



**CENTRO DE INVESTIGACIÓN Y DE ESTUDIOS
AVANZADOS DEL INSTITUTO POLITÉCNICO NACIONAL**

DEPARTAMENTO DE TOXICOLOGÍA

**“Linajes del cromosoma Y asociados al riesgo de Cáncer
de próstata en varones mestizos mexicanos”**

T E S I S

Que presenta

IBT. ESMERALDA DE JESÚS ÁLVAREZ TOPETE

**Para obtener el grado de
MAESTRA EN CIENCIAS
EN LA ESPECIALIDAD DE TOXICOLOGÍA**

Co-Tutoras

Dra. María del Rocío Gómez Ortega

Dra. Luisa Torres Sánchez

Asesores:

Dra. Emma S. Calderón Aranda

Dr. Marco Antonio Meraz Ríos

México, D.F., Octubre de 2018

El presente trabajo se realizó en el Laboratorio 14 del Departamento de Toxicología del Centro de Investigación y de Estudios Avanzados del Politécnico Nacional (CINVESTAV), bajo la dirección de la Dra. María del Rocío Gómez Ortega.

Durante el desarrollo de esta investigación se contó con el financiamiento del Consejo Nacional de Ciencia y Tecnología (CONACyT) a través del proyecto de FONDO SECTORIAL DE INVESTIGACION EN SALUD Y SEGURIDAD SOCIAL FOSISS No 261268.

También se contó con el financiamiento del Instituto Nacional de Salud Pública (INSP) a través de los proyectos “Factores de riesgo genéticos y no genéticos para cáncer de próstata en hombres mexicanos” del Fondo Sectorial SS/IMSS/ISSSTE-CONACyT Convocatoria SALUD- 140482 y “RIESGO DE CÁNCER DE PRÓSTATA ASOCIADO A LA SEROPREVALENCIA DE INFECCIÓN POR TRICOMONAS VAGINALIS Y AL POLIMORFISMO rs10993994 DEL GEN MSMB, EN MÉXICO” del Fondo Sectorial SS/IMSS/ISSSTE-CONACyT Convocatoria SALUD-272810.

Esta tesis está dedicada a la memoria de mi hermano
José Antonio Álvarez Topete, a quien llevaré siempre en mi corazón,
TE AMARÉ SIEMPRE CON TODAS MIS FUERZAS
DONDE QUIERA QUE ESTES.

AGRADECIMIENTOS

Agradezco a Dios por permitirme terminar una etapa más en mi vida y guiarme en los momentos de más dificultad y debilidad.

Gracias al amor de mi vida: Miguel García, por siempre apoyarme en cada momento de estrés y angustia, por amarme incondicionalmente y por mostrarme que el amor se expresa de mil formas y entre ellas con la dedicación y perseverancia.

Agradezco también a mis padres: Antonio, Luis Manuel y María, quienes me inculcaron valores y principios, confiaron en mí y creyeron en cada meta que me propuse, ellos fueron los principales promotores de mis sueños.

Agradezco a toda la familia García (Edmundo, Josefina, Andrea, Nina, Zay, Roberto, Melanie) por abrirme las puertas de su corazón enseñándome que la humildad es uno de los valores más hermosos.

Agradezco a mis directoras de tesis la Dra. Rocio Gómez Ortega y la Dra. Luisa Torres Sánchez por su apoyo incondicional en la elaboración del proyecto y sabios consejos, también agradezco a mis asesores académicos: Dra. Emma Calderon Aranda y Marco Antonio Meraz Rios por su comprensión y amistad.

Gracias a mis mejores amigas Nadina y Lauriana, con quienes viví grandes aventuras, sin ustedes mi vida en CINVESTAV hubiera sido gris, las quiero mucho.

Finalmente agradezco a mis compañeros de generación y compañeros de laboratorio, así como a cada persona que formo parte de mi trayectoria, fue un honor conocerlos.

INDICE

| | |
|---|----|
| RESUMEN | 9 |
| ABSTRACT | 10 |
| GLOSARIO | 11 |
| INTRODUCCIÓN | 13 |
| 1.2. Embriogénesis de la glándula prostática y su relación con las hormonas | 14 |
| 1.3 Fisiopatología del Cáncer de Próstata | 16 |
| 1.4 Estadios clínicos del Cáncer de Próstata | 18 |
| 1.5 Epidemiología del Cáncer de Próstata..... | 19 |
| 1.6 Factores de riesgo..... | 22 |
| 1.7 Arquitectura genética del cromosoma Y | 25 |
| 1.8 Microsatélites | 27 |
| 1.9 Linajes ancestrales y Cáncer de Próstata..... | 28 |
| JUSTIFICACIÓN | 31 |
| HIPOTESIS | 32 |
| OBJETIVOS | 32 |
| 4.1 Objetivo general | 32 |
| 4.2 Objetivos particulares..... | 32 |
| METODOLOGÍA | 33 |
| 5.1 Diseño de estudio | 33 |
| 5.1.1 Definición de casos | 33 |
| 5.1.2 Definición de control..... | 33 |
| 5.1.3 Entrevista realizada a cada participante del estudio | 34 |
| 5.1.4 Descripción de variables | 35 |
| 5.2 Estrategia experimental | 36 |
| 5.2.1 Recolección de la muestra. | 36 |
| 5.2.2 Extracción de DNA | 36 |
| 5.2.3 Cuantificación, pureza e integridad del DNA | 36 |
| 5.3 Amplificación por reacción en cadena de la polimerasa | 36 |

| | |
|--|-----------|
| 5.3.1 Principio | 36 |
| 5.3.2 Procedimiento..... | 37 |
| 5.4 Electroforesis Capilar | 37 |
| 5.4.1 Principio | 37 |
| 5.4.2 Procedimiento..... | 38 |
| 5.5 Determinación de los linajes ancestrales | 40 |
| 5.8 Análisis de redes..... | 41 |
| 5.9 Comparación con otras poblaciones | 41 |
| 5.10 Asociación entre CP y alelos de los 17 marcadores Y-STRs | 41 |
| RESULTADOS | 43 |
| 6.1 Características de las poblaciones de estudio..... | 43 |
| 6.2 Diversidad genética poblacional global. | 44 |
| 6.2.1 Diversidad alélica. | 45 |
| 6.2.2 Diversidad haplotípica global..... | 45 |
| 6.2.4 Análisis filogenético..... | 48 |
| 6.3 Análisis Molecular de Varianza (AMOVA)..... | 54 |
| 6.4. Análisis de escala multidimensional (MDS). | 54 |
| 6.4.1 Comparación con otras poblaciones..... | 55 |
| 6.5 Asociación entre el linaje ancestral y el cáncer de próstata. | 59 |
| 6.6 Asociación entre linaje ancestral y la agresividad del cáncer de próstata. | 59 |
| 6.7 Asociación entre linaje ancestral y la edad de presentación del cáncer de próstata 60 | |
| DISCUSIÓN | 62 |
| CONCLUSIONES | 69 |
| PERSPECTIVAS | 70 |
| BIBLIOGRAFIA | 71 |
| INFORMACIÓN SUPLEMENTARIA | 82 |

ÍNDICE DE FIGURAS

| | |
|---|----|
| Figura 1. Anatomía de la próstata humana y constitución de células epiteliales. | 14 |
| Figura 2. Embriogénesis de la próstata..... | 15 |
| Figura 3. Cascada de señalización del Receptor de Andrógenos (AR por sus siglas en inglés Androgen Receptor)..... | 16 |
| Figura 4. Progresión de las etapas del cáncer de próstata humano..... | 17 |
| Figura 5. Incidencia y mortalidad del cáncer de próstata a nivel mundial por cada 100,000 habitantes comparado con otros tipos de cáncer.. | 20 |
| Figura 6. Incidencia del cáncer de próstata en México por cada 100,000 habitantes. ... | 21 |
| Figura 7. Mortalidad del cáncer de próstata en México por cada 100,000 habitantes.. | 21 |
| Figura 8. Distribución del CP en la República Mexicana. Casos registrados por cada 100,000 hombres con edades iguales o mayores a 25 años. | 22 |
| Figura 9. Arquitectura del cromosoma Y..... | 26 |
| Figura 10. Haplogrupos del cromosoma Y..... | 30 |
| Figura 11. Distribución de los casos de acuerdo con la agresividad y a la edad al momento del diagnóstico..... | 44 |
| Figura 12. Red filogenética del haplogrupo E. | 50 |
| Figura 13. Red filogenética del macrohaplogrupo F (haplogrupos G,I, y J) en la población de estudio utilizando 15 Y-STRs.. | 51 |
| Figura 15. Red filogenética del haplogrupo J.A. Se muestra la red filogenética de acuerdo a los linajes derivados de J. B. | 53 |
| Figura 17. Análisis de escalamiento multidimensional del linaje ancestral E entre los grupos de estudio (Casos y controles) y un CG (Control Genómico) , empleando 15 Y-STRs..... | 56 |
| Figura 18. Análisis de escalamiento multidimensional del linaje ancestral I entre los grupos de estudio (Casos y controles) y un CG (Control Genómico), empleando 15 Y-STRs..... | 57 |
| Figura 19. Análisis de escalamiento multidimensional del linaje J entre los grupos de estudio (Casos y controles) y un CG (Control Genómico), empleando 15 Y-STRs.. | 58 |

ÍNDICE DE TABLAS

| | |
|---|----|
| Tabla 1. Sistema de Gradación de Gleason. | 19 |
| Tabla 2. Concentraciones séricas y edad de inicio de la pubertad en diferentes etnicidades..... | 24 |
| Tabla 3. Clasificación de los microsatélites según el tamaño de la repetida..... | 27 |
| Tabla 4. Clasificación de los microsatélites con respecto a la estructura conservada del motif..... | 27 |
| Tabla 5. Características de los marcadores Y-STR..... | 38 |
| Tabla 6. Condiciones de reacción para la realización de la PCR..... | 39 |
| Tabla 7. Condiciones de amplificación empleadas para la genotipificación..... | 39 |
| Tabla 8. Condiciones de corrida de la electroforesis capilar..... | 40 |
| Tabla 9. Características seleccionadas en la población de estudio de acuerdo con la condición de caso-control..... | 43 |
| Tabla 10. Frecuencias alélicas de los 17 marcadores Y-STR encontradas entre los casos (Ca) incidentes de CP y controles (Co) poblacionales de varones mestizos mexicanos..... | 46 |
| Tabla 12. Distribución de haplogrupos o linajes ancestrales de acuerdo con casos y controles..... | 48 |
| Tabla 13. Distribución de haplogrupos o linajes ancestrales de acuerdo con casos y controles..... | 48 |
| Tabla 14. Asociación entre el clan ancestral y el riesgo de cáncer de próstata..... | 59 |
| Tabla 15. Asociación entre el clan ancestral y la agresividad del cáncer de próstata evaluada por la escala de Gleason al momento del diagnóstico..... | 60 |
| Tabla 17. Asociación entre clan ancestral y edad al momento del diagnóstico restringido por edad..... | 61 |

