¹³ Los continentes que presentan mayor incidencia son los de América y Europa. La tasa de mortalidad varía de acuerdo con el país, de los cuales el Perú presenta una de las mayores tasas cercana al 50%.

El tema de la incidencia de este cáncer está relacionado con múltiples factores genéticos y demográficos, como la edad, raza, historia familiar y susceptibilidad genética. Estos son algunos de los factores de riesgos con los cuales se relaciona la aparición de esta enfermedad. En su mayoría, la edad es un factor predominante, debido a que esta enfermedad aparece en hombres adultos entre los 50-70 años. Por otro lado, los afroamericanos presentan una mayor tasa de incidencia que los americanos de ascendencia europeas.¹⁴ De esta manera, se puede entender las variaciones con respecto a las tasas de incidencia tanto para cada continente, así como para los países dentro de cada continente.

2.4.3. Métodos de Monitoreo

Para el monitoreo y detección del cáncer de próstata, existen tres procedimientos clínicos: Examinación rectal digital (ERD), prueba del nivel de antígeno prostático específico (PSA) en sangre, y biopsias sistémicas guiadas por ultrasonido transrectal (USTR).

La ERD es el método de monitoreo del primer nivel de atención en salud. Consiste en el tacto directo de la glándula prostática mediante la inserción de un dedo del doctor o enfermera que realiza la prueba clínica. La persona que realiza la prueba siente la textura, forma o tamaño de la próstata para determinar si existe la presencia de un tumor. La próstata normal se siente firme con una superficie regular y lisa, y es simétrica.¹⁵ Cuando la próstata presenta una anormalidad (nódulo o surco) o una ligera asimetría, es indicador de la presencia de cáncer de próstata. Aunque la ERD es una buena herramienta para la predicción y diagnóstico de esta enfermedad, el número de casos detectados son muy bajos. Un problema con esta técnica es que presenta una baja reproducibilidad debido a variabilidad de los observadores, es decir, la persona que realiza el monitoreo. Asimismo, se ha reportado una baja sensibilidad, por lo que la ERD no debería ser el único método de monitoreo empleado.¹¹

El PSA es una proteasa inhibidora presente en la sangre y que es producida únicamente por la glándula prostática. El nivel de PSA en la sangre usualmente se relaciona con el volumen de la próstata.¹⁶ Esta proteína es un indicador específico del estado del órgano, pero no es estrictamente un indicador de tumor. En otras palabras, altos niveles de PSA (a partir de 4 ng/mL) pueden estar asociados con la infección, inflamación, ampliación o cáncer en la próstata. Si bien el nivel de PSA nos indica la presencia de cáncer de próstata, no es posible identificar si el tumor resultará maligno o benigno. Por este motivo, es que el uso de PSA origina un sobrediagnóstico que resulta en la obtención de biopsias innecesarias o negativas.¹¹

El nivel de PSA en conjunto con la ERD sirve para detectar el cáncer en estadios tempranos. No obstante, para confirmar la presencia de esta enfermedad se debe realizar una biopsia. Este procedimiento se realiza mediante el uso de USTR (**Figura 3**), el cual consiste en el uso de una aguja delgada que es dirigida por imágenes de

ultrasonido.¹⁷ El inconveniente es que se trata de una técnica invasiva y puede generar algunas complicaciones, como infecciones en el tracto urinario de los pacientes que se someten a este análisis. Sin embargo, es la técnica más utilizada para la obtención de biopsias que serán analizadas histológicamente cuando el tamizaje mediante PSA y ERD sugieren la posibilidad de una neoplasia.

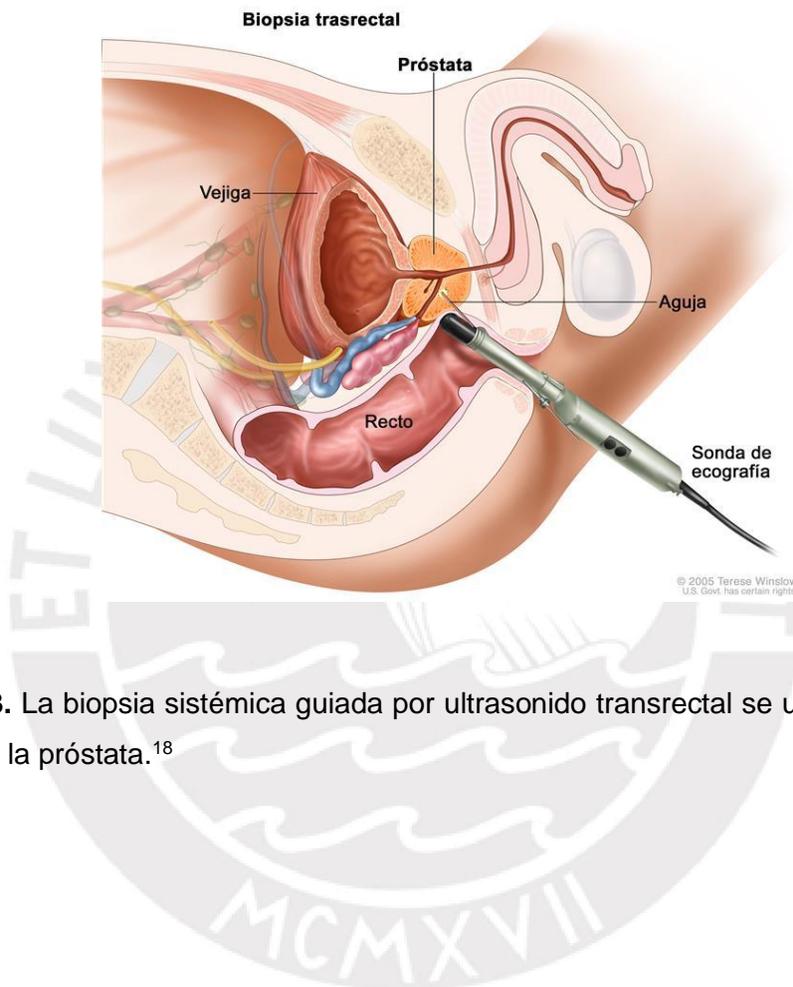


Figura 3. La biopsia sistémica guiada por ultrasonido transrectal se utiliza para extraer tejido de la próstata.¹⁸

2.5. Biomarcadores

Las cuatro unidades básicas de información de un organismo se encuentran en el genoma, transcriptoma, proteoma y metaboloma. En este sentido, los biomarcadores son una importante herramienta en la medicina individualizada ya que nos permite obtener una característica cuantificable y evaluada como indicador de un proceso biológico normal, procesos patológicos o respuesta farmacológica a una intervención terapéutica. Un biomarcador para el tema de cáncer puede tener aplicaciones relevantes como mejorar el diagnóstico y el estadio, indicar la prognosis de la enfermedad o seleccionar a los pacientes para diferentes opciones de tratamiento.¹¹

Para nuestro estudio, se utilizaron bases de datos de vías de señalización, como la del Kyoto Encyclopedia of Genes and Genomes (KEGG)¹⁹, para el estudio de las cascadas bioquímicas de interés y los genes involucrados. Lo que nos interesa de los genes son sus niveles de expresión, y para la obtención de esta información existen bases de datos como cBioportal web²⁰ o SurvExpress²¹, en la cual se pueden encontrar diversos conjuntos de datos sobre estudios asociados a una amplia variedad de cánceres. Lo resaltante de estas bases de datos es que también proporcionan información clínica de los pacientes, lo cual será muy útil para el análisis estadístico de los genes. Es común que, en estos estudios, se obtengan genes que sean falsos positivos, es decir, que no tengan relación directa con la enfermedad. Por este motivo, se deben realizar análisis estadísticos adecuados para su identificación y posterior uso como marcadores pronósticos.²²

El resultado de esto vendría a ser lo que se conoce como una firma genética, lo que equivaldría a un puntaje para diferenciar grupos de riesgos entre pacientes. En la medicina se utilizan mucho estos “scores” utilizando parámetros clínicos y patológicos, como los niveles de PSA o puntaje de Gleason, para encontrar rangos que engloban ciertas características. No obstante, estos marcadores de pronóstico, en su mayoría, no predicen adecuadamente el resultado para los pacientes.

2.6. Vías de señalización bioquímicas vinculadas al cáncer

En biología, las vías de señalización consisten en interacciones secuenciales en biomoléculas que generan un cierto producto o cambio en la función o en patrones de expresión celular. Asimismo, estas vías pueden activar y desactivar genes, o estimular a una célula para moverse, diferenciarse e incluso morir (**Figura 4**). Tanto desde el interior como desde el exterior del tejido, las células están recibiendo constantemente señales químicas generadas por estímulos tales como una lesión, infección, estrés o incluso la presencia o falta de nutrientes. Para reaccionar y ajustarse a estas señales, las células envían y reciben señales a través de vías bioquímicas. Las moléculas que forman las vías bioquímicas interactúan con señales, así como entre sí, para llevar a cabo sus tareas designadas.²³

El interés de los investigadores por el estudio de estas vías bioquímicas sobre las enfermedades humanas está vinculado con la identificación de genes, proteínas y otras moléculas asociados a estas vías que puedan servir como biomarcadores de diagnóstico, de pronóstico o como blancos para desarrollar mejores y más eficaces medicamentos.

El cáncer de próstata ha sido vinculado con la vía de señalización Wnt que también juega un papel importante en el desarrollo embrionario y en la diferenciación o pérdida de pluripotencia de células madre.

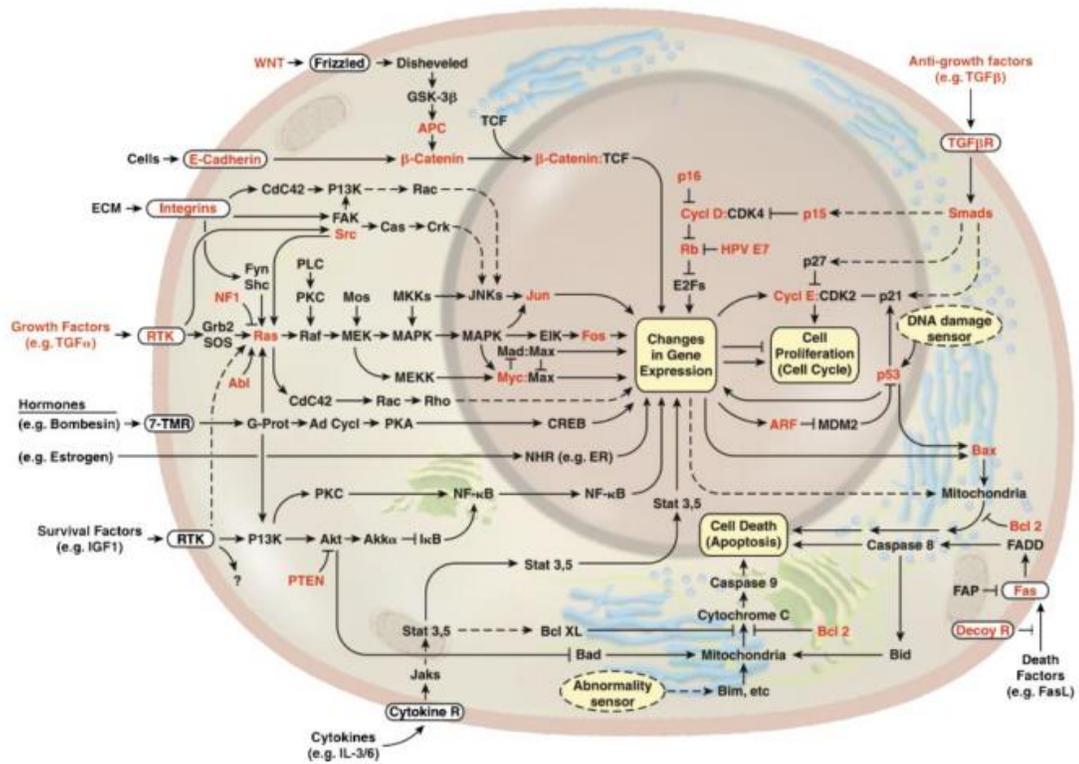


Figura 4. El Circuito integrado de vías de señalización dentro de una célula permite entender los estímulos que se presentan en el interior de la célula. Los genes de color rojo son aquellos que comúnmente presentan una función alterada en las células cancerígenas.²