RESUMEN

El cáncer de próstata (CP) es la segunda neoplasia más común en hombres y la quinta causa de muerte por cáncer a nivel mundial, su frecuencia presenta una amplia variedad y parece estar relacionada con los grupos étnicos. Los hombres de origen africano o afroamericano presentan aproximadamente 2 a 3 veces más riesgo de desarrollar cáncer de próstata que los de origen asiático; mientras que los latinos presentan un riesgo intermedio. Este fenómeno parece ser estar relacionado con diferencias en el metabolismo y en la exposición a andrógenos endógenos, de los diferentes grupos étnicos. El cromosoma Y, el cual presentan una herencia haploide, ocupa un papel central en el desarrollo de los órganos sexuales masculinos tales como la próstata y algunos polimorfismos localizados en este cromosoma sexual, así como el linaje ancestral del cromosoma se han sido relacionados con la incidencia y progresión del CP, en poblaciones europeas y asiáticas. Debido a que en México no existen estudios al respecto, en el presente trabajo se exploró la asociación entre los linajes patrilineales y el riesgo de CP en 150 casos incidentes de CP y 383 controles poblaciones, ambos mestizos mexicanos, no relacionados entre sí y residentes de la Ciudad de México. Mediante electroforesis capilar se realizó la discriminación alélica de diecisiete marcadores hipervariables presentes en la región no recombinante del cromosoma Y. Los haplogrupos más frecuentemente observados fueron el Q (~42%), R1b (~30%) y E1b1b (~8%) y ninguno de ellos exhibió una asociación significativa con el riesgo de CP, sino con la agresividad del tumor y la edad al momento del diagnóstico. En comparación con el haplogrupo Q, los portadores del haplogrupo E1b1b presentaron una mayor frecuencia de cáncer menos agresivo ($Gleason \leq 6$) (OR = 4.28; $95\%IC = 1.45 - 12.04$, $p = 0.008$) y una asociación marginal con una edad de diagnóstico mayor o igual a 60 años o de inicio tardío (OR = 1.96; $95\%IC = 0.93-4.10$, $p = 0.07$). Nuestros resultados apoyan el papel de la ascendencia africana en el desarrollo del cáncer de próstata. El haplogrupo E1b1b tiene su origen en el Noroccidente de África y estudios previos con Judíos Ashkenazis y hombres japoneses portadores de este haplogrupo, son consistentes con nuestros resultados.

ABSTRACT

Prostate cancer is the second most common neoplasm in men and the fifth cause of death worldwide. This neoplasm exhibits a multifactorial aetiology where the age, the family history, the hormonal levels, and the ethnicity contributing to their development. Y-chromosome plays a central role in the development of the male sex organs such as the prostate. Hence, different polymorphisms, located in this sexual chromosome, have been related to the incidence and progression of this cancer. Prior studies in Asian and Europe populations have associated the male haplogroup with prostate cancer risk. Nonetheless, these studies are scanty in the Mexican population. Herein, a relation between the Y-chromosome haplotype and the prostate cancer risk was explored using seventeen hypervariable markers. The haplotypes were determined in 533 unrelated mestizo men (150 incidents of prostate cancer and 383 controls) residents in Mexico City. The haplogroups Q (~ 42%), R1b (~ 30%), and E1b1b (~ 8%) exhibited the highest frequencies in the studied population; none of them contributed with the prostate cancer risk. Then, the results were stratifying by haplogroup and aggressiveness degree. The haplogroup E1b1b presented a significant association with the least aggressive cancer (Gleason \leq 6) (OR = 4.28, $_{95\%}$ IC = 1.45 - 12.04, $p = 0.008$). This ancestral lineage also presented a marginal association with the early-onset prostate cancer development (OR = 1.96, $_{95\%}$ CI = 0.93-4.10, $p = 0.07$). Our results support the role of African descent in the development of prostate cancer. Haplogroup E1b1b has its origin in the Northwest of Africa and previous studies with Ashkenazi Jews and Japanese men carrying this haplogroup, are consistent with our results

GLOSARIO

| | |
|-----------|--|
| AGEB: | Áreas Geoestadísticas Básicas |
| AMOVA: | Analysis of molecular variance |
| AR: | Androgen Receptor |
| ASR: | Age-standardized rates |
| CG: | Control Genómico |
| CaMa: | Cáncer de Mama |
| CP: | Cáncer de Próstata |
| CP-IT | Cáncer de Próstata de inicio temprano |
| CP-ITO | Cáncer de Próstata de inicio tardío |
| CYP3A4: | Citocromo P-450 3A4 |
| DHT: | Dihydrotestosterone |
| DNA: | Deoxyribonucleic acid |
| E2: | Estradiol |
| ELAC2: | ElaC Ribonuclease Z 2 |
| FBI: | Federal Bureau of Investigation |
| HG: | Haplogrupo |
| HGM | Hospital General de México |
| Hsp: | Heat Shock Proteins |
| Ht: | Haplotipo |
| IMSS | Instituto Mexicano del Seguro Social |
| INC | Instituto Nacional de Cancerología |
| INNSZ | Instituto Nacional de Medicina y Ciencias de la Nutrición Salvador Zubirán |
| ISSSTE | Instituto de Seguridad y Servicios Sociales de los Trabajadores del Estado |
| JAK/STAT: | Janus Kinase/signal transducers and activators of transcription |
| MAPK: | Mitogen activated protein kinases |
| Mb: | Mega bases |
| MHL: | Minimal Haplotype Loci |
| MPD: | Median Pairwise Difference |
| NKX3.1: | NK3 Homeobox 1 |
| NRY: | Non-recombining Region of the Y chromosome |
| PAR1: | Pseudoautosomic region 1 |
| PAR2: | Pseudoautosomic region 2 |
| PCR: | polymerase chain reaction |
| PI3K/AKT: | Phosphoinositide 3-kinase/ AKT |
| PIA: | Proliferative Inflammatory Atrophy |
| PIN: | Prostatic Intraepithelial Neoplasia |
| PSA: | Prostate-Specific Antigen |
| PTEN: | Phosphatase and Tensin |
| RNASEL: | Ribonuclease L |
| SNP: | Single Nucleotide Polymorphisms |
| SOX9: | SRY-Box 9 |
| SRY: | Sex-Determining Region Y |

| | |
|--------|----------------------------------|
| SSA | Secretaria de Salud |
| STR: | Short Tandem Repeat |
| Y-STR: | Y-chromosome-Short Tandem Repeat |
| ZC: | Zona Central |
| ZP: | Zona Periférica |
| ZT: | Zona de Transición |

INTRODUCCIÓN

1.1 La glándula prostática.

La próstata es una glándula exocrina ubicada en la cavidad pélvica por debajo de la vejiga y , rodeando a la uretra (Taylor, 2017). Esta glándula, presente sólo en el sexo masculino, es la encargada de producir las secreciones seminales que constituyen una gran parte del volumen del semen compuesto por fructuosa y glucosa (los cuales funcionan como nutrientes para los espermatozoides), así como, enzimas que trabajan para descomponer las proteínas del semen después de la eyaculación y productos químicos alcalinos que neutralizan las secreciones vaginales ácidas promoviendo así la supervivencia de los espermatozoides (Hermabessière & Taillandier, 2002).

Esta glándula está conformada básicamente por tres zonas (Figura 1): la zona periférica, que ocupa un 70% del volumen glandular, la zona central, cruzada por los conductos eyaculadores que supone un 25% de la glándula y finalmente la zona transicional, la cual rodea a la uretra con un 5% del volumen glandular (Timms, 2008). Estas regiones están compuestas por ácinos epiteliales, un grupo de células especializadas organizadas a modo de un conjunto de lóbulos, dispuestas en una red estromal fibromuscular (Taylor, 2017). Estos ácinos están altamente organizados y en ellos se encuentran diversos tipos celulares, como las células lumbinales que conforman el 60% del número total de células epiteliales, mientras que el 40% restante está conformado por células madre, células de transición y células basales; las cuales son descritas a continuación (Packer & Maitland, 2016):

- *Células lumbinales*. Son un subtipo celular dependiente de andrógenos que produce y libera proteínas secretoras de la próstata en el lumen, incluido el Antígeno Prostático Específico (PSA, por sus siglas en inglés *Prostate-Specific Antigen*) y serinas, así como la expresión del receptor de andrógenos, CD24 y CD26 (Rybak, Bristow, & Kapoor, 2015). Algunos estudios han demostrado que al haber una desregulación en este tipo celular se generan tumores con un único fenotipo de adenocarcinoma (Park et al., 2016).
- *Células basales*. Este tipo celular se localiza entre las células lumbinales y la membrana basal, expresan CD44, p53 y proteínas CK5/14. Se ha mostrado que este linaje celular genera lesiones tumorales que exhiben una mezcla de regiones escamosas y adenocarcinomas en la glándula prostática (Rybak et al., 2015).

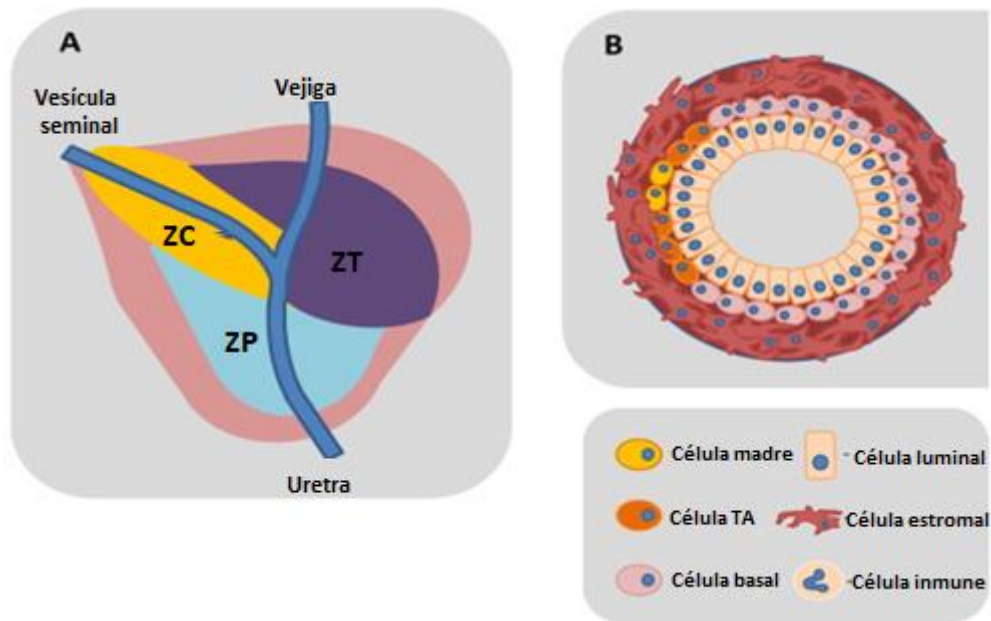


Figura 1. Anatomía de la próstata humana y constitución de células epiteliales. A) Conformación de la glándula prostática: la zona central (ZC), la zona periférica (ZP) y la zona de transición (ZT). B) Representación esquemática de un acino prostático normal y su celularidad (Tomado y modificado de Packer et al., 2016).