2.6.1. Vía de señalización Wnt

La vía Wnt es un sistema que dirige una amplia variedad de respuestas celulares en el desarrollo de los tejidos, así como el mantenimiento de células madre.²⁴ Este sistema de señales se divide en las canónicas o vía dependiente de β -catenina, y las no canónicas, o vía independiente de β -catenina. Sin embargo, solo nos centraremos en las vías canónicas porque están vinculadas con las señales que regulan la transcripción de genes.²⁵ La molécula principal, como ya se mencionó, en esta vía es la β -catenina, la cual se encuentra en tres escenarios celulares: la membrana, el citoplasma y el núcleo. Es por este motivo que la característica más resaltante de esta vía es la acumulación y translocación de la adherencia de esta proteína en el núcleo.²⁶

El mecanismo por el cual se desarrolla la vía Wnt ha sido ampliamente estudiado por lo que se conocen los componentes involucrados, así como las interacciones entre ellos hasta alcanzar el núcleo (**Figura 5**).²⁷ Esta vía es activada por señales Wnt extracelulares, las cuales van a generar señales hacia el interior de la célula y desencadenan distintas cascadas bioquímicas. Para producir estas señales, este sistema consta de alrededor de 30 ligandos extracelulares, conocidos como proteínas Wnt, las cuales se unen al receptor transmembranoso Frizzled (FZD) y co-receptor LRP5-LRP6 en la membrana plasmática. Existen algunos inhibidores de esta interacción como las proteínas Frizzled secretadas (SFRPs, por sus siglas en inglés), la familia de proteínas Dickkopfs (DKKs) y el factor inhibidor Wnt 1 (WIF1). De esta manera, en ausencia de ligandos Wnt, la β -catenina citoplasmática es reclutada hacia un complejo de destrucción, en el cual interactúa con axinas y la proteína APC. Lo que continúa es la fosforilación N-terminal de la β -catenina por las quinasas CK1 α y GSK3 β , para su subsiguiente degradación por una ubiquitina. El bajo nivel de β -catenina asegura la represión transcripcional de los genes objetivo Wnt mediante el reclutamiento del co-represor Groucho hacia los factores de transcripción LEF-TCF.

Bajo la presencia de ligandos Wnt, LRP5 y LRP6 son fosforilados por CK1 α y GSK3 β , y las moléculas Dishevelled (DVL) son reclutadas hacia la membrana plasmática para interactuar con el receptor FZD y otras moléculas DVL. La interacción de la axina con los LRP5 y LRP6 fosforilados, y Dishevelled polimerizado lleva a la inactivación del complejo de destrucción y subsecuentemente hacia la estabilización de la β -catenina y su translocación hacia el núcleo. En el núcleo, esta proteína forma un complejo activo transcripcional con los factores de transcripción LEF y TCF al desplazar al co-represor Groucho e interactuar con otros co-activadores como las proteínas BCL9, Pygopus (Pygo) y CBP.

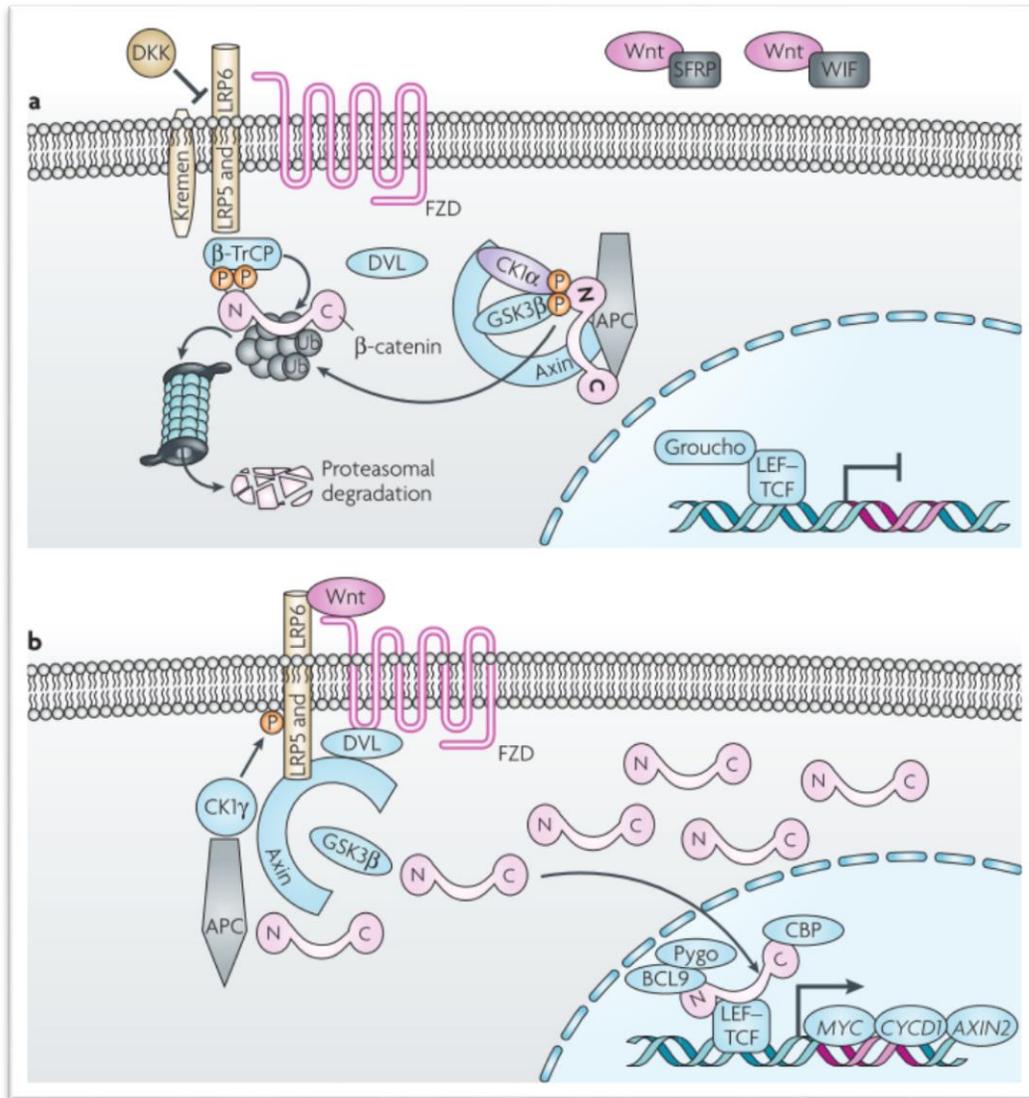


Figura 5. La vía de señalización canónica Wnt. a) En ausencia de ligando Wnt y b) en presencia del ligando Wnt.²⁷

La importancia de enfocarnos en esta vía de señalización para este estudio es que está relacionada con el desarrollo de enfermedades neoplásicas, como el cáncer de próstata. Diversos estudios relacionados a esta vía demuestran que existen una variedad de mutaciones que afectan la estabilidad de la proteína β-catenina, y que desencadenan el

desarrollo de una variedad de cánceres humanos.²⁷ Algunos ejemplos son la pérdida de función de los genes AXIN1 y AXIN 2, que evitan un acoplamiento adecuado del complejo de destrucción y, por ende, la β -catenina no puede ser marcada para su posterior degradación. Asimismo, la mutación del gen CTNNB1, el cual codifica a la β -catenina, presenta un aumento de su función, lo que evita la fosforilación de esta proteína y que sea acumulada en el citoplasma. De esta manera, se presenta a la vía Wnt como aquella que puede autorregular un flujo de retroalimentación negativa, lo cual puede resultar, como ya se mencionó, en la progresión de tumores, al activar genes asociados a la proliferación celular, como MYC y ciclina D1.

Asimismo, la vía Wnt está vinculada con el desarrollo y mantenimiento de las células madre de distintos tejidos y órganos. Por este motivo, las señales inadecuadas de esta vía pueden desencadenar la desregulación de células madre en el órgano de la próstata.

2.6.2. Células madre de cáncer y su efecto en la recurrencia del cáncer

La idea de que un tumor es una masa de células malignas homogéneas es equivocada para el cáncer de próstata sin importar su origen celular. Estos tumores son ecosistemas complejos que contienen células tumorales, al igual que otros tipos de células que pueden influenciar en la función de un tumor como conjunto. Avances en secuenciación genómica han demostrado que el cáncer dentro de un solo paciente es una mezcla heterogénea de subconjuntos (subclones) genéticamente distintos que aparecen debido a presiones evolutivas. En este contexto, la hipótesis de lo que se conoce como células madre de cáncer propone una posible explicación de por qué los pacientes que reciben tratamientos clínicos, cuyo objetivo es eliminar la mayor cantidad de células con alto potencial replicativo, pueden recurrir en esta enfermedad.²⁸

2.6.3. Células madre de cáncer

La característica que mejor define a las células madre consiste en la auto-renovación celular. Éste es un proceso biológico clave en el cual, mediante la división celular, una célula madre produce una célula hija (división asimétrica) o dos células hijas (división simétrica) que retienen la capacidad de auto-renovarse, para así asegurar el mantenimiento de la población de células madres o su expansión para crecimientos

clonales de largo plazo.²⁸ Las células madre de cáncer (CSC, por sus siglas en inglés) al igual que las células madre sanas presentan esta capacidad, pero es desregulada para el caso de las CSC. De esta manera, tenemos una población de CSC que se diferencian de las células cancerígenas que presentan una alta capacidad de proliferación. Las terapias clínicas, como quimioterapia o radioterapia, pueden disminuir el tamaño del tumor al eliminar las células cancerígenas, pero las CSCs permanecen intactas lo cual permite que la reaparición de la neoplasia. Posteriormente, las nuevas células cancerígenas van a presentar nuevas mutaciones que posiblemente las haga resistentes a los distintas terapias clínicas.²⁸

2.6.3.1. Recurrencia del cáncer

La existencia del modelo de CSC ha sido muy útil en el estudio de recurrencia en cáncer de leucemia y cáncer de mamá, y se encuentran reportadas en la literatura científica descrita por Kreso y sus colaboradores.²⁸ Por su parte, existe evidencia que sugiere la relevancia del estudio de recurrencia en cáncer de próstata desde una perspectiva de células madre y pluripotencia. Estos estudios vincularon la vía Wnt con la desregulación de las células madre de próstata adultas, las cuales podrían adquirir la capacidad de autorrenovarse durante la oncogénesis.²⁹

Otros estudios recientes han mostrado interés en la obtención de biomarcadores mediante técnicas de alto rendimiento como la secuenciación de ADN³⁰ y detección de proteínas involucradas mediante espectrometría de masas.³¹ Los estudios son realizados con el propósito de relacionar esta información molecular con la información clínica, de la cual ya se dispone. Se utiliza clínicamente el puntaje de Gleason y los valores de PSA como puntos de corte para dividir a los pacientes en grupos de riesgos, pero los resultados no siempre son los pronosticados. Por este motivo, la intención es obtener información genómica con el propósito de generar modelos de predicción más eficientes. De esta manera, se obtendría lo que comúnmente se conoce como firma genética³², con la cual se generaría mejores puntos de cortes para la prognosis del cáncer de próstata. En otras palabras, este análisis tendría un mayor nivel de pronóstico y los biomarcadores identificados podrían ser relacionados con esta enfermedad.

Estos estudios son importantes ya que cuando a un paciente se le considera en un grupo de riesgo alto, usualmente se les interviene quirúrgicamente para removerles la próstata mediante una prostactetomía radical. Sin embargo, según el estudio realizado por Komisarof³² y colaboradores, el 33% de los pacientes que participaron en el estudio recayeron en esta enfermedad. Por lo tanto, es importante la identificación de estos biomarcadores para un mejor diagnóstico y pronóstico, de manera que se realicen los tratamientos adecuados según cada paciente y se evite además destinar recursos a pacientes que no necesitan las intervenciones.



2.7. Imagenología

Varias condiciones médicas, no solo para casos de cáncer, puede ser diagnosticadas al remover una muestra de tejido para realizar alguna prueba de laboratorio. Esta extracción de tejidos, proveniente de cualquier órgano del cuerpo, se le conoce como biopsia y debe ser examinada por un patólogo bajo un microscopio.¹⁸ La información que proporciona una biopsia es sobre la estructura tisular que se puede obtener de una evaluación microscópica guiada por técnicas de tinción histológica. Entre ellas, La tinción por **H&E** es el método estándar que se utiliza comúnmente para estudios en base a tejidos.³³ La técnica permite tanto mejorar el contraste como discernir entre los núcleos y el citoplasma en los tejidos al teñirlos de diferentes colores. La hematoxilina tiñe el núcleo celular de color azul oscuro, mientras que la eosina tiñe el citoplasma y la mayoría de fibras tisulares de color rojizo (tonalidades de rosado, anaranjado y rojo).³⁴ Este proceso de dos colores ayuda a representar mejor la morfología de los tejidos y células, y facilita el diagnóstico a través del microscopio.

El método comúnmente utilizado para clasificar histológicamente los tejidos prostáticos es el sistema de clasificación de Gleason. El tejido es clasificado en cinco grados (**Figura 6**), numerados desde el 1 hasta el 5. El grado aumenta con el incremento del nivel de malignidad, que está vinculado con la agresividad del cáncer. Finalmente, la heterogeneidad del cáncer de próstata se describe mediante el puntaje de Gleason, el cual es la suma entre el primer y segundo grado más predominantes en la sección de tejido. De esta manera, se logra cuantificar el grado de agresividad del cáncer al utilizar la tinción por H&E con lo cual se determina el grado de diferenciación entre células normales y células malignas. Sin embargo, esta práctica histológica en el diagnóstico del cáncer y la clasificación de Gleason es propensa a la subjetividad y la reproducibilidad limitada, debido a su gran confianza en la interpretación humana.³⁵



Figura 6. Sistema de clasificación de Gleason.³⁵ 1) Pequeñas y uniformes glándulas, 2) más estroma entre las glándulas, 3) márgenes claramente infiltrativos, 4) masas irregulares de glándulas neoplásicas y 5) formación ocasional de glándulas.

La imagenología por espectrometría de masas (IMS, por sus siglas en inglés) es una tecnología que genera mapas de espectros de masas de compuestos químicos para conocer la distribución de moléculas en los tejidos. Presenta ventajas únicas de alta sensibilidad, amplio rango de moléculas, especificidad molecular y la flexibilidad de manejar una variedad grande de analitos en una misma plataforma. Lo interesante de esta técnica es que nos permite la traducción de información molecular en morfología tisular. Esta combinación de las capacidades analíticas de la espectrometría de masas (MS, por sus siglas en inglés) con información microscópica ayudará al entendimiento de procesos moleculares que se estén desarrollando en un tipo específico de célula

dentro de un tejido.³⁶ Desde el punto de vista clínico, representa un salto tecnológico desde las biopsias que proporcionan morfología tisular de manera robusta, pero limitada información molecular.

Dentro de las diversas técnicas de ionización para MS que pueden utilizarse para el análisis directo de tejidos, la ionización por desorción láser asistida por una matriz (MALDI, por sus siglas en inglés) ha guiado el desarrollo de aplicaciones biológicas y clínicas para la IMS, y es una de las técnicas más comúnmente utilizadas. De esta manera, las áreas de investigación en medicina y patología muestran un gran interés por esta técnica ya que permite la identificación de una variedad de proteínas, lípidos y metabolitos que pueden ser utilizados como biomarcadores para el diagnóstico, así como el pronóstico y efectividad de la terapia a través de ensayos moleculares de las biopsias obtenidas de los pacientes.³⁷

Actualmente, los estudios patológicos basados en microscopía utilizan la tinción con anticuerpos específicos, en una técnica conocida como inmunohistoquímica, para la identificación de biomarcadores proteómicos en las biopsias de los pacientes. Sin embargo, una vez realizado este procedimiento, la sección del tejido ya no puede volver a analizarse y solo se puede identificar una molécula por ensayo.³⁶ Asimismo, es una técnica de alto costo ya que se requieren anticuerpos específicos para cada antígeno a estudiar, por lo que no sería una técnica adecuada para implementarse en ensayos clínicos. La ventaja de MALDI IMS es que realiza análisis sin etiquetas y permite la obtención de mapas moleculares, es decir, la distribución topográfica de los analitos estudiados (**Figura 7**). Cabe mencionar que en cada medición por MALDI IMS se pueden obtener múltiples moléculas. Por último, al utilizar un método de ionización blando, la sección del tejido de la biopsia se recupera y se puede preservar para posteriores estudios ya que es una técnica no destructiva.

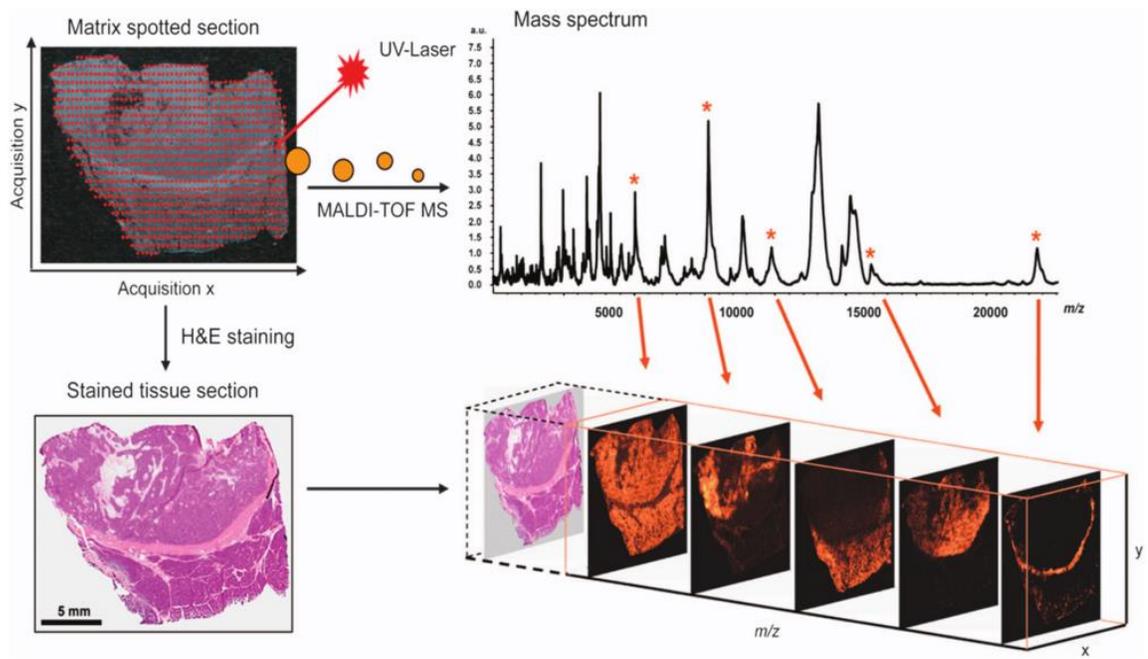


Figura 7. Principio del funcionamiento de MALDI IMS³⁶

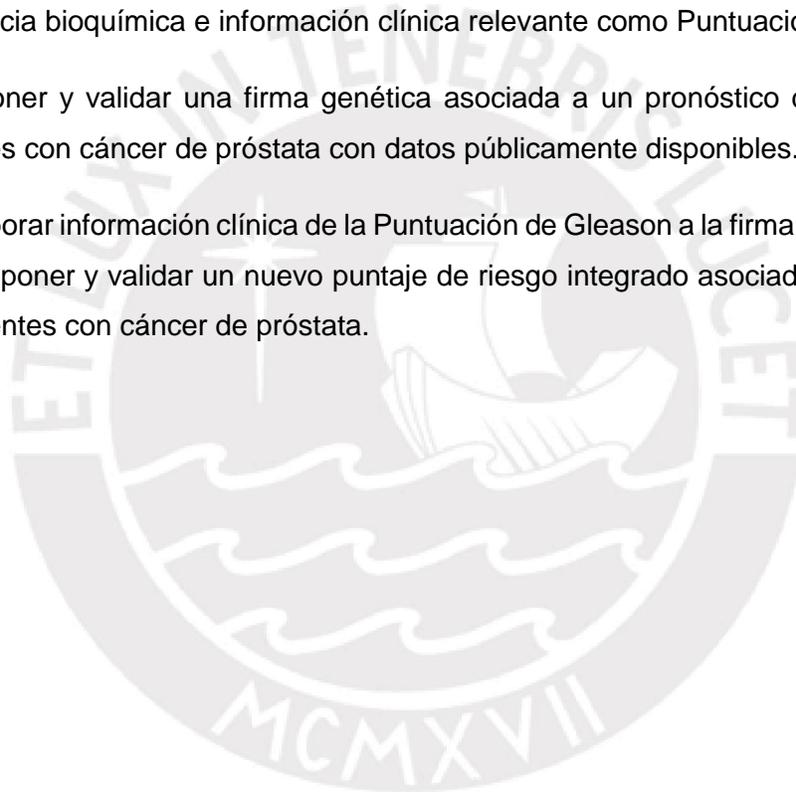
3. OBJETIVOS

3.1. Objetivo General

Identificar y validar biomarcadores genómicos para el pronóstico de recurrencia en cáncer de próstata.

3.2. Objetivos Específicos

1. Identificar y curar series de datos de cáncer de próstata públicamente disponibles que contengan niveles de genes de vías de señalización asociadas a recurrencia, tiempo de recurrencia bioquímica e información clínica relevante como Puntuación de Gleason.
2. Proponer y validar una firma genética asociada a un pronóstico de recurrencia en pacientes con cáncer de próstata con datos públicamente disponibles.
3. Incorporar información clínica de la Puntuación de Gleason a la firma genética validada para proponer y validar un nuevo puntaje de riesgo integrado asociado a la recurrencia en pacientes con cáncer de próstata.