1.2. Embriogénesis de la glándula prostática y su relación con las hormonas

La determinación sexual embrionaria masculina está dada por el cromosoma Y, así como, por algunos genes específicos como los SRY (*Sex-Determining Region Y* por sus siglas en inglés). SRY incrementa la expresión de la proteína SOX9 (*SRY-Box 9* por sus siglas en inglés), responsable de la determinación testicular, la diferenciación de las células de Sertoli y Leydig, las cuales tienen un papel central en la producción de hormonas esteroideas (Rey, Josso, & Racine, 2000).

En los humanos, las glándulas accesorias masculinas son las vesículas seminales, las glándulas bulbo uretrales y la próstata (Makaya, 2015). La próstata comienza su desarrollo embrionario aproximadamente a las siete semanas de gestación, y alcanza su madurez funcional y morfológica entre los 18 a 20 años de edad (López & Angulo, 2000). El desarrollo de esta glándula comienza cuando la testosterona fetal a través de los receptores de andrógenos estimula el mesénquima del seno urogenital y la proliferación del epitelio subyacente (Makaya, 2015). Hacia la semana diez del desarrollo, comienzan a crecer las primeras yemas celulares epiteliales de la próstata, las

cuales empiezan a ser notorias a las doce semanas de gestación (Rey & Copelli, 2010). Durante las semanas quince y dieciséis del desarrollo embrionario, la altura del epitelio glandular disminuye, probablemente por efecto de los estrógenos maternos, apareciendo las células triangulares, similares a las células basales de la próstata del adulto, (Figura 2); el crecimiento de estas células permanece constante y se detiene a los dos meses posteriores al nacimiento (López & Angulo, 2000; Rey & Copelli, 2010).

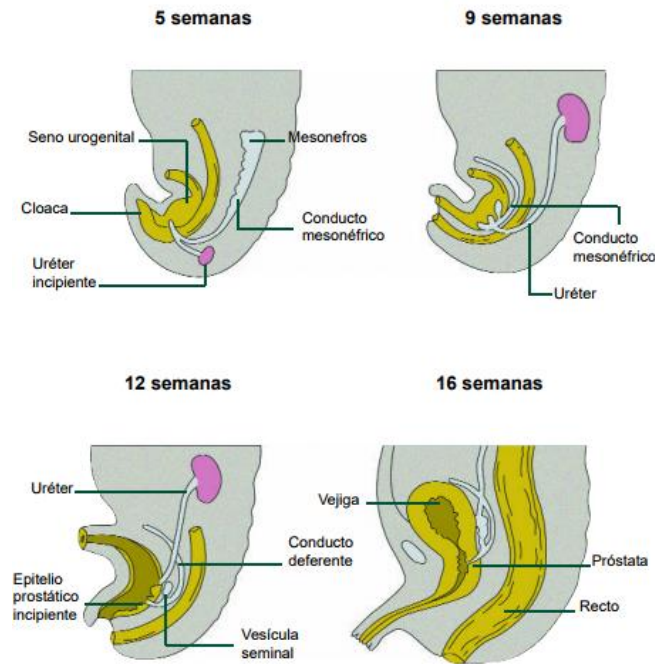


Figura 2. Embriogénesis de la próstata. La glándula se desarrolla en la base de la vejiga a medida que crece desde la uretra y se une con el mesénquima circundante. Tomado de López y Angulo, 2010.

La estructura de la próstata permanece sin cambios en el periodo prepuberal, pero durante la pubertad, inicia cambios morfológicos promovidos principalmente por la 5α -dihidrotestosterona (DHT, por sus siglas en inglés *Dihydrotestosterone*), andrógeno con mayor poder mitogénico (Wu et al., 2013). Estudios previos han señalado que la DTH, actúa principalmente sobre los receptores de andrógenos de las células luminales, induciendo la transcripción de genes diana implicados en respuestas biológicas tales como el crecimiento y la supervivencia de las células prostáticas, así como en la producción del antígeno prostático específico (Figura 3) (Hermabessière & Taillandier, 2002). La DHT se asocia también con cambios proliferativos benignos y malignos en la glándula prostática, así como con una variedad de complicaciones patológicas como el

agrandamiento de la próstata, causada por un aumento en el número de células derivadas de hiperplasia celular y una apoptosis reducida (Carson & Rittmaster, 2003).

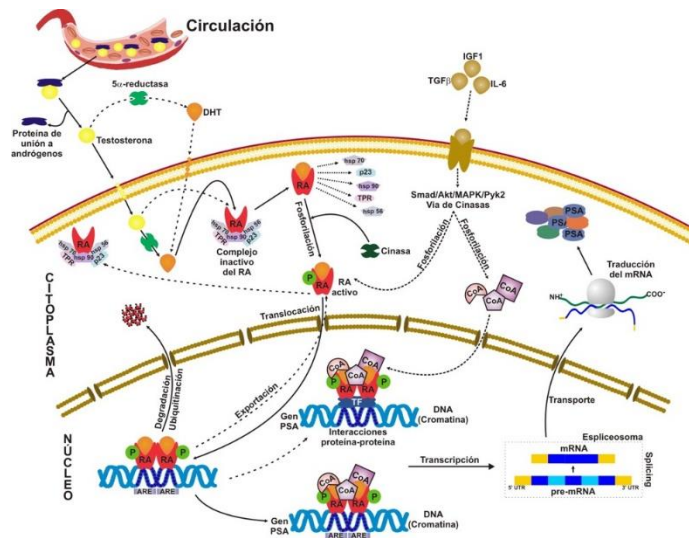


Figura 3. Cascada de señalización del Receptor de Andrógenos (AR por sus siglas en inglés *Androgen Receptor*). El efecto de los andrógenos empieza con el transporte de la testosterona, la cual es metabolizada a DHT, por la enzima 5α-reductasa, transportándose por difusión pasiva al citoplasma. En el citoplasma, monómeros inactivos del AR se encuentran unidos a proteínas chaperonas como Hsp (por sus siglas en inglés *Heat Shock Proteins*). Cuando la DHT se une al AR, éste sufre un cambio conformacional, las Hps se disocian y el AR es fosforilado, permitiendo que el complejo hormona-receptor se transloque al núcleo. En el núcleo, el AR forma dímeros y se une al DNA (por sus siglas en inglés *Deoxyribonucleic acid*) acoplándose a coactivadores o factores de transcripción que ayudan a modular la transcripción del gen. Adaptado de Luke y Coffey, 1994; Heinlein y Chang, 2004; Weigel y Moore, 2007.

1.3 Fisiopatología del Cáncer de Próstata

El Cáncer de próstata (CP) puede iniciar en la porción acinar de la glándula prostática o en los componentes del sistema ductal que drenan a los ácinos (Mustafa et al., 2016). El inicio se caracteriza por una hiperplasia progresiva de la glándula prostática, relacionada con la proliferación de fibroblastos y elementos glandulares epiteliales cerca de la uretra (Packer & Maitland, 2016). Como consecuencia de este crecimiento, la fisiología normal del sistema glandular se modifica (Figura 4), dando lugar a lesiones pre-neoplásicas como la atrofia inflamatoria prostática (PIA, por sus siglas en inglés *Proliferative Inflammatory Atrophy*), caracterizada por una inflamación crónica, cuya frecuencia aumenta con la edad; esta lesión puede progresar a una neoplasia intraepitelial

prostática (PIN, por sus siglas en Inglés *Prostatic Intraepithelial Neoplasia*) (Athanas, B. et al., 2003).

La PIN, se caracteriza por tener células luminales y basales con apariencia casi normal y concentraciones de PSA bajos, por lo que se considera de bajo grado (Velcheti, Karnik, Bardot, & Prakash, 2008). No obstante, la PIN puede también mostrar anomalías genéticas y moleculares similares a las de las células cancerosas, por lo que se consideran de alto grado (Shen et al., 2010). Las lesiones precursoras (PIA y PIN de alto grado) por lo general evolucionan hacia adenocarcinoma invasivo (el tipo más común de CP presente en el 95% de los casos), con pérdida de la membrana basal y cuyo crecimiento modifica la arquitectura normal del sistema glandular prostático, aumenta el tamaño del órgano y obstruye progresivamente la uretra, alterando así el patrón de la eliminación urinaria (Packer & Maitland, 2016; Velcheti et al., 2008).

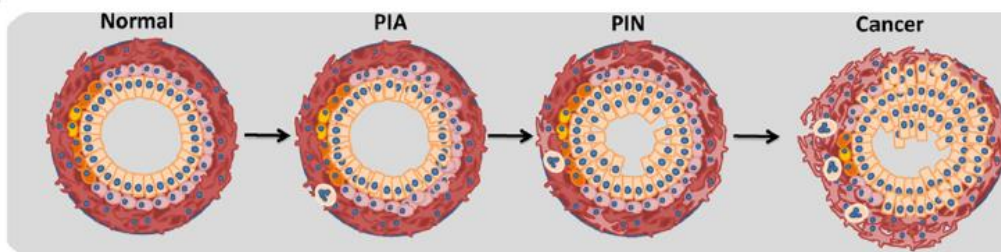


Figura 4. Progresión de las etapas del cáncer de próstata humano. PIA (*Proliferative inflammatory atrophy*, por sus siglas en inglés), PIN (*Prostatic Intraepithelial Neoplasia* por sus siglas en inglés) Tomado de Abate-Shen 2010 y Packer et al. 2016.

El inicio del CP, se vincula con cambios moleculares como la activación de proto-oncogenes e inhibición de genes supresores tumorales implicados en la señalización celular como *PTEN*, (*Phosphatase and Tensin*, por sus siglas en Inglés) y *NKX3.1* (*NK3 Homeobox 1* por sus siglas en inglés), cuya expresión induce la formación de lesiones tipo PIN (Velcheti et al., 2008). Por su parte en la progresión del CP, destaca la alteración en las vías de señalización de PI3K/AKT (*Phosphoinositide 3-kinase/ AKT* por sus siglas en inglés), JAK/STAT (*Janus Kinase/signal transducers and activators of transcription* por sus siglas en inglés) y MAPK (*Mitogen activated protein kinases* por sus siglas en inglés) (da Silva et al., 2013).