4. PARTE EXPERIMENTAL

4.1. Metodología

4.1.1. Obtención y validación de firma genética para pronóstico de recurrencia en cáncer de próstata

4.1.1.1. Generación de la lista maestra de genes

La generación de la lista maestra de genes involucra a la vía de señalización WNT (KEGG entry: hsa04310) y la vía de señalización que regula la pluripotencia de células madre (KEGG entry: hsa04550). La información de los genes involucrados para cada cascada bioquímica se encontró disponible en la base de datos de la Kyoto Encyclopedia of Genes and Genomes (KEGG). Una vez extraídos los genes para cada caso, se revisó la lista de genes que se extrajeron y se eliminaron los genes repetentes. La lista inicial de genes contaba con 242 genes.³⁸

4.1.1.2. Selección de Cohortes

Se realizó la búsqueda de información genómica públicamente disponible en las bases de datos de cBioPortal web y Gene Expression Omnibus (GEO) sobre adenocarcinomas primarios de próstata. Se seleccionaron pacientes que presentaban el nivel de expresión de genes e información clínica, como la puntuación de Gleason, niveles de antígeno prostático específico, supervivencia libre de recurrencia (RFS, por sus siglas en inglés, Recurrence Free Survival) y el estadio del paciente (libre de enfermedad o recurrente).

Se buscó identificar una cohorte de descubrimiento y al menos dos de validación. En la **Tabla 1**, se muestran las cohortes que se utilizaron para el estudio, así como el número de pacientes de cada una. Como cohorte de descubrimiento, se utilizó la TCGA provisional, y como cohortes de validación se utilizaron dos grupos de pacientes del Memorial Sloan-Kettering Cancer Center (MSKCC).

Tabla 1. Cohortes seleccionadas para el estudio

Cohorte	Tipo de cohorte	Número de pacientes disponibles
The Cancer Atlas Genome (TCGA provisional)	Descubrimiento	499
MSKCC (GSE21032)	Validación	216
MSKCC (GSE54691)	Validación	104

4.1.1.3. Selección de genes y pacientes

Se descargaron y ordenaron en tablas con el programa Microsoft Excel ® los datos genómicos y clínicos. Antes de empezar con el análisis, se excluyeron los pacientes que presentaban estadio metastásico (41 casos) y los que contaban con niveles de expresión de genes nulos (1 caso). Finalmente, la cohorte de descubrimiento se empezó a trabajar con 457 pacientes.

Una vez con la cohorte de descubrimiento definida, se procedió a preseleccionar el grupo de genes que participarían del estudio. Se efectuó la exclusión de genes que presentaban valores nulos o ambiguos. Posteriormente, se calcularon los coeficientes de variación (CV) para de cada gen y se excluyeron aquellos que presentaban $CV < 50\%$. La lista de genes preseleccionados se presentará en la sección de Resultados (**Figura 8**).

4.1.1.4. Normalización de los niveles de expresión de genes

La información de niveles de expresión de genes descargada se normalizó para ver la baja expresión o sobreexpresión de los genes para cada paciente. El primer paso que

se realizó fue la transformación al logaritmo en base 2 (Log2) de los valores de la expresión de los genes. Paso siguiente, se calcularon los valores de la mediana para cada gen. Los valores ambiguos (logaritmos de valores negativos) fueron reemplazados con el valor de la mediana para el gen correspondiente. Finalmente, los valores se centraron en la mediana, es decir, se restó el valor de la mediana al valor transformado en log2. Este procesamiento se realizó con el programa Microsoft Excel® 2010.

4.1.1.5. Obtención de la firma genética para pronóstico de recurrencia en cáncer de próstata

Con la lista final de genes y la base de datos de la cohorte de descubrimiento se realizó un análisis estadístico para identificar aquellos genes relacionados con la recurrencia de cáncer de próstata y así obtener una firma de genes que sirva para pronóstico.

La base de datos que se encontraba en una hoja de cálculos de Excel se copió al programa estadístico IBM® SPSS Statistics 23. Se evaluó la información genética con un análisis univariado de regresión de Cox relacionado con la sobrevida libre de recurrencia. Se seleccionaron aquellos genes estadísticamente significativos con un $P < 0.001$. Los genes seleccionados se sometieron a un método de análisis multivariado de regresión de Cox paso a paso para determinar aquellos genes que sirvan como factores independientes pronósticos. Todos estos pasos realizados para la identificación de los genes asociados a la sobrevida libre de recurrencia fueron para evitar los falsos positivos, es decir, aquellos genes cuyos valores de expresión son estadísticamente significativos pero que no tienen ninguna relación con la recurrencia de este cáncer.

Con los genes seleccionados se obtuvo una combinación lineal de los valores normalizados de la expresión de genes con sus respectivos coeficientes de regresión. Esta combinación lineal se presentará en la sección de resultados y se le otorga el nombre de firma pronóstico.

4.1.1.6. Estimación de la sobrevida libre de recurrencia

Una vez establecida la firma pronóstico, se emplearon los valores de expresión de genes para calcular el riesgo de recurrencia para cada paciente en la cohorte de descubrimiento. Posteriormente, se calculó el valor de la mediana de la cohorte para utilizarla como punto de corte, y así dicotomizar a los pacientes en un grupo de mayor y otro de menor riesgo.

Se estimó la sobrevida libre de recurrencia con el método de Kaplan-Meier. El evento que se definió para el estudio fue la recurrencia o progreso en cáncer de próstata y la censura de la variable fue hacia la derecha. Con la puntuación de riesgo calculada y el punto de corte se graficaron las curvas de supervivencia (**Figura 9**) y se utilizó la prueba Log-rank para comparar ambos grupos de riesgo. Un valor $P < 0.05$ se consideró estadísticamente significativo, lo que quiere decir que sí se puede distinguir ambos grupos de riesgo en la cohorte de pacientes.

4.1.1.7. Validación de la firma pronóstico

La validación de la firma pronóstico se realizó utilizando las bases de datos de validación, cada una por separado (**Tabla 1**). La primera cohorte (GSE21032) de validación contaba con 216 pacientes con adenocarcinoma primario de próstata, pero se excluyeron aquellos que no presentaban la puntuación de Gleason o niveles de expresión de genes. En total solo se evaluaron 130 pacientes para esta cohorte. En el caso de la segunda cohorte (GSE54691), no hubo que excluir ningún paciente por lo que el número de casos fue igual al inicial.

En el caso de las bases de datos de validación, solo se descargaron los niveles de expresión de los genes presentes en la firma pronóstico. Las cohortes 1 y 2 se descargaron de las bases de datos públicas cBioPortal y GEO, respectivamente. En el caso de la cohorte 2, se seleccionaron las sondas correspondientes a cada uno de los genes seleccionados, pero solo se escogió el valor más alto. Posteriormente se realizó la normalización de los datos tal como se mencionó en la **Sección 4.1.1.4**.

Se reemplazaron los niveles de expresión de los genes en la ecuación de la **Sección 4.2.1.2** y se obtuvo el puntaje de riesgo para cada cohorte. Para establecer los grupos de riesgo, se utilizó (1) el valor de mediana correspondiente a cada grupo de pacientes como punto de corte, y se optó por utilizar en un análisis independiente, (2) el valor de la mediana de la cohorte de descubrimiento. Nuevamente, se utilizó el método de Kaplan Meier para estimar la supervivencia libre de recurrencia de las cohortes de validación y se aceptó como estadísticamente significativo aquellos valores de coeficientes $P < 0.05$.

4.1.1.8. Integración de la Puntuación de Gleason para las cohortes estudiadas

Se buscó integrar al puntaje de riesgo de la Ecuación 1 que involucra los tres genes con un parámetro clínico que fue la puntuación de Gleason (GS, por sus siglas en inglés). Se verificó la distribución de la puntuación de Gleason en las 3 cohortes (**Tabla 2**).

Tabla 2. Distribución de la Puntuación de Gleason en las tres cohortes estudiadas

Puntuación de Gleason	<u>Frecuencias</u>		
	TCGA provisional	GSE21032	GSE54691
6	37	41	16
7	229	74	78
8	59	8	4
9	128	7	6
10	4	0	0
TOTAL	457	130	104

La puntuación de Gleason se recodificó en tres nuevos grupos como se muestra en la **Tabla 3**. La recodificación se realizó debido a que a partir de un valor de 6 se le asigna como tumor, y se buscaba diferenciar a partir de la agresividad en cada caso asumiendo que los valores menores a 6 suelen ser hiperplasias benignas y los mayores a 7, cánceres muy agresivos.

Tabla 3. Recodificación de la puntuación de Gleason

Puntuación de Gleason	Recodificación
GS = 6	1
GS = 7	2
GS > 7	3

La nueva recodificación de la puntuación de Gleason de la cohorte de descubrimiento se sometió a un análisis multivariado de regresión de Cox con el puntaje de riesgo obtenido del nivel de expresión de genes. Se obtuvieron coeficientes de regresión para formula una puntuación integrada. Se calculó la mediana de este nuevo puntaje de riesgo integrado y se utilizó como punto de corte para las cohortes de descubrimiento y las de validación. Finalmente, se utilizó el método de Kaplan Meier para comparar las curvas de supervivencia con el puntaje de riesgo integrado en todas las cohortes. Se considero estadísticamente significativo el valor de coeficiente $P < 0.05$.



4.2. Resultados

4.2.1. Obtención y validación de firma genética para pronóstico de recurrencia en cáncer de próstata

4.2.1.1. Selección de genes y pacientes

De los 242 genes de la lista maestra, se excluyeron aquellos que presentaban niveles de expresión nulos en la cohorte de descubrimiento. Para esta etapa se excluyeron 11 genes. Posteriormente, se calculó el coeficiente de variación (CV) para cada gen y así excluir aquellos que presentaban un valor menor a 50% en la cohorte de descubrimiento. Con este último paso, se excluyeron 105 genes y la lista final de genes seleccionados pasó a ser de 126 genes.³⁸ La **Figura 8** muestra el diagrama de flujo de trabajo para la selección de genes para la firma de pronóstico. Asimismo, en la **Tabla 4** se presenta el número de pacientes empleados para cada cohorte.

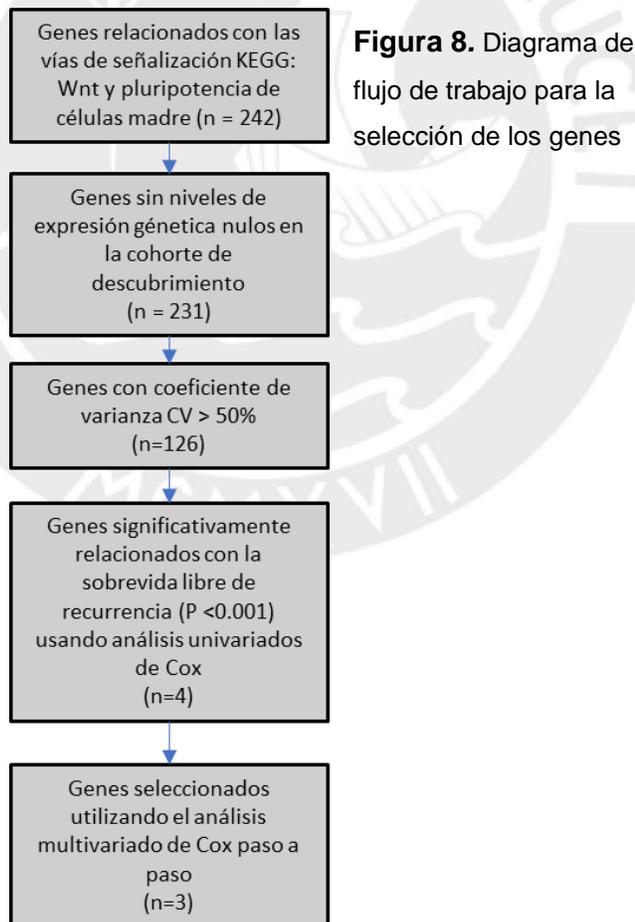


Figura 8. Diagrama de flujo de trabajo para la selección de los genes

Tabla 4. Pacientes empleados de las cohortes seleccionadas para el estudio

Cohorte	Tipo de cohorte	Número de pacientes empleados
The Cancer Atlas Genome (TCGA provisional)	Descubrimiento	457
MSKCC (GSE21032)	Validación	130
MSKCC (GSE54691)	Validación	104

4.2.1.2. Obtención de la firma de pronóstico de genes

El análisis univariado de regresión de Cox para los 126 genes vinculados a la recurrencia del cáncer de próstata identificó cuatro genes relacionados estadísticamente con la RFS ($P < 0.001$). Posteriormente, el análisis multivariado de regresión de Cox que se le realizó a los cuatro genes del paso anterior seleccionó a aquellos genes como factores pronósticos independientes: FZD3, FZD5 y PLCB2. Los genes seleccionados para cada etapa se presentan en la **Tabla 5**.

Tabla 5. Genes seleccionados para cada etapa del análisis

Análisis Univariado	Análisis Multivariado
FZD3	FZD3
FZD5	FZD5
LIFR	PLCB2
PLCB2	

En la **Tabla 6**, se presentan la relación de riesgo (HR, por sus siglas en inglés), el valor P y los coeficientes de regresión para cada uno de los genes seleccionados.

Tabla 6. Valores relación de riesgo, significancia estadística y coeficiente de regresión para los genes seleccionados

Gen	Cohorte de Descubrimiento HR (95% CI)	Valor de P	Coefficiente de Regresión
FZD3	0.559 (0.410 – 0.762)	0.000235	-0.581
FZD5	0.751 (0.583 – 0.967)	0.026252	-0.287
PLCB2	1.542 (1.218 – 1.951)	0.000317	0.433

HR = Hazard ratio (por sus siglas en inglés)

La puntuación de pronóstico de los tres genes resultó de la suma de la multiplicación de los niveles de expresión normalizados de los genes por su respectivo coeficiente de regresión como se describe en la siguiente ecuación:

$$\text{Firma pronóstico: } -0.581 * \text{FZD3} - 0.287 * \text{FZD5} + 0.433 * \text{PLCB2} \quad (1)$$

La mediana del puntaje en la cohorte de descubrimiento fue de 0.0331.

4.2.1.3. Estimación de la sobrevida libre de recurrencia en base a la firma pronóstico

Con el método de Kaplan-Meier se estimó la RFS al graficar las curvas de supervivencia de los dos grupos de riesgos en la cohorte de descubrimiento. En la **Figura 9**, se muestran las curvas de supervivencia para ambos grupos y el valor fue $P = 0.000001$ ($P < 0.05$). La media del tiempo de sobrevida para el grupo de bajo riesgo fue de 132.2 meses, mientras que el valor del grupo de alto riesgo fue de 68.3 meses (**Tabla 7**).

Tabla 7. Medias del tiempo de sobrevida para las cohortes estudiadas

Cohorte	Grupo de Riesgo	Número de pacientes	Media del tiempo libre de recurrencia en meses (IC al 95%)
The Cancer Atlas Genome (TCGA provisional)	Bajo riesgo	226	132.195 (118.387 – 146.002)
	Alto riesgo	225	67.277 (58.045 – 78.509)
MSKCC (GSE21032)	Bajo riesgo	65	107.730 (95.284 – 120.176)
	Alto riesgo	65	106.202 (89.321 – 123.082)
MSKCC (GSE54691)	Bajo riesgo	66	99.329 (89.251 – 109.407)
	Alto Riesgo	38	74.1 (59.830 – 88.371)

*IC: Intervalo de confianza

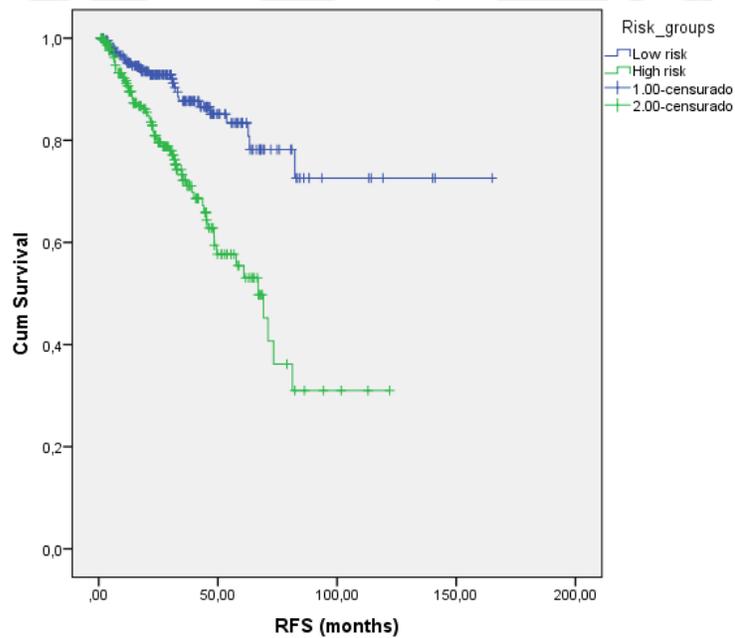


Figura 9. Curvas de supervivencia para la cohorte de descubrimiento. Grupo de bajo riesgo (Azul). Grupo de alto riesgo (Verde). Para ambos casos se utiliza la censura hacia la derecha (rayas verticales)

Para validar la firma de pronóstico de genes se graficaron las curvas de supervivencia con el método de Kaplan-Meier para verificar si podía diferenciar entre los grupos de riesgo para las cohortes de validación. La **Figura 10a** corresponde a la cohorte de validación que consta de 130 pacientes con adenocarcinomas primarios de próstata. La media del tiempo de sobrevida fue 107.7 y 106.2 meses para los grupos de bajo y alto riesgo, respectivamente (Tabla 7).

Por su parte, la **Figura 10b** corresponde a la cohorte de validación que consta de 104 pacientes con la misma enfermedad mencionada anteriormente. La media del tiempo de sobrevida fue de 99.3 y 74.1 meses para los grupos de bajo y alto riesgo, respectivamente (Tabla 7).

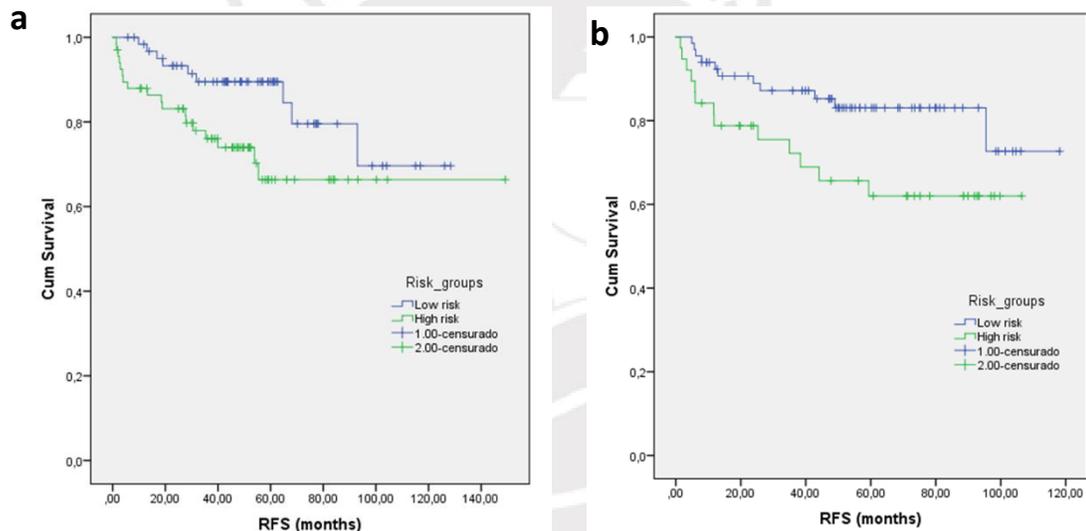


Figura 10. Curvas de supervivencia para las Cohortes de validación por el método de Kaplan-Meier. (a) Cohorte de 130 pacientes ($P = 0,037$) y (b) Cohorte de 104 pacientes ($P=0,035$) con punto de corte en la mediana de la puntuación de riesgo de la cohorte de descubrimiento. Los grupos de bajo riesgo se encuentran de color azul, mientras que los de alto riesgo, de color verde.

Se utilizó la mediana de la puntuación de riesgo de la cohorte de descubrimiento como punto de corte para ambas cohortes de validación para establecer dos grupos riesgos. Los valores P fueron 0.037 y 0.035, lo que significó que eran estadísticamente

significativos al ser menores a 0.05. De esta manera, sí se pudo diferenciar a los pacientes en dos grupos de riesgo.

4.2.1.4. Integración de la Puntuación de Gleason

El análisis multivariado entre la puntuación integrada de riesgo en base a los niveles de expresión de genes y la recodificación de la puntuación de Gleason, nos otorgó coeficientes de regresión para analizar la contribución de cada uno en la nueva puntuación integrada (**Tabla 8**).

Tabla 8. Coeficientes de regresión para la nueva puntuación integrada de riesgo

	Coeficiente de regresión
Puntuación de riesgo (PR)	0.770
Gleason recodificado (GR)	1.075

A continuación, se presenta la Ecuación 2 que relaciona las dos variables mencionadas anteriormente para dar como resultado una puntuación integrada de riesgo:

$$\text{Puntuación integrada} = 0.770 * \text{PR} + 1.075 * \text{GR} \quad (2)$$

Con la puntuación integrada de riesgo definida, se calculó la mediana en la cohorte de descubrimiento y se obtuvo un valor de 2.4621. Este valor se utilizó como punto de corte para establecer los grupos de riesgos en las cohortes de descubrimiento y validación. Finalmente, se obtuvieron las curvas de supervivencia con el método de Kaplan-Meier tanto para la cohorte de descubrimiento (**Figura 11**), así como para las cohortes de validación, para verificar la efectividad del puntaje de riesgo integrado (**Figura 12 y 13**). Los valores del coeficiente P fueron menores a 0.05 por lo que los grupos establecidos con el puntaje integrado en las cohortes de validación sí presentan diferencias en el desenlace del paciente. En la **Tabla 9** se presentan los valores de la media de tiempo de sobrevida para los grupos de riesgo en cada cohorte.

Tabla 9: Medias del tiempo de sobrevida para las cohortes estudiadas empleando la puntuación integrada

Cohorte	Grupo de Riesgo	Número de pacientes	Media del tiempo libre de recurrencia en meses (IC al 95%)
The Cancer Atlas Genome (TCGA provisional)	Bajo riesgo	227	151.099 (143.251– 158.947)
	Alto riesgo	224	62.027 (53.405 – 70.649)
MSKCC (GSE21032)	Bajo riesgo	93	122.708 (107.952 – 137.464)
	Alto riesgo	37	66.849 (52.154 – 81.543)
MSKCC (GSE54691)	Bajo riesgo	94	99.662 (91.459 – 107.865)
	Alto Riesgo	10	12.415 (3.050 – 21.779)

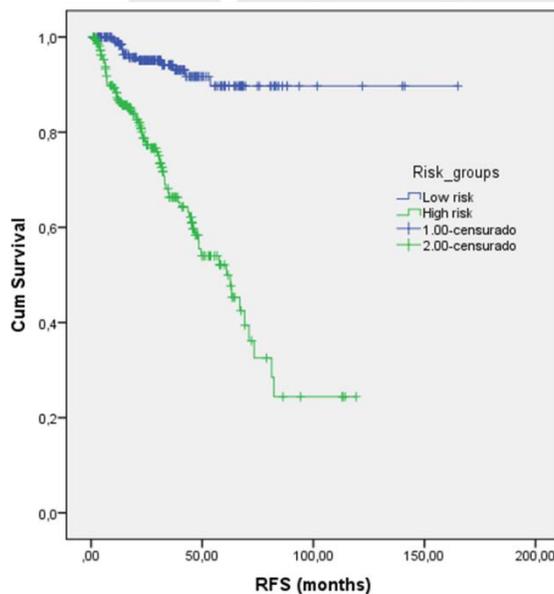


Figura 11. Curva de Supervivencia basada en la nueva puntuación integrada para la cohorte de descubrimiento (457 pacientes empleados y $P < 0.05$). El punto de corte utilizado fue el valor propio de la mediana de esta cohorte. Los grupos de bajo riesgo se encuentran de color azul, mientras que los de alto riesgo, de color verde.