1.4 Estadios clínicos del Cáncer de Próstata

Los cambios fenotípicos en la celularidad de la próstata pueden clasificarse mediante el sistema de gradación o puntaje de *Gleason*, propuesto por *Donal Gleason* y *Mellinger* en 1974, y que es actualmente el sistema más empleado en todo el mundo (Bolaños & Chacon, 2017). Este sistema determina que tan alejada de la morfología normal se encuentran las células tumorales. En la medida en que las células cancerosas se asemejan menos al tejido sano el crecimiento del tejido es mucho más rápido y el cáncer se considera que tiene un comportamiento más agresivo (Pierorazio, Walsh, Partin, & Epstein, 2013). El puntaje de *Gleason* además es una herramienta que permite planificar el tratamiento del paciente, y estima la supervivencia de los pacientes (Humphrey, 2004). Con base en esta escala, se reconocen cinco patrones de crecimiento, los que se consolidan en cinco grados (una escala del uno al cinco), que van de bien diferenciado a pobremente diferenciado. Cada patrón se utiliza para generar una puntuación histológica que puede variar de dos (1+1) a diez (5+5) (Humphrey, 2004) (Tabla 1). Para obtener este puntaje, el patólogo selecciona dos zonas de la biopsia de la próstata y observa detalladamente la arquitectura celular bajo el microscopio, para posteriormente asignar a cada una de las zonas un número (del uno al cinco), los valores obtenidos se suman entre sí obteniendo una puntuación histológica (Pierorazio et al., 2013). El puntaje dos es indicativo de un cáncer poco agresivo y con un crecimiento lento, presentando principalmente niveles de PSA <10 ng/mL, mientras que el puntaje diez indica un cáncer altamente agresivo e invasivo con niveles de PSA > 10 ng/mL (Bolaños & Chacon, 2017).

Tabla 1. Sistema de Gradación de *Gleason*.

| Puntaje Gleason | Grado | Diferenciación glandular | Pronóstico |
|--------------------------------------|--------------|--|--|
| Patrón 1 1+1=2 3+3=6 | Grado I | Glándulas bien delimitadas, uniformes, con acinos de tamaño mediano. | Cáncer poco agresivo Crecimiento lento PSA <10 ng/ml |
| Patrón 2 3 + 4 = 7 | Grado II | Predominantemente glándulas bien formadas, con un menor componente de glándulas fusionadas y | Cáncer de mediana agresividad Riesgo intermedio de que el cáncer crezca y se disemine, PSA 10 y 20 ng/ml. |
| Patrón 3 4 + 3 = 7 | Grado III | Glándulas fusionadas y mal formadas. | Cáncer de mediana agresividad Riesgo intermedio de que el cáncer crezca y se disemine, PSA 10 y 20 ng/ mL |
| Patrón 4 4+4= 8 3+5=8 5+3=8 | Grado IV | Glándulas fusionadas, malformadas y pobremente diferenciadas en su mayoría, presentan una | Cáncer agresivo Riesgo alto de que el cáncer crezca y se disemine PSA 20ng/mL |
| Patrón 5 4+5=9 5+4=9 5+5=10 | Grado V | Ausencia de diferenciación glandular | Cáncer de alta agresividad Riesgo muy alto de que el cáncer se disemine a otros órganos. |

PSA= Por sus siglas en inglés Prostatic-Specific Antigen. Adaptado de Bolaños y Chacon et al., 2017; Pierorazio et al.,2013

1.5 Epidemiología del Cáncer de Próstata

El CP es una enfermedad compleja (multifactorial, poligénica y heterogénea) que a nivel mundial es el segundo cáncer más común en hombres mayores de 50 años (31.1 por 100,000 habitantes) y la quinta causa de muerte por cáncer en hombres (7.8 por cada 100,000 habitantes) (Figura 5) (Globocan 2012; García et al., 2007) y representa el 15% de todas las neoplasias en hombres (Ferlay y col., 2012).

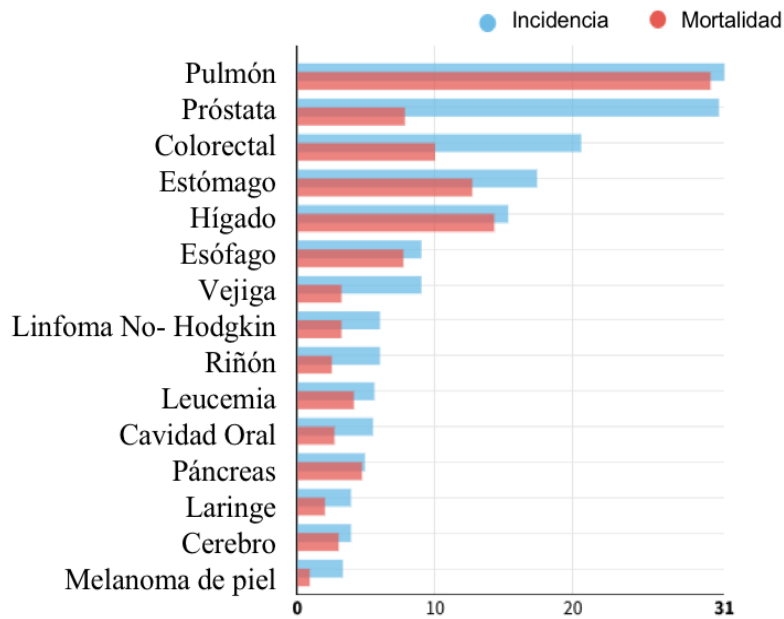


Figura 5. Incidencia y mortalidad del cáncer de próstata a nivel mundial por cada 100,000 habitantes comparado con otros tipos de cáncer. Tomado de Globocan 2012 (<http://globocan.iarc.fr/Default.aspx>).

En México, el CP es la primera causa de cáncer y mortalidad por cáncer en hombres, con una incidencia de 27.3 casos por cada 100 mil habitantes y 11.3 muertes por cada 100 mil habitantes; mostrando ésta última un incremento de aproximadamente 2% anual durante los últimos 10 años (Figuras 6 y 7) (Ferlay et al., 2015).

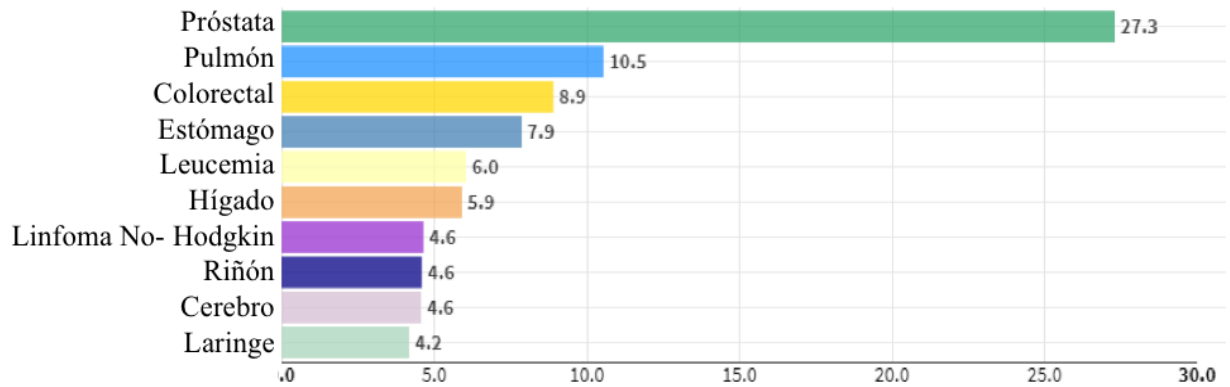


Figura 6. Incidencia del cáncer de próstata en México por cada 100,000 habitantes. Tomado de Globocan 2012 (<http://globocan.iarc.fr/Default.aspx>).

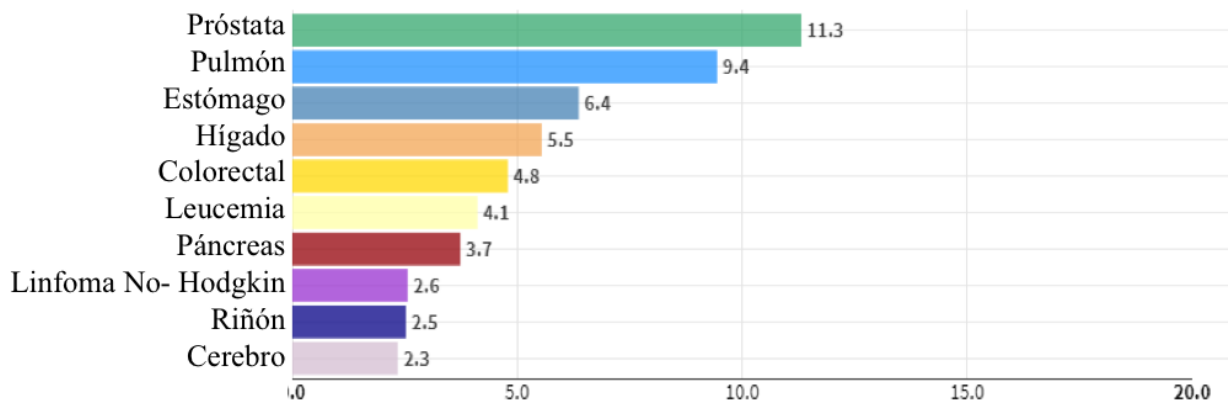


Figura 7. Mortalidad del cáncer de próstata en México por cada 100,000 habitantes. Tomado de Globocan 2012 (<http://globocan.iarc.fr/Default.aspx>).

Aunque en México no se cuenta con un registro nacional de cáncer, los datos del registro histopatológico de cáncer de 2002 muestra una amplia variabilidad en el número de casos de CP reportados por las diferentes regiones del país; el Noreste, el Occidente y el Centro-Norte son las regiones geográficas con mayores tasas de morbilidad estandarizadas por edad (ASR, por sus siglas en inglés *Age-standardized rates*) (HNM, 1993-2002), lo que sugiere que pueden existir diferencias en la prevalencia de ciertos factores de riesgo para la enfermedad. (Figura 8).

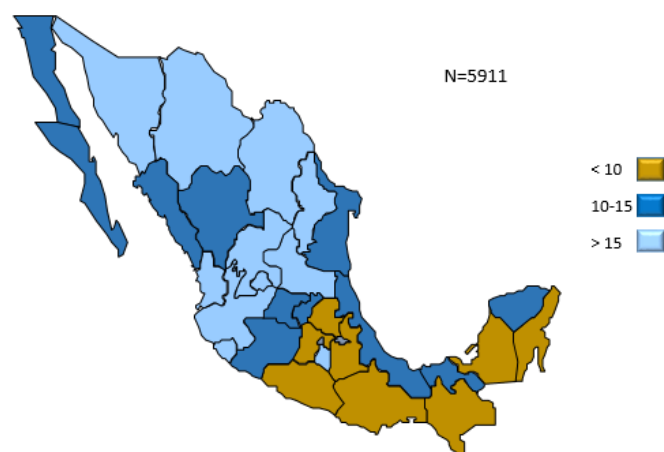


Figura 8. Distribución del CP en la República Mexicana. Casos registrados por cada 100,000 hombres con edades iguales o mayores a 25 años. Tomado del Registro Histopatológico de Neoplasias Malignas, DGAE/SSA, México 1993-2002.

1.6 Factores de riesgo

El CP, como enfermedad compleja, ha sido relacionado con diferentes factores de riesgo, entre ellos aquellos que se consideran modificables como las exposiciones a ciertos contaminantes ambientales (*v.gr.*, el cadmio, los hidrocarburos aromáticos policíclicos, los bisfenoles, y pesticidas (Aimola et al., 2012; Chang & Delzell, 2016; Cheong et al., 2016; Parent, Désy, & Siemiatycki, 2009)), algunos de los cuales se han relacionado con daño a las células epiteliales causado por un elevado estrés oxidativo, provocando una inflamación crónica y con ello la atrofia precoz de la próstata (De Marzo et al., 2007). Del mismo modo algunos compuestos como los bisfenoles, ftalatos y pesticidas, logran imitar y antagonizar hormonas endógenas, alterando la síntesis y metabolismo hormonal, proceso crítico en el desarrollo del CP (Sibbinqi et al., 2002). Características relacionadas con el estilo de vida como (*v.gr.*, la dieta, la obesidad, el sedentarismo, el consumo de alcohol y el tabaquismo), han sido asociadas con el desarrollo de esta patología (Crawford, 2015; Grönberg, 2003).

Con respecto a los factores no modificables se sabe que la edad, el factor genético, la carga hormonal y la etnicidad son los que se encuentran claramente establecidos (Bostwick et al., 2004; Wigle, Turner, Gomes, & Parent, 2008). El incremento en la edad supone uno de los principales factores de riesgo; 70% de los

varones con más de 65 años padece cambios histológicos en la glándula prostática que son consecuencia de un incremento del estrés oxidativo, desencadenando una serie de procesos pro-carcinogénicos (Pienta KJ et al., 1993). En lo referente a los factores genéticos, algunos estudios sugieren que éstos contribuyen hasta con un 42% con el desarrollo del CP, siendo genes como RNASEL (por sus siglas en Inglés, *Ribonuclease L*), ELAC2 (por sus siglas en inglés *ElaC Ribonuclease Z 2*) y AR (*Androgen Receptor* por sus siglas en inglés) los que se han asociado con un mayor riesgo de presentar CP (Beuten et al., 2010; Gómez et al., 2016). De igual manera la historial familiar de CP, y de cáncer de mama (CaMa), han demostrado que incrementan hasta dos veces más el riesgo de presentar CP (Brandt, Sundquist, & Hemminki, 2012) aumentando hasta más de tres veces el riesgo relativo (RR= 2.54 _{95%IC}=1.72-3.75), dependiendo del número de familiares afectados (Thomas et al., 2012).

Como se mencionó previamente, tanto el CaMa como el CP son cánceres hormono-dependientes, por lo que las hormonas sexuales se han considerado un factor de riesgo importante en el desarrollo y progresión del CP (Chen, Page, Chen, & Giovannucci, 2008; Wigle et al., 2008). Particularmente los andrógenos, juegan un papel crítico en el crecimiento, desarrollo, diferenciación y mantenimiento del epitelio prostático (Severi et al., 2006) e influyen en el desarrollo sexual masculino y en la presencia de características sexuales secundarias durante la pubertad (*v.gr.*, aumento en la estatura, acné y salida del vello facial así como púbico entre otras) (Bonilla et al., 2016). Los estrógenos, por su parte, participan de manera activa -al igual que la testosterona- en el crecimiento prostático, sobre todo en el desarrollo de las células epiteliales y en eventos morfogénicos tempranos (Nelles, Hu, & Prins, 2011).

Con respecto a la etnia, a nivel mundial se observa una incidencia variable con respecto al grupo étnico (Eeles et al., 2013). La mayor incidencia se ha reportado entre los Afrodescendientes con 248.5 casos nuevos de CP por cada 100,000 habitantes, mientras que entre Europeos, la incidencia es de 156.7 casos por cada 100,000 habitantes. Las incidencias más bajas se encuentran en hispanos (138 casos por cada 100,000 habitantes) y Asiáticos (~94 casos por cada 100,000 habitantes) (CDC, 2014). Diferencias en genes involucrados en el metabolismo de las hormonas sexuales pueden ser potencialmente responsables de esas diferencias étnicas.

Las concentraciones séricas de testosterona, testosterona libre y DHT son mucho más altas en hombres africanos seguidos por los hombres hispanos y europeos respectivamente y las más bajas se han observado en asiáticos (Litman, Bhasin, Link,

Araujo, & McKinlay, 2006) (Tabla 2), esto se ha explicado a causa de una disminución de la actividad de la 5 α -reductasa, enzima encargada de convertir la testosterona a DHT.

Tabla 2. Concentraciones séricas y edad de inicio de la pubertad en diferentes etnicidades

| Etnia | Pubertad (años) | Testosterona (ng/dL) | DHT (ng/dL) | E2 (ng/dL) | Referencia |
|-------------------|-----------------|----------------------|-------------|------------|------------------------------|
| Africana n=531 | 9-9.5 | 454 | 47.8 | 2.13 | (Litman et al.,2006) |
| Hispana n=648 | 10-10.5 | 441 | 44.4 | 1.16 | (Lopez et al., 2013) |
| Europea n=702 | 10-10.5 | 434 | 43.6 | 1.94 | (Lorente et al., 2015) |
| Asiática n=550 | 10.5-11 | 420 | 41.4 | 0.1522 | (Lichtensztajn et al., 2014) |

*DHT (Dihidrotestosterona), E2 (Estradiol)

La aparición de los caracteres sexuales y algunas características genéticas también varían de acuerdo a la ascendencia étnica (Bonilla et al., 2016; Ragavan et al., 2004). En hombres (Tabla 2) de ascendencia Africana la edad de inicio de la pubertad es más temprana (rango:9 – 9.5 años), seguida por los Hispanos y Europeos (rango:10 – 10.5 años) y finalmente los Asiáticos (rango 10.5 – 11 años), lo cual coincide con la incidencia de CP en estos grupos (Litman et al., 2006; Lopez et al., 2013) y con la evidencia que asocia a la maduración sexual a edades tempranas con el riesgo de presentar CP (Bonilla et al., 2016).

La potencial diferencia en la actividad de algunas enzimas metabolizadoras puede deberse a cambios genéticos que pueden ser heredados y que se presentan de manera diferencial en ciertos grupos étnicos aumentando así el riesgo de desarrollar la enfermedad (Lichtensztajn et al., 2014). Uno de los cambios genéticos directamente asociado con el CP, es el gen del receptor de andrógenos (AR) ubicado en el locus (Xq11-q12). La actividad transcripcional del RA difiere de acuerdo con los grupos étnicos y esto depende del número de repeticiones del *motif* (CAG)₈₋₃₁, localizado en el exón uno del AR (Ackerman et al., 2012). Un menor número de repeticiones del *motif* (≤ 19 repeticiones), lo cual, es mucho más frecuente en sujetos de ascendencia africana, que en Europeos, Asiáticos y Latinos (Ackerman et al., 2012; Ding, Xu, Menon, Reddy, & Barrack, 2004; Harada et al., 2010) se ha asociado con un mayor riesgo de CP (Gómez et al., 2016).

Polimorfismos en el gen que codifica para la proteína CYP3A4, involucrada en el metabolismo y probablemente en la desactivación de la testosterona, también presenta diferencias entre grupos étnicos. En un estudio multiétnico con voluntarios sanos se observó que al menos una copia del alelo G estuvo presente en 81% de los Afroamericanos; mientras este alelo no estuvo presente entre los sujetos asiáticos. Cuando se compararon Afroamericanos con y sin CP, los sujetos con cáncer tuvieron dos veces más frecuencia de portadores del genotipo GG; este genotipo también se asoció con CP más agresivos principalmente en hombres con una edad mayor a los 65 años al momento del diagnóstico del cáncer (Paris et al., 1999).

La íntima relación entre las hormonas esteroideas, los diferentes orígenes étnicos y la susceptibilidad al CP han llevado a sugerir la posible participación del cromosoma Y en la etiología de este tipo cáncer (Sucheston et al., 2012). El cromosoma Y es portador de genes relevantes para el desarrollo masculino tales como el gen SRY, importante en la diferenciación sexual masculina, el gen TSPY (por sus siglas en inglés, *Testis Specific Protein, Y-Linked*) involucrado en la espermatogénesis, y el gen PRY2 (por sus siglas en inglés *PTPN13-Like, Y-Linked 2*) que solo se expresa en testículos, entre otros. En los últimos años, estos genes han sido relacionados con el incremento en el desarrollo de ciertas enfermedades relacionadas con el sexo masculino como la infertilidad, el cáncer de testículo y el cáncer de próstata (Kido & Lau, 2015).

1.7 Arquitectura genética del cromosoma Y

El cromosoma Y es uno de los dos cromosomas sexuales y el más pequeño de los cromosomas humanos (60 megabases, Mb). Abarca más de 59 millones de bloques de construcción de DNA (Por sus siglas en inglés *Deoxyribonucleic acid*) y representa casi el 2% del DNA del genoma haploide (Tiirikka & Moilanen, 2015). Citogenéticamente, el cromosoma Y está formado por dos brazos (uno largo *Yq* y otro corto *Yp*), separados por el centrómero (Colaco & Modi, 2018) (Figura 9). Este cromosoma se encuentra organizado en tres regiones, dos de ellas pseudoautosómicas (o teloméricas) conocidas como PAR1 y PAR2 (por sus siglas en inglés *Pseudoautosomic Region 1 y 2*) de 2.4 y 0.4 Mb, respectivamente. PAR1 está situado en el extremo del brazo *Yp*, mientras que PAR2 se localiza en el extremo del brazo *Yq*. Estas regiones son relativamente cortas y se aparean y recombinan con los PAR del cromosoma X durante

la meiosis (Colaco & Modi, 2018; Kayser, 2017). La tercera región que conforma al cromosoma Y es la región masculino-específica denominada NRY (por sus siglas en Inglés *Non-recombining Region of the Y chromosome*). NRY está ausente de recombinación, pues el cromosoma X carece de homología para recombinarse con el cromosoma Y. Consecuentemente, sigue un modo de herencia holándrica, en la que el haplotipo resultante se hereda de padres a hijos con muy escasas modificaciones (Ross, Bentley, & Tyler-Smith, 2006). A su vez, NRY está conformada por dos sub-regiones (Kayser, 2017; Ross et al., 2006):

1. Región heterocromática: Es una región altamente condesada de aproximadamente 40 Mb, la cual es transcripcionalmente inactiva.
2. Región eucromática: Es una región de 23 Mb que se encuentra ligeramente compactada y es transcripcionalmente activa. Es en esta región es donde se encuentran diversos genes que contribuyen a la formación de las gónadas, la regulación de la espermatogénesis, y el desarrollo del cerebro entre otros (Hanson et al., 2006). En esta región eucromática también se localizan diferentes tipos de polimorfismos altamente variables como los microsatélites, también conocidos como STR (Por sus siglas del inglés: *Short Tandem Repeat*) (Tiirikka & Moilanen, 2015).