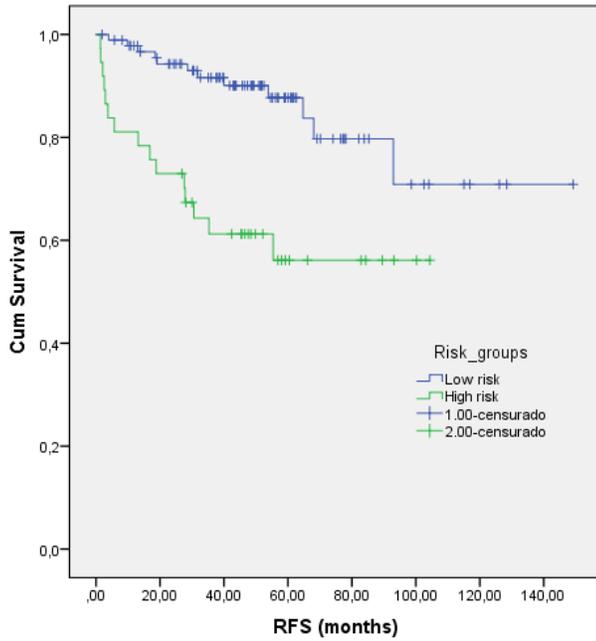


Figura 12. Curva de Supervivencia basada en la nueva puntuación integrada para la cohorte de validación con 130 pacientes ($P < 0.05$). El punto de corte utilizado fue el valor de la mediana de la cohorte de descubrimiento.

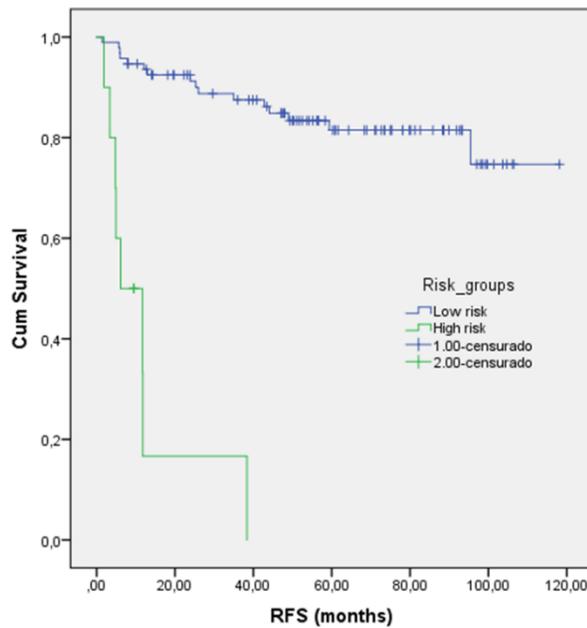


Figura 13. Curva de Supervivencia basada en la nueva puntuación integrada para la cohorte de validación con 104 pacientes ($P < 0.05$). El punto de corte utilizado fue el valor de la mediana de la cohorte de descubrimiento.

5. DISCUSIÓN DE RESULTADOS

5.1. Firma pronóstico y estimación mediante el método de Kaplan Meier

La metodología estadística reportada previamente por Pinto y colaboradores²², quienes obtuvieron una firma pronóstica en cáncer de mama a partir de la observación de que ese cáncer se caracteriza por cambios a nivel del sistema inmune, fue modificada en este estudio para probar la hipótesis novedosa de que cascadas bioquímicas involucradas con la regeneración de células madre podrían estar asociados con la recurrencia en el cáncer de próstata. Los tres genes encontrados fueron FZD3, FZD5 (los receptores de clase frizzled 3 y 5), y PLCB2 (fosfolipasa C beta 2) utilizando la cohorte de descubrimiento. El valor de la relación de riesgo que se obtuvo para estos genes es informativo sobre el rol que cumple cada uno de estos donde aquellos que presenten un HR < 1 tienen un efecto protector, mientras que en el caso de un HR > 1 está asociado con un peor desenlace (efecto negativo). De esta manera, los genes FZD3 y FZD5 tuvieron un efecto protector, mientras que PLCB2 tuvo un efecto negativo (**Tabla 6**).

La firma pronóstico que se obtuvo a partir de la combinación lineal de los niveles de expresión de cada gen con su respectivo coeficiente de regresión proporcionó un puntaje de riesgo para cada paciente. El valor de la mediana del puntaje de riesgo para la cohorte de descubrimiento se utilizó como punto de corte para dicotomizar a los pacientes de las otras dos cohortes y así verificar si nuestra firma pronóstico era válida para una adecuada identificación y separación de pacientes de mejor o peor prognosis. Con el estimador de Kaplan-Meier se estimó el tiempo libre de recurrencia y se observó que, para ambas cohortes de validación, las curvas de supervivencia presentaban una separación estadísticamente significativa ($P < 0.05$). Con esta información, se pudo dicotomizar adecuadamente a los pacientes en dos grupos: uno de alto riesgo (peor prognosis) y otro de bajo riesgo (mejor prognosis).

En la **Tabla 7** se presentan los valores de la media del RFS con su respectivo intervalo de confianza al 95% para cada una de las cohortes. Lo resaltante de esta información es que, en promedio, para el caso de la cohorte de descubrimiento, los pacientes con bajo riesgo podrían vivir aproximadamente el doble de tiempo que aquellos pacientes

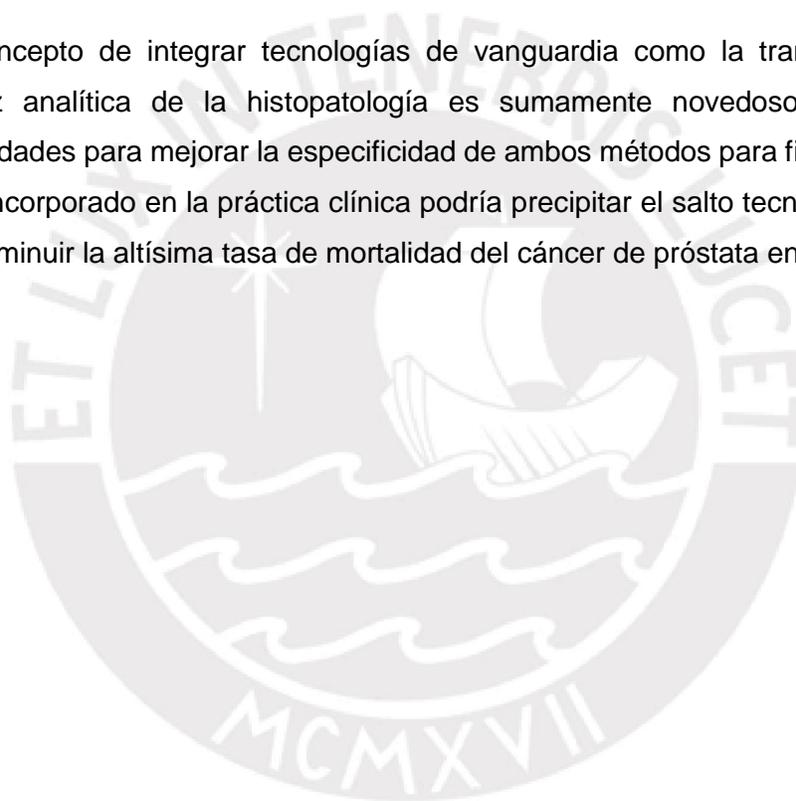
del grupo de alto riesgo antes de que el cáncer reaparezca. No obstante, para el caso de la cohorte de validación 1, a pesar de que se pudo diferenciar los grupos de riesgo, ambos presentaban valores de la media de RFS muy cercanos, con lo cual habría una mínima diferencia de tiempo antes de que reaparezca el cáncer en los pacientes de alto y bajo riesgo. Por último, la cohorte de validación 2 presentó una media de RFS para el grupo de bajo riesgo de 99.3 meses, mientras que para el grupo de alto riesgo fue 74.1 meses. Esto reafirma que los pacientes de bajo riesgo tendrán un mayor tiempo antes de que reaparezca el cáncer.

Para mejorar la identificación de pacientes de buen o mal pronosis, se incorporó un parámetro clínico a la firma pronóstico. El parámetro clínico utilizado que se conoce como la puntuación de Gleason, es un sistema de clasificación de la agresividad del cáncer de próstata en base a información sobre la estructura celular de regiones de tumores tomadas en las biopsias. Los pacientes que evaluamos tuvieron valores entre 6 a 10, ya que bajo estas condiciones el tumor es considerado agresivo. La ecuación resultante fue el puntaje integrado, la cual tenía una contribución de los niveles de expresión de genes (firma pronóstico) y una contribución del parámetro clínico. De esta manera, se calculó nuevamente la puntuación de riesgo integrado para los pacientes de la cohorte de descubrimiento. Posteriormente, se calculó la mediana de este puntaje de riesgo y se utilizó como punto de corte para las cohortes de validación. Nuevamente se utilizó el estimador de Kaplan-Meier para verificar si la nueva puntuación de riesgo era adecuada para diferenciar a los pacientes en grupos de buen o mal pronosis. Como se puede observar en **las Figuras 10, 11 y 12**, la separación entre las curvas de supervivencia se vio mejorada al incluir el parámetro clínico. Es necesario mencionar que los valores de P estuvieron muy por debajo de 0.05, lo cual indicaba que la separación de los grupos de riesgo era estadísticamente significativa. De esta manera, se pudo concluir que con la puntuación integrada se aumenta la capacidad de discriminar a los pacientes y dividirlos en dos grupos de riesgo.

Los valores de la media de RFS para los grupos de riesgo presentaron una notoria y amplia diferencia en todas las cohortes estudiadas (**Tabla 9**). Para el caso de la cohorte de descubrimiento, la media de RFS de las pacientes con bajo riesgo fue más del doble que para los pacientes de alto riesgo. De manera similar, sucedió para el caso de la

cohorte de validación, en el cual los pacientes de bajo riesgo presentaban una media de RFS de aproximadamente el doble al de los pacientes de alto riesgo. La diferencia entre el RFS de los grupos de riesgo para la cohorte de validación 2 fue la más pronunciada: la media de RFS para el grupo de bajo riesgo fue de 99.7 meses, mientras que para el grupo de alto riesgo fue de 12.4 meses. Esta información indica que los pacientes de bajo riesgo van a tener en promedio un mayor tiempo antes de que recurran en cáncer y por lo tanto no van a tener que recibir tratamientos tan agresivos. Por su parte, los pacientes de alto riesgo van a tener que recibir un tratamiento más personalizado debido a que es muy probable que recurran en cáncer en un menor tiempo.