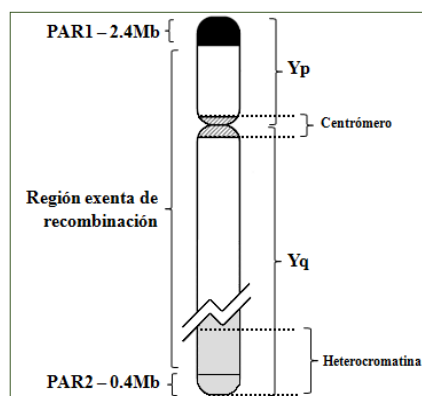


Figura 9. Arquitectura del cromosoma Y. PAR1: Región pseudoautosómicas 1; PAR2: región pseudoautosómicas 2; Yp: Brazo corto del cromosoma Y; Yq: Brazo largo del cromosoma Y.

1.8 Microsatélites

Los STR son secuencias cortas (*motifs*) repetidas en tándem altamente polimórficas. El *motif* de estas secuencias van de uno hasta seis pares de bases (bp, por sus siglas en inglés *base pair*) y se encuentran localizados a lo largo de todo el genoma y se conservan entre diferentes especies (Aranguren-Méndez, J. A., Román-Bravo, R., Isea, W., Villasmil, 2004). Los STR se clasifican dependiendo del tamaño del *motif* de repetición y también con base en su estructura conservada (González, 2003; Tóth, Gáspári, & Jurka, 2000) (Tabla 3 y 4).

Tabla 3. Clasificación de los microsatélites según el tamaño de la repetida.

| Tipo de microsatélite | Tipo de repetida |
|-----------------------|-----------------------|
| Mononucleótidos | (A) _n |
| Dinucleótidos | (AT) _n |
| Trinucleótidos | (CAG) _n |
| Tetranucleótidos | (GAAG) _n |
| Pentanucleótidos | (TACGT) _n |
| Hexanucleótidos | (TGCTAG) _n |

Tabla 4. Clasificación de los microsatélites con respecto a la estructura conservada del *motif*.

| Tipo de microsatélite | Tipo de repetida |
|------------------------|--|
| Puros o perfectos | (AC) _n |
| Puro interrumpido | (CA) _n AA(CA) _n |
| Compuesto | (GT) _n (TG) _n |
| Compuesto interrumpido | (CT) _n (GT) _n CTA(GT) _n |
| Complejos | (ACC) _n TG(GA) _n (TTA) |

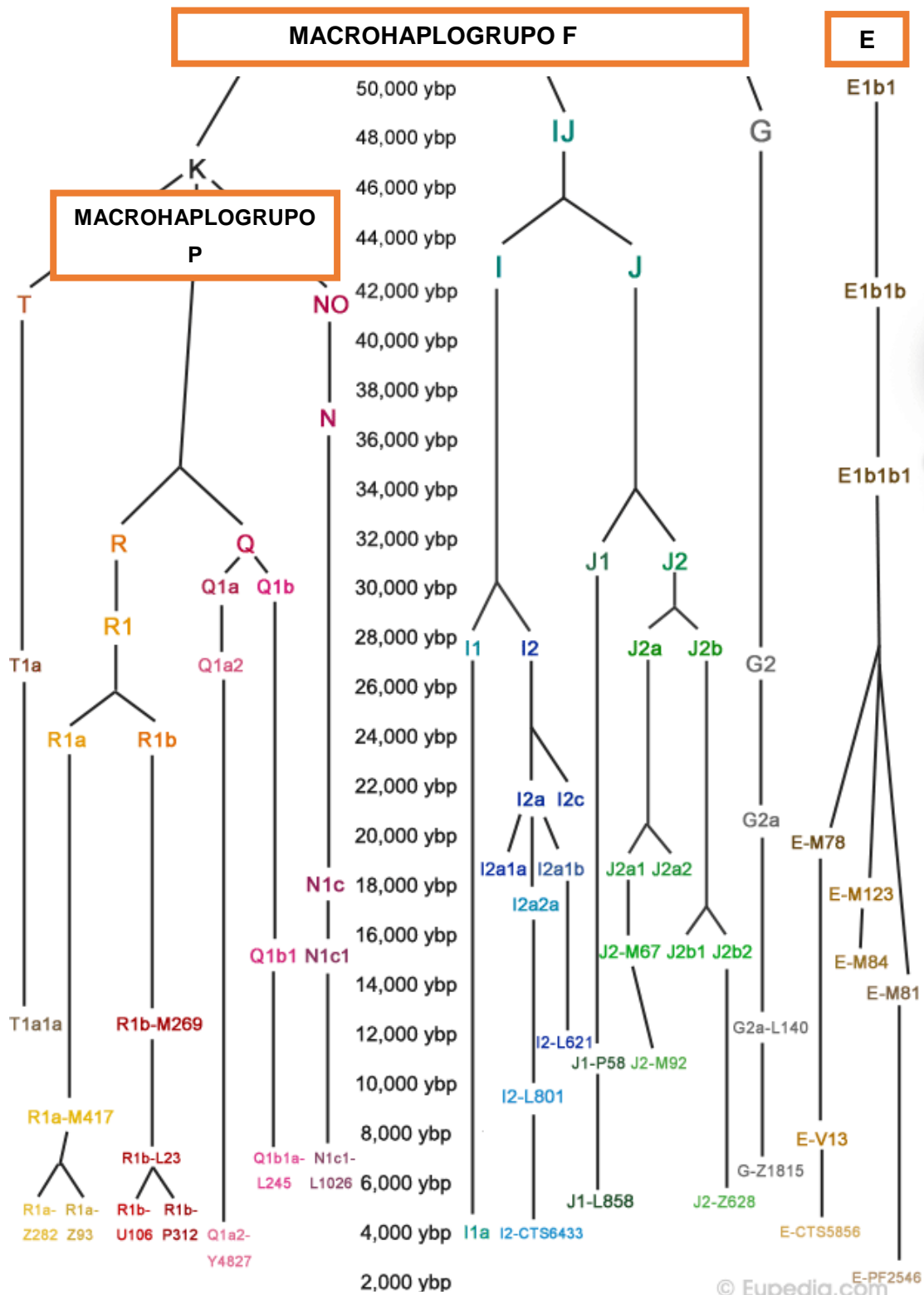
Los STRs específicos del cromosoma Y son conocidos como Y-STR (por sus siglas en Inglés *Y-chromosome short tandem repeats*). Estos polimorfismos hipervariables presentan una herencia haploide y consecuentemente uniparental, por lo que permiten hacer el seguimiento del linaje ancestral masculino al heredarse de generación en generación (Diaz, 2010; Salazar-Flores et al., 2010; C. Santana et al., 2015). Dado que el número de repeticiones en tándem puede ser muy variable entre individuos, los Y-STRs son herramientas genéticas muy útiles en la genética forense, para la determinación de relaciones de parentesco y en la determinación de los tiempos de coalescencia de un determinado linaje, permitiendo asignar la ascendencia de los linajes paternos (haplogrupos o clanes ancestrales) con una alta especificidad geográfica y consecuentemente relacionarla con los orígenes étnicos (Kayser, 2017).

El cromosoma Y al igual que los testículos y la próstata es específico del sexo masculino, lo cual, sugiere que este cromosoma puede tener una participación en el desarrollo del CP (Dasari et al., 2001). Adicionalmente, el cromosoma Y es extraordinariamente pobre en genes, lo que, sugiere que está sujeto a una baja presión selectiva sobre todo en las zonas no codificantes, razón por la cual podría acumular más mutaciones en comparación con otros cromosomas y por lo tanto podría ser de gran utilidad en el diagnóstico de enfermedades complejas como el CP (Jordan, Hanlon, Al-Saleem, Greenberg, & Tricoli, 2001). Estudios citogenéticos han correlacionado la presencia de alteraciones, aberraciones, rearrreglos, ganancias y pérdidas en el cromosoma Y con el desarrollo y progresión de este tipo de cáncer (Forsberg, 2017). La alta heredabilidad del CP y la incidencia diferencial de éste entre los diferentes grupos étnicos, aunado al hecho de que el cromosoma Y tiene un tipo de herencia patrilineal, hacen pensar que los linajes ancestrales podrían contribuir con la susceptibilidad para desarrollar CP y que posiblemente exista una relación entre el linaje, los niveles séricos de andrógenos y potencialmente con la aparición de ciertas características sexuales.

1.9 Linajes ancestrales y Cáncer de Próstata

Basados en las particularidades del cromosoma Y, diferentes grupos de investigación en el mundo han sugerido que los diferentes haplogrupos podrían relacionarse con la susceptibilidad, resistencia, incidencia y desarrollo de algunas enfermedades, entre ellas el CP (Cannon-Albright et al., 2014a). Debido a que los haplogrupos del cromosoma Y son geográficamente específicos, los hombres de

diferentes grupos étnicos tienen diferentes linajes y potencialmente diferentes predisposición a desarrollar CP (Cannon-Albright et al., 2014b). En este sentido, los linajes pertenecientes al macro-haplogrupo DE frecuentemente encontrado en África, pero también con una presencia importante en el Medio Oriente y en Europa (Elhaik et al., 2014; Sucheston et al., 2012), se han relacionado con un mayor riesgo a desarrollar CP (Ewis et al., 2006). El macro-haplogrupo P, del cual derivan los linajes más prominentes de nuestro país, así como el macro-haplogrupo F, que incluye linajes presentes en la población mexicana, también han sido relacionados con un mayor riesgo de presentar CP (Figura 10) (Lindström, Adami, Adolfsson, & Wiklund, 2008; Nargesi et al., 2011; Skaletsky et al., 2003). Sin embargo, a la fecha no hay estudios que relacionen los haplogrupos mexicanos con la susceptibilidad a desarrollar CP.



*ybp (por sus siglas en inglés, *Year Before Present*)

Figura 10. Haplogrupos del cromosoma Y. En esta figura se representan a los macrohaplogrupos F, DE y P así como, la filogenia de cada uno, los cuales tienen una distribución importante en México. Tomado de Eupedia (<https://www.eupedia.com/>)

De igual manera, evidencia reciente sugiere que algunos alelos de los diferentes *loci* del Y-STR han mostrado cierta asociación con el desarrollo de CP (Hasanain, Noran, Imad, & Muhanned, 2015). Particularmente, en población portuguesa, los portadores del alelo 13 presente en el *locus* DYS393 presentaron dos veces más riesgo de CP (OR = 1.97, 95%IC = 1.26-3.07, $p = 0.015$) (Carvalho, Pinheiro, & Medeiros, 2010). Por su parte, los portadores la frecuencia del alelo 14 del *locus* DYS439 en la población de Malasia fue tres veces mayor (OR=3.57, 95% IC = 1.84-6.96, $p < 0.05$) en los sujetos con CP (Nargesi et al., 2011). La expresión diferencial de los genes del cromosoma Y parecen tener también un papel fundamental en el desarrollo del CP (Dasari et al., 2001; Kido & Lau, 2015). Consecuentemente el cromosoma Y pueden ser un medio para escanear los genomas en *locus* conocidos o desconocidos por alguna enfermedad y aportar información de un perfil genético de riesgo (Ewis et al., 2006).

JUSTIFICACIÓN

En México, el CP representa la primera causa de incidencia y mortalidad (27.3 y 11.3 por cada 100,000 respectivamente), con mayor frecuencia reportada en las regiones geográficas del Noreste, Occidente y Centro-Norte del país. La gran mayoría de los cánceres de próstata en México son clasificados como de alto riesgo (Gleason >7) al momento del diagnóstico, lo que repercute en una baja supervivencia (\approx cinco años). La alta heredabilidad del CP y la incidencia diferencial de éste entre los diferentes grupos étnicos podrían contribuir con la susceptibilidad para desarrollar esta patología.

La población mexicana posee una elevada diversidad genética patrilineal como consecuencia de las poblaciones parentales que la originaron (descendientes de Europeos, Africanos y linajes Nativo Americanos) y de las recientes migraciones árabes enriqueciendo la etnogénesis patrilineal de nuestro país, por lo tanto la multiétnicidad de nuestra población nos provee de una oportunidad única para investigar la influencia del *background* étnico y de los diferentes polimorfismos que lo integran con el fin de encontrar posibles biomarcadores de riesgo que contribuyan a la identificación de grupos potencialmente en riesgo de presentar el CP.

HIPOTESIS

La frecuencia del cáncer de próstata en varones mestizos mexicanos varía de acuerdo con el linaje ancestral paterno.

OBJETIVOS

4.1 Objetivo general

Determinar la asociación entre los diferentes linajes del cromosoma Y, y los alelos provenientes del estudio de 17 *loci* Y-STR con la presencia de cáncer de próstata, mediante un estudio de casos y controles en varones Mestizos Mexicanos no relacionados, residentes de la Ciudad de México.

4.2 Objetivos particulares

- ❖ Identificar los linajes patrilineales de 533 hombres Mestizos (150 casos incidentes de CP y 383 controles) mediante 17 Y-STRs.
- ❖ Determinar la diversidad haplotípica de la población de estudio y otros estadísticos genéticos como el número de alelos (k), el número de haplotipos únicos (h), el valor *mean pairwise difference* (MPD) y el cálculo de la varianza genética intra-poblacional.
- ❖ Conocer la relación filogenética de los haplotipos dentro de la población y con otras poblaciones.
- ❖ Establecer una asociación entre el linaje ancestral y el riesgo de CP, su agresividad y la edad al momento del diagnóstico.

METODOLOGÍA

5.1 Diseño de estudio

Entre noviembre del 2011 y agosto del 2014 se llevó a cabo un estudio de casos incidentes de CP y controles poblacionales, ambos residentes en la Ciudad de México. Todos los individuos tenían una edad comprendida entre los 42 y 91 años, sin antecedentes de ningún otro tipo de cáncer al momento del reclutamiento, y sin relación de parentesco aparente.

5.1.1 Definición de casos

Los casos fueron 402 sujetos con diagnóstico de CP incidente histológicamente confirmados, sin ninguna restricción con respecto a su estadio clínico, identificados en la consulta de urología de seis hospitales de la Ciudad de México; dos hospitales de segundo nivel de atención (el Hospital Regional Dr. Carlos Mc Gregor, del Instituto Mexicano del Seguro Social, IMSS y el Hospital López Mateos, del Instituto de Seguridad y Servicios Sociales de los Trabajadores del Estado, ISSSTE) y cuatro hospitales de tercer nivel de atención (el Hospital General de México, HGM, el Instituto Nacional de Cancerología INC, el Instituto Nacional de Medicina y Ciencias de la Nutrición Salvador Zubirán, INNSZ -todos ellos pertenecientes a la Secretaría de Salud (SSA)-, y el Hospital de Oncología del Centro Médico Nacional Siglo XXI, del IMSS).

Con base al puntaje de Gleason al momento del diagnóstico, los casos se clasificaron según su agresividad como: bien diferenciado (Gleason ≤ 6), moderadamente diferenciado (Gleason=7) y pobremente diferenciado (Gleason ≥ 8). Por otro lado, los sujetos cuyo CP fue diagnosticado antes de alcanzar la edad de 60 años, fueron agrupados como CP de inicio temprano (CP-IT), mientras que los sujetos que fueron diagnosticados por CP a edades iguales o mayores a los 60 años se consideraron como CP de inicio tardío (CP-ITO).

5.1.2 Definición de control

Por cada caso se seleccionaron dos sujetos pareados por edad (± 5 años), sin antecedentes de CP y que no estuvieran bajo estudio por sospecha de este o por síntomas urológicos como: disuria, hematuria, o concentraciones de antígeno prostático mayores a 4ng/mL. Los controles fueron identificados en su casa posterior a la identificación del caso. Para ello, se procedió a la identificación de los controles siguiendo el marco

muestral de viviendas que se usa en el Sistema de Encuestas Nacionales de Salud. Este marco se divide en Áreas Geoestadísticas Básicas (AGEB), que son agrupamientos convencionales de 20 a 80 manzanas en las zonas urbanas. Se seleccionaron 33 AGEBS de la Ciudad de México y en cada una de ellas se muestrearon diez manzanas. La selección se realizó de manera sistemática, iniciando por la esquina noreste de cada manzana visitando de manera consecutiva todas las casas hasta encontrar un sujeto que cumpliera con los criterios de inclusión. En caso de que más de un hombre cumpliera con las características requeridas, se llevó a la selección de uno al azar. Si el potencial control no estaba presente al momento de la visita, se realizaron hasta tres intentos para localizarlo; si los tres intentos fueron infructuosos, la búsqueda se abandonó y se procedió a pasar a otra casa. A cada sujeto elegible se le explicó la finalidad y en qué consistiría su participación en el estudio. Los sujetos que no aceptaron participar contestaron una entrevista breve, en la cual, se recabó información acerca de: la edad, el estado civil, la escolaridad y el lugar de nacimiento. El estudio se realizó siguiendo los principios establecidos en la Declaración de Helsinki y contó con la aprobación del Comité de Ética del Instituto Nacional de Salud Pública y de cada uno de los hospitales participantes.

Para fines del presente estudio se realizó una selección de 150 casos y 383 controles, los cuales contaban con muestra suficiente de DNA para realizar las determinaciones genéticas.

5.1.3 Entrevista realizada a cada participante del estudio

A cada participante se le realizó una entrevista estructurada ([Anexo 1](#)), la cual contenía información sociodemográfica (edad, nivel educativo, estado civil, lugar de nacimiento, ocupación habitual a lo largo de la vida y tiempo de residencia en la Ciudad de México). El cuestionario además incluyó preguntas sobre los antecedentes de cáncer en familiares en primer grado (cáncer de próstata, de mama, de ovario y de colón), historia de la vida sexual (edad de la primera relación sexual, número de parejas sexuales, sexo con sexo-servidoras) y antecedentes de infecciones de transmisión sexual (tipo de infección, número de episodios y edad al momento de la infección). También se preguntó acerca de la historia personal de enfermedades crónicas (dislipidemias, diabetes, e hipertensión), así como la historia de tabaquismo, dieta en los tres años previos a la entrevista (para los controles) o al momento del diagnóstico (para los casos) y actividad física a lo largo de la vida. Al finalizar la entrevista a cada participante se le midió la talla,

el peso, la circunferencia de cintura y cadera, además de solicitarle una muestra de sangre venosa.

5.1.4 Descripción de variables

Variable Independiente: Haplotipo y haplogrupo de cada individuo, alelos para cada *locus* de estudio.

Variabes Dependientes: Cáncer de próstata, agresividad del cáncer de próstata y edad de inicio del CP.

Potenciales Variables confusoras: Dentro de estas variables tenemos a la edad, hábito de fumar y antecedentes heredo-familiares.

5.1.5 Descripción operacional de las variables

Variable: Cáncer de próstata

Definición conceptual: Crecimiento anormal y desordenado de las células del epitelio glandular prostático, generando cambios citológicos. Estos cambios son medidos mediante la Escala de Gleason ([Tabla 1](#)) (Pierorazio et al., 2013).

Operacionalización: Casos= 1: El número de sujetos con cáncer de próstata incidente.

Variable: Agresividad del cáncer de próstata

Definición conceptual: Grado de diferenciación celular en el tejido prostático.

Operacionalización: *Gleason* 5-6= 0 : carcinoma bien diferenciado; *Gleason* 7= 1: carcinoma moderadamente diferenciado; *Gleason* 8-10=2 : carcinoma poco diferenciado.

Variable: Edad

Definición conceptual: Número de años cumplidos al momento del diagnóstico de cáncer de próstata.

Operacionalización: Edad < 60 años= 0 : Cáncer de inicio temprano; Edad > 60 años= 1: Cáncer de inicio tardío.

Variable: Haplogrupo

Definición conceptual: Linaje ancestral determinado por una combinación haplotípica, el cual está relacionado con una región geográfica específica (Athey, 2006).

Operacionalización: Se clasificaron de acuerdo a los dieciocho linajes ancestrales encontrados, tomando como referencia para el análisis multivariado el linaje Nativo Americano.

5.2 Estrategia experimental

5.2.1 Recolección de la muestra.

De cada participante, se obtuvieron 10 mL de sangre venosa periférica en tubos *Vacutainer* (Becton Dickinson, Franklin Lakes, NJ, EE.UU.), que contenían ácido etilendiamino tetracético (EDTA) como anticoagulante. A partir de esta muestra, se aislaron las células mononucleares de sangre periférica mediante un gradiente de densidad de Ficoll-hypaque (Hystopaque, Sigma Chemical Co., Sigma-Aldrich, St Louis, MO, EE.UU.).

5.2.2 Extracción de DNA

El aislamiento de material genético se llevó a cabo mediante el reactivo TRIzol (Thermo Fisher Scientific, Suwanee, GA, EE.UU.) siguiendo las instrucciones del fabricante.