Este concepto de integrar tecnologías de vanguardia como la transcriptómica y la robustez analítica de la histopatología es sumamente novedoso y abre nuevas oportunidades para mejorar la especificidad de ambos métodos para fines pronósticos y de ser incorporado en la práctica clínica podría precipitar el salto tecnológico necesario para disminuir la altísima tasa de mortalidad del cáncer de próstata en el Perú.



5.2. Genes involucrados en las cascadas bioquímicas

El presente estudio se enfocó en evaluar el poder predictivo de recurrencia utilizando genes relacionados con la regeneración de células madres. Las cascadas bioquímicas estudiadas fueron las vías de señalización Wnt y la vía que regula la pluripotencia de células madres. Ambas cascadas están relacionadas con el mantenimiento de células madres, lo cual fue determinante para seleccionarlas en nuestro análisis. De esta manera, se puede evaluar la hipótesis que estas vías, al verse comprometidas, pueden resultar en la resistencia de enfermedades neoplásicas después de recibir tratamientos, como es el caso del cáncer de próstata. Los genes seleccionados son nuevos biomarcadores que serán discutidos en los siguientes capítulos.

5.2.1. TP53

El gen TP53 es el encargado de codificar la proteína p53, la cual cumple el rol de supresora de tumor. Esto quiere decir que regula la división celular al evitar que las células crezcan o se dividan rápidamente o de una manera descontrolada.

La proteína p53 se localiza en el núcleo de las células y se enlaza directamente al ADN. Cuando el ADN en una célula es dañado por algún agente externo (reactivos tóxicos, radiación), esta proteína determina si el ADN será reparado o la célula dañada se someterá a apoptosis. Al detener la proliferación de células con ADN mutado o dañado, p53 ayuda a prevenir el desarrollo de tumores.

Las mutaciones somáticas en el este gen son los cambios genéticos más comunes en el cáncer humano.³⁹ La mayoría de estas mutaciones cambian un único aminoácido en la proteína p53, lo cual lleva a la producción de una versión alterada de la proteína. Esta proteína no puede controlar la proliferación celular y es incapaz de activar la apoptosis en las células mutadas. Como resultado de esto, las células con ADN dañado pueden continuar dividiéndose de manera descontrolada, lo que lleva al crecimiento de un tumor.

Por este motivo, es que este gen se ha estudiado mucho como biomarcador para la enfermedad del cáncer.⁴⁰ En nuestro estudio, se realizó un análisis preliminar de la base de datos, el cual seleccionó al gen TP53 como un buen indicador de pronóstico.⁴¹ Sin embargo, con esa firma pronóstico no se logró una adecuada separación de pacientes

con bajo y alto riesgo. Esto quizás se deba a que estas alteraciones genéticas son más frecuentes en estadios avanzados o subtipos de cánceres de comportamiento agresivo. Para un segundo análisis, se realizó la exclusión de aquellos pacientes con metástasis, seguido de una exclusión de los genes que presentaran un coeficiente de variación menor a 50%. De esta manera, los pacientes seleccionados presentaron un cáncer localizado en la próstata y el gen TP53 fue excluido del análisis. El resultado fue una firma pronóstico sin el gen TP53, la cual tuvo un mejor poder predictivo para la identificación de pacientes de bajo y alto riesgo. La exclusión realizada en la cohorte de descubrimiento fue una medida necesaria, ya que los genes que presentaban una baja heterogeneidad demostraban que los niveles de expresión de los pacientes estudiados no variaban lo suficiente como para ser estudiados como marcadores predictivos de recurrencia.

5.2.2. FZD3 y FZD 5

Los genes FZD3 y FZD5 son aquellos que pertenecen a la familia frizzled, los cuales codifican proteínas que cumplen la función de receptor transmembranal. La cascada bioquímica Wnt se inicia cuando una proteína Wnt se unen al complejo formado por la proteína receptora FZD y al co-receptor LRP5/6. La señal resultante en la vía canónica Wnt resulta en la activación de genes específicos involucrados en la supervivencia, proliferación o invasión celular.

Existen diversos estudios que han reportado que estos dos genes presentaron una regulación positiva en varios tipos de cánceres.⁴² Para el caso de la proteína FZD3 fueron cáncer de pulmón, leucemia primaria aguda y linfoblástica, mieloma linfoma y sarcoma de Ewing. Por otro lado, esta condición se presentó en carcinoma de células renales y en tejidos de cáncer de próstata avanzado para el caso de FZD5.

En el caso de leucemia linfocítica crónica, FZD3 activó varias vías de señalización Wnt/ beta-catenina en presencia de Wnt3 y LRP6. Para el caso de cáncer de mama, FZD3 activó la vía de señalización G α (s) / cAMP/ PKA en presencia de Wnt5a e inhibió la migración celular. Estos estudios muestran que FZD3 puede activar o inhibir a las células cancerígenas en presencia de ligandos y LRP.⁴²

FZD5 ha sido estudiado para casos avanzados de cáncer de próstata. Un estudio reporta que existe una alta expresión del ligando WNT5A al igual que la proteína FZD5.⁴³ Esto indica que los genes involucrados en la vía de señalización Wnt tienen una alta regulación, a diferencia de un tumor primario de próstata.

Estos biomarcadores ya han sido considerados como potenciales biomarcadores terapéuticos para tratamientos de cáncer o como indicadores de cánceres avanzados. El aporte de nuestro estudio es vincular y utilizar estos genes como biomarcadores sobre recurrencia. De esta manera, se logra mejorar la función de los marcadores, ya no solo para cáncer de próstata avanzados, sino para poder diferenciar entre pacientes, en un estadio temprano, aquellos que tengan un mejor o peor pronóstico y que el médico determine las terapias idóneas.



5.2.3. PLCB2

El gen PLCB2 codifica a una proteína clave en el metabolismo de los lípidos de inositol. Esta proteína se conoce como fosfolipasa C β 2 y se encarga de catalizar la hidrólisis del fosfatidilinositol 4,5-bisfosfato a los segundos mensajeros inositol 1,4,5-trisfosfato (IP3) y diacilglicerol. Desempeña un papel crucial en múltiples vías transmembranales de transducción de señales que regulan numerosos procesos celulares, que incluyen proliferación y motilidad.

Los pocos estudios realizados sobre esta proteína han sido realizados en líneas celulares de cáncer de mama.⁴⁴ A partir de este estudio se conoce que la enzima cumple una función estructural ya que tiene un rol en el mecanismo de transducción responsable de la morfología de la célula. Los fosfoinosítidos regulan la actividad de las proteínas de unión a las actinas durante los cambios del citoesqueleto que conducen a la migración. De esta manera, la regulación positiva de esta enzima genera una mayor concentración de estos metabolitos, lo cual va a generar un control desregulado del citoesqueleto de actina; y como consecuencia habrá una mayor motilidad e invasión celular.

Lo mencionado anteriormente puede ser de mucha utilidad al momento de estudiar biopsias de tejido prostático, ya que la información que nos proporcione el puntaje de Gleason se podría complementar con la distribución topográfica de estos metabolitos. El uso de MALDI IMS podría ser de utilidad para estudiar las biopsias de tumor, para incorporar información dimensional con información molecular. El aumento de la resolución de los estudios histológicos proporcionaría una adecuada diferenciación entre un cáncer primario o metastásico, lo cual sería de suma utilidad al momento de seleccionar las terapias adecuadas para el paciente. De esta manera, un aproximamiento al estudio de biopsias de tumor mediante parámetros moleculares mejoraría el diagnóstico de los pacientes ya que se llegaría a una clasificación más acertada del estadio del tumor o el grado de metástasis.

5.2.4 Vinculación de los genes identificados

Los tres genes identificados (FZD3, FZD5 y PLCB2) se introdujeron a una herramienta bioinformática conocida como STRING⁴⁵ para observar la correlación entre las proteínas asociadas a estos genes entre sí y como afectan la cascada bioquímica. La **Figura 14** muestra las asociaciones entre proteínas que se lograron identificar.



Figura 14. Red funcional de asociación de proteínas al introducir los genes FZD3, FZD5 y PLCB2 a la base bioinformática de STRING.

Las proteínas Wnt comprenden una amplia familia de moléculas de señalización que orquestan e influyen en una gran variedad de procesos biológicos y de desarrollo celular. Estas proteínas interactúan con los receptores Frizzled a través de sus dominios ricos en cisteínas en el extremo N de los receptores, lo cual lleva a la iniciación de las distintas vías de señalización hacia adentro de la célula gracias a la proteína de unión a nucleótidos de guanina (proteína G), la cual es codificada por el gen GNAQ. Esta proteína heterotrimérica está compuesta por tres subunidades ($G\alpha$, $G\beta$ y $G\gamma$) y se encuentra acoplada a los receptores transmembranales.⁴⁶ La activación del receptor cataliza el intercambio de GDP por GTP unido a la subunidad α inactiva de la proteína G en el lado intracelular de la membrana. El resultado es un cambio conformacional y una disociación del complejo en una subunidad α activa y un dímero β - γ . La presencia de estas moléculas continúa activando distintos efectores posteriores, tal es el caso de las enzimas fosfolipasas C. Este es uno de los efectores que regula las concentraciones intracelulares de los mensajeros secundarios como el diacilglicerol y que conducen a una respuesta fisiológica.⁴⁶

Este modelo permite vincular las proteínas asociadas a los genes identificados en nuestro estudio para la obtención de la firma pronóstico. Como ya se sabe, la comunicación celular mediada por Wnt y Fzd es crucial para el desarrollo embrionario apropiado y la diferenciación de células madre. También en el organismo adulto, las células madre están reguladas por WNT y también están surgiendo investigaciones que surgieren que las Wnt tienen un papel importante en la regulación de la fisiología básica.⁴⁷ De esta manera, podemos evaluar e identificar algunas partes de las cascadas bioquímicas involucradas para proponer potenciales agentes terapéuticos que interrumpan la auto-renovación indefinidas de células cancerígenas.



CONCLUSIONES

- La información clínica y los niveles de expresión de genes de los pacientes con adenocarcinoma primario de próstata de la cohorte de descubrimiento permitió la identificación de tres genes como potenciales biomarcadores genómicos asociados a la sobrevida libre de recurrencia.
- Con los biomarcadores identificados se obtuvo una firma pronóstico que se utilizó para calcular un puntaje de riesgo para cada paciente. El valor de la mediana de la firma pronóstico se utilizó como punto de corte en las cohortes de validación para dicotomizar a los pacientes en grupos de alto y bajo riesgo. Aquellos pacientes que estuvieron por debajo del valor de la mediana fueron considerados de bajo riesgo, mientras que aquellos que estuvieron por encima fueron considerados de alto riesgo.
- El estimador de Kaplan-Meier permitió validar nuestra firma pronóstico ya que estadísticamente se pudo diferenciar ambos grupos de riesgos en las dos cohortes de validación empleadas. De esta manera, se concluye que nuestra firma pronóstico permitiría diferenciar aquellos pacientes de bajo y alto riesgo de recurrir en cáncer, con el propósito de asignarles las terapias adecuadas para su tratamiento.
- Por último, se logró incorporar el parámetro clínico de la Puntuación de Gleason a la firma en base a información genética, de manera que se obtuvo un puntaje integrado de riesgo utilizando nuevamente la cohorte de descubrimiento. La mediana del puntaje de riesgo integrado se utilizó para separar a los pacientes en dos grupos de riesgo. La ventaja de incorporar este parámetro clínico es que aumenta el poder de discriminar a los pacientes con mejor o peor pronóstico. Finalmente, se logró validar el puntaje integrado con el método de Kaplan-Meier, lo cual nos indicó que es un puntaje de riesgo adecuado y robusto para determinar que pacientes tiene una alto o bajo riesgo de recurrir en cáncer de próstata.

- Los biomarcadores genómicos que se determinaron en este trabajo podrían ser vinculados con biomarcadores metabolómicos para estudios en base a tejidos. En este sentido, lo que se podría realizar para un posterior estudio sería la determinación de estos metabolitos presentes en biopsias de tumor mediante el uso de la técnica de imagenología por espectrometría de masas. El propósito de esto sería estudiar las biopsias y vincular la información molecular con la dimensional. De esta manera, se podría obtener una mejor resolución en los estudios histológicos.
- La identificación de estos marcadores permitirá que estudios posteriores identifiquen posibles agentes terapéuticos para interrumpir la auto-renovación indefinida de células cancerígenas en próstata, de manera que se evite la recurrencia de esta enfermedad neoplásica.



6. BIBLIOGRAFÍA

1. Klausner, R. D. The fabric of cancer cell biology-Weaving together the strands. *Cancer cell* **1**, 3–10 (2002).
2. Hanahan, D. & Weinberg, R. A. The hallmarks of cancer. *Cell* **100**, 57–70 (2000).
3. Ross, D. W. *Introduction to Oncogenes and Molecular Cancer Medicine*. (2012).
4. Hahn, W. C. & Weinberg, R. A. Rules For Making Human Tumor Cells. *The New England journal of medicine* **347**, 1593–1603 (2002).
5. Hanahan, D. & Weinberg, R. A. Hallmarks of cancer: The next generation. *Cell* **144**, 646–674 (2011).
6. Horgan, R. P. & Kenny, L. C. Omic technologies: proteomics and metabolomics. *The Obstetrician & Gynaecologist* **13**, 189–195 (2011).
7. Chen, Y. P. P. *Bioinformatics Technologies*.1-12 (Springer, 2005).
8. Moore, J. H. Bioinformatics. *Journal of Cellular Physiology* **213**, 365–369 (2007).
9. Boja, E. S., Kinsinger, C. R., Rodriguez, H. & Srinivas, P. Integration of omics sciences to advance biology and medicine. *Clinical Proteomics* **11**, 45 (2014).
10. Zhu, Y., Williams, S. & Zwiggelaar, R. Computer technology in detection and staging of prostate carcinoma: A review. *Medical Image Analysis* **10**, 178–199 (2006).
11. Bolla, M. & van Poppel, H. *Management of Prostate Cancer - A Multidisciplinary Approach*. (2012). doi:10.1007/978-3-642-27597-5
12. Bell, K. J. L., Del Mar, C., Wright, G., Dickinson, J. & Glasziou, P. Prevalence of incidental prostate cancer: A systematic review of autopsy studies. *International Journal of Cancer* **137**, 1749–1757 (2015).
13. Ferlay, J., Soerjomataram, I., Ervik, M., Dikshit, R., Eser, S., Mathers, C., Rebelo, M., Parkin, D. M., Forman, D. & Bray, F. GLOBOCAN 2012 v1.0, Cancer

- Incidence and Mortality Worldwide: IARC CancerBase. No. 11 [Internet]. Lyon, France: International Agency for Research on Cancer. **11**, <http://globocan.iarc.fr> (2013).
14. Bashir, M. N. Epidemiology of prostate cancer. *Asian Pacific Journal of Cancer Prevention* **16**, 5137–5141 (2015).
 15. Shirley, A. & Brewster, S. The Digital Rectal Examination. *The Journal of Clinical Examination* **11**, 1–12 (2011).
 16. Scarpato, K. R. & Albertsen, P. C. *Prostate-Specific Antigen Screening Guidelines*. Prostate Cancer (Elsevier Ltd, 2016). doi:10.1016/B978-0-12-800077-9/00013-X
 17. Ismail, M. T. & Gomella, L. G. Transrectal prostate biopsy. *Urologic Clinics of North America* **40**, 457–472 (2013).
 18. U.S. Department of Health and Human Services. NCI Dictionary of Cancer Terms - National Cancer Institute. Available at: <https://www.cancer.gov/publications/dictionaries/cancer-terms/search?contains=false&q=biopsy>. (Accessed: 5th June 2018)
 19. Kanehisa, M., Furumichi, M., Tanabe, M., Sato, Y. & Morishima, K. KEGG: new perspectives on genomes, pathways, diseases and drugs. *Nucleic Acids Research* **45**, D353–D361 (2017).
 20. Cerami, E., Gao, J., Dogrusoz, U., Gross, B. E., Sumer, S. O., Aksoy, B. A., Jacobsen, A., Byrne, C. J., Heuer, M. L., Larsson, E., Antipin, Y., Reva, B., Goldberg, A. P., Sander, C. & Schultz, N. The cBio Cancer Genomics Portal: An open platform for exploring multidimensional cancer genomics data. *Cancer Discovery* **2**, 401–404 (2012).
 21. Aguirre-Gamboa, R., Gomez-Rueda, H., Martínez-Ledesma, E., Martínez-Torteya, A., Chacolla-Huaringa, R., Rodriguez-Barrientos, A., Tamez-Peña, J. G. & Treviño, V. SurvExpress: An Online Biomarker Validation Tool and Database for Cancer Gene Expression Data Using Survival Analysis. *PLoS ONE* **8**, (2013).

22. Pinto, J. A., Araujo, J., Cardenas, N. K., Morante, Z., Doimi, F., Vidaurre, T., Balko, J. M. & Gomez, H. L. A prognostic signature based on three-genes expression in triple-negative breast tumours with residual disease. *npj Genomic Medicine* **1**, 15015 (2016).
23. National Human Genome Research Institute (NHGRI). Biological Pathways. Available at: <https://www.genome.gov/27530687/biological-pathways-fact-sheet/>. (Accessed: 11th September 2017)
24. Greenow, K. & Clarke, A. R. Controlling the Stem Cell Compartment and Regeneration In Vivo: The Role of Pluripotency Pathways. *Physiological Reviews* **92**, 75–99 (2012).
25. Saito-Diaz, K., Chen, T. W., Wang, X., Thorne, C. A., Wallace, H. A., Page-McCaw, A. & Lee, E. The way Wnt works: Components and mechanism. *Growth Factors* **31**, 1–31 (2013).
26. Komiya, Y. & Habas, R. Wnt signal transduction pathways. *Organogenesis* **4**, 68–75 (2008).
27. Klaus, A. & Birchmeier, W. Wnt signalling and its impact on development and cancer. *Nature Reviews Cancer* **8**, 387–398 (2008).
28. Kreso, A. & Dick, J. E. Evolution of the cancer stem cell model. *Cell Stem Cell* **14**, 275–291 (2014).
29. Blum, R., Gupta, R., Burger, P. E., Ontiveros, C. S., Salm, S. N., Xiong, X., Kamb, A., Wesche, H., Marshall, L., Cutler, G., Wang, X., Zavadil, J., Moscatelli, D. & Wilson, E. L. Molecular signatures of prostate stem cells reveal novel signaling pathways and provide insights into prostate cancer. *PLoS ONE* **4**, (2009).
30. Singh, D., Febbo, P. G., Ross, K., Jackson, D. G., Manola, J., Ladd, C., Tamayo, P., Renshaw, A. A., Amico, A. V. D., Richie, J. P., Lander, E. S., Loda, M., Kantoff, P. W., Golub, T. R. & Sellers, W. R. Gene expression correlates of clinical prostate cancer behavior. **1**, 203–209 (2002).

31. Hwang, S.-I., Thumar, J., Lundgren, D. H., Rezaul, K., Mayya, V., Wu, L., Eng, J., Wright, M. E. & Han, D. K. Direct cancer tissue proteomics: a method to identify candidate cancer biomarkers from formalin-fixed paraffin-embedded archival tissues. *Oncogene* **26**, 65–76 (2007).
32. Komisarof, J., Mccall, M., Newman, L., Bshara, W., James, L., Morrison, C. & Land, H. A four gene signature predictive of recurrent prostate cancer. *Oncotarget* **7**, (2016).
33. Ruifrok, A. C. & Johnston, D. A. Quantification of histochemical staining by color deconvolution. *Analytical and quantitative cytology and histology* **23**, 291–9 (2001).
34. Bancroft, J. D., Layton, C. & Suvana, S. K. *Bancroft's theory and practice of histological techniques*. (Churchill Livingstone, 2013).
35. Tabesh, A., Teverovskiy, M., Pang, H.-Y., Kumar, V. P., Verbel, D., Kotsianti, A. & Saidi, O. Multifeature Prostate Cancer Diagnosis and Gleason Grading of Histological Images. *IEEE Transactions on Medical Imaging* **26**, 1366–1378 (2007).
36. Aichler, M. & Walch, A. MALDI Imaging mass spectrometry: current frontiers and perspectives in pathology research and practice. *Laboratory Investigation* **95**, 422–431 (2015).
37. Caprioli, R. M. Imaging mass spectrometry: molecular microscopy for enabling a new age of discovery. *Proteomics* **14**, 807–9 (2014).
38. Casado, F. List of genes related to stem cell behavior. *Harvard Dataverse* (2018). Available at: <https://doi.org/10.7910/DVN/VWQOI0>. (Accessed: 13th September 2018)
39. Olivier, M., Hollstein, M. & Hainaut, P. TP53 Mutations in Human Cancers: Origins, Consequences, and Clinical Use. *Cold Spring Harbor Perspectives in Biology* **2**, a001008–a001008 (2010).

40. Olivier, M., Hollstein, M. & Hainaut, P. TP53 mutations in human cancers: origins, consequences, and clinical use. *Cold Spring Harbor perspectives in biology* **2**, 1–17 (2010).
41. Espinoza-Portocarrero, M. A., Araujo, J. M., Pinto, J. A. & Casado, F. L. In silico validation of a prostate cancer recurrence prognostic signature based on pathways related to stem cells. *Journal of Clinical Oncology* **35**, e23205–e23205 (2017).
42. Ueno, K., Hirata, H., Hinoda, Y. & Dahiya, R. Frizzled homolog proteins, microRNAs and Wnt signaling in cancer. *International Journal of Cancer* **132**, 1731–1740 (2013).
43. Thiele, S., Rauner, M., Goettsch, C., Rachner, T. D., Benad, P., Fuessel, S., Erdmann, K., Hamann, C., Baretton, G. B., Wirth, M. P., Jakob, F. & Hofbauer, L. C. Expression profile of WNT molecules in prostate cancer and its regulation by aminobisphosphonates. *Journal of Cellular Biochemistry* **112**, 1593–1600 (2011).
44. Brugnoli, F., Grassilli, S., Lanuti, P., Marchisio, M., Al-Qassab, Y., Vezzali, F., Capitani, S. & Bertagnolo, V. Up-modulation of PLC- β 2 reduces the number and malignancy of triple-negative breast tumor cells with a CD133+/EpCAM+phenotype: A promising target for preventing progression of TNBC. *BMC Cancer* **17**, 1–11 (2017).
45. Swiss Institute of Bioinformatics, Center for Protein Research & European Molecular Biology Laboratory. STRING: functional protein association networks. Available at: <https://string-db.org/>. (Accessed: 3rd July 2018)
46. Ross, E. M. G α q and phospholipase C- β : Turn on, turn off, and do it fast. *Science Signaling* **4**, 1–3 (2011).
47. Dijksterhuis, J. P., Petersen, J. & Schulte, G. WNT/Frizzled signalling: receptor-ligand selectivity with focus on FZD-G protein signalling and its physiological relevance: IUPHAR Review 3. *British journal of pharmacology* **171**, 1195–209 (2014).