5.2.3 Cuantificación, pureza e integridad del DNA

La concentración y pureza de las muestras se evaluaron mediante un espectrofotómetro Nano Drop 2000 (Thermo Fisher Scientific, Wilmington, DE, EE.UU.) tomando en cuenta una relación de absorbancias 260/280 mayor a 1.5. Las muestras que presentaron concentraciones menores a 2ng/ μ L se cuantificaron usando el fluorómetro Qubit 3.0 (Thermo Fisher Scientific, Wilmington, DE, EE.UU.) y empleando el *kit* de ensayo Qubit dsDNA (Thermo Fisher Scientific, Wilmington, DE, EE.UU.) el cual es altamente selectivo para DNA de doble cadena. Todas las determinaciones se hicieron siguiendo las instrucciones sugeridas por el fabricante.

5.3 Amplificación por reacción en cadena de la polimerasa

5.3.1 Principio

La reacción en cadena de la polimerasa o PCR (por sus siglas en inglés *Polymerase Chain Reaction*) es una reacción enzimática *in vitro* que amplifica millones de veces una secuencia específica de DNA durante varios ciclos repetidos en los que la secuencia blanco es copiada fielmente (Lorenz, 2012). Para ello, la PCR aprovecha la actividad de la enzima DNA polimerasa que tiene la capacidad de sintetizar naturalmente el DNA en las células (Tamay de Dios, Ibarra, & Velasquillo, 2013)

Los elementos importantes en la reacción son el templado o molde (DNA), la enzima DNA polimerasa, los oligonucleótidos o *primers*, los desoxirribonucleótidos trifosfatados o dNTPs (adenina, timina, citosina y guanina), el ión magnesio (Mg^{++}), una

solución amortiguadora o *buffer* y el agua. Todos estos elementos interactúan en tres etapas principales de las que se compone la PCR: desnaturalización, hibridación y extensión (Tamay de Dios et al., 2013).

Los equipos en donde se realiza la PCR son llamados termocicladores, los cuales están diseñados para establecer un sistema homogéneo en donde las condiciones de temperatura y el tiempo necesarios no se modifiquen en cada uno de los ciclos. Para confirmar que la reacción ha sido exitosa se lleva a cabo un corrimiento electroforético en geles de agarosa (Mullis K., 1986).

5.3.2 Procedimiento

La obtención del haplotipo de los individuos estudiados se realizó mediante la amplificación simultánea de diecisiete marcadores microsatélite de cromosoma Y (Y-STR), descritos en la [Tabla 5](#). Esta amplificación se llevó a cabo a través de una PCR punto final empleando el termociclador Veriti® de 96 pozos (Applied Biosystems, Carlsbad, CA, EE.UU) haciendo usando del *kit AmpFSTR Yfiler* (Applied Biosystems, Carlsbad, CA, EE.UU.), siguiendo las condiciones de amplificación propuestas por el fabricante. Simultáneamente con los 17 Y-STRs, se llevó a cabo la amplificación del DNA control 007 (Applied Biosystems, Carlsbad, CA, EE.UU.), el cual sirvió como control de calidad para evaluar la eficiencia de la etapa de la amplificación y la asignación alélica.

Las condiciones de reacción para cada una de las muestras se estandarizaron a un volumen final de 6.25 µL. Los volúmenes de reacción de la PCR, así como las condiciones programadas en el termociclador se muestran en la [Tabla 6 y 7](#).

5.4 Electroforesis Capilar

5.4.1 Principio

La electroforesis capilar (EC) es una técnica analítica que permite la separación y cuantificación de una amplia gama de analitos sujetos a un campo eléctrico (de 100 a 500 V/cm). Esta técnica permite el análisis de fragmentos de DNA con pesos moleculares muy parecidos, los que al estar unidos a marcas fluorescentes (6-FAM™, NED™, PET® , VIC®) son detectados diferencialmente por un láser de argón que induce la excitación de los diferentes fluorocromos a diferentes longitudes de onda (λ), permitiendo así el análisis simultáneo de múltiples fragmentos de acuerdo a sus pesos moleculares y al

tiempo de migración a través de un capilar de sílice con un diámetro de 50µm (Magaña, et al 2009).

5.4.2 Procedimiento

Para la realización de esta técnica se empleó el analizador genético 3130XL Applied Biosystems (Applied Biosystems, Carlsbad, CA, EE.UU.). Las condiciones de corrida se aprecian en la [Tabla 8](#). Como control de calidad se empleó un “*Ladder*” con todos los alelos reportados y validados a la fecha (Applied Biosystems, Carlsbad, CA, EE.UU.), el que sirvió para la asignación precisa de todos los alelos que integraban el haplotipo de cada individuo, el que fue visualizado mediante el uso del Software GenMapper® ID v3.2 (Applied Biosystems Carlsbad, CA., EE.UU.), validado por el FBI (*Federal Bureau of Investigation* por sus siglas en inglés).

Tabla 5. Características de los marcadores Y-STR.

| <i>Locus</i> | <i>Motif de repetición</i> | Rango de repeticiones | <i>DYE*</i> |
|--------------|----------------------------|-----------------------|-------------|
| DYS456 | AATA | 11-20 | 6-FAM™ |
| DYS389 I | TCTG | 9-17 | |
| DYS389 II | TCTA | 24-34 | |
| DYS390 | [TCTA][TCTG] | 17-28 | |
| DYS458 | GAAA | 12-21 | VIC® |
| DYS19 | AGAT | 10-19 | |
| DYS385a/b | TCTA | 7-28 | |
| DYS393 | AGAT | 8-17 | NED™ |
| DYS391 | TCTA | 6-14 | |
| DYS439 | AGAT | 8-15 | |
| DYS635 | [TATC][TATG] | 17-27 | |
| DYS392 | TAT | 6-20 | |
| YGATA H4 | TAGA | 8-15 | PET® |
| DYS437 | [TCTA][TCTG] | 12-17 | |
| DYS438 | TTTTTC | 6-14 | |
| DYS448 | AGAGAT | 17-27 | |

* Los colores representan el color del fluorocromo con que están marcados los *primers* de amplificación.

Tomado de Manual Usuario Applied Biosystems, *Amp FLSTR® Y-Filer™*.

Tabla 6. Condiciones de reacción para la realización de la PCR.

| Reactivos | Concentración de los reactivos | Volumen de reacción por muestra |
|--------------------|--------------------------------|---------------------------------|
| Y-Filer PCR Mix | 1X | 2.3 μ L |
| Y-Filer Primer Set | 100 μ M | 1.25 μ L |
| TaqGold | 5 U/ μ L | 0.2 μ L |
| Agua | - | 2 μ L |
| DNA | 2 ng/ μ L | 0.5-1 μ L |
| Volumen final | - | 6.25 μ L |

Tabla 7. Condiciones de amplificación empleadas para la genotipificación.

| Condiciones de PCR | Temperatura | Tiempo | Ciclos |
|---------------------------|-------------|------------|-----------|
| Desnaturalización inicial | 95° C | 11 minutos | 1 ciclo |
| Desnaturalización | 94° C | 1 minuto | 35 ciclos |
| Alineamiento | 61° C | 1 minuto | |
| Extensión | 72° C | 1 minuto | |
| Extensión final | 60° C | 80 minutos | 1 ciclo |
| Enfriamiento | 4° C | ∞ | 1 ciclo |

Al llevar a cabo la inyección de los amplicones, se utilizó un marcador interno de peso molecular GeneScan LIZ-500 (Applied Biosystems, Carlsbad, CA, EE.UU.) que se empleó para obtener el peso molecular de los alelos en bp (*base pair* por sus siglas en inglés) el cual abarca un rango de tamaño de 35-500 bp, conteniendo dieciséis fragmentos de una sola cadena de DNA: 35, 50, 75, 100, 139, 150, 160, 200, 250, 300, 340, 350, 400, 450, 490 y 500 bases. Este marcador ha sido evaluado como un estándar de tamaño de carril interno y proporciona resultados precisos, los que son obtenidos mediante una curva de calibración empleando los dieciséis pesos descritos previamente.

Tabla 8. Condiciones de corrida de la electroforesis capilar.

| Condiciones | Características |
|----------------------------|------------------------|
| Voltaje | 15 Kv |
| Temperatura | 60° C |
| Formamida HI-DI | 9 µL |
| Polímero | POP 7 |
| Marcador de peso molecular | LIZ-500 |
| Tiempo de inyección | 23 segundos |
| Tiempo de corrida | 45 minutos por arreglo |

5.5 Determinación de los linajes ancestrales

Para la determinación del linaje ancestral se creó una base de datos con el haplotipo (Ht) de cada una de las muestras analizadas. Cada Ht fue introducido en los programas *Haplogroup Predictor* (www.hprg.com/hapest5/) (Athey, 2005, 2006) y *NEVGEN* (<http://www.nevgen.org/>). Estos programas asignaron el haplogrupo (HG) usando el número de repetidas (alelos) de cada uno de los diecisiete marcadores Y-STR estudiados, empleando para ellos inferencias bayesianas, tomando los datos de la base de datos *Family Tree DNA* (FTDNA, por sus siglas en inglés) (Tabla 5). La bondad de ajuste y la probabilidad son los parámetros estadísticos que fueron considerados para soportar el grado de acoplamiento de los datos con respecto a los valores teóricos; la probabilidad indica la posibilidad con la que esa asignación se está llevando a cabo. Con el fin de contrarrestar problemas de comparación múltiple con los 17 marcadores Y-STR se realizó la corrección por Bonferroni ($p=0.003$).

5.6 Descripción de la población de estudio

Para la descripción de la población de estudio se consideraron las características sociodemográficas y clínicas como escolaridad, estado civil, lugar de nacimiento y antecedentes familiares de CP tanto de los casos incidentes como de los controles, a partir de pruebas de Chi2 o t-Student.

5.7 Análisis estadístico descriptivo de los marcadores moleculares

La estadística genética consistió en la determinación de las frecuencias alélicas y haplotípicas, el número de alelos (k), el número de haplotipos únicos (h), el valor mean pairwise difference (MPD) así como la diversidad haplotípica (D) presente en la población. También se realizó el cálculo de la varianza genética intrapoblacional (V_p).

Los estadísticos descritos se obtuvieron empleando ARLEQUIN v3.5 (Excoffier and Lischer. 2010). El número de haplotipos únicos fue estimado por conteo directo.

5.8 Análisis de redes

Con el fin de conocer la relación filogenética de los haplotipos dentro de la población y con otras poblaciones, así como los patrones de diversidad y las edades de los haplogrupos, se construyeron redes haplotípicas usando Network 4.6.1.3 (Bandelt et al., 1999) y se editaron en Network Publisher v1.3.0.0 (Fluxus Technology Ltd). En la construcción de la red, cada marcador fue ponderado por el inverso de su varianza, determinado por la diversidad de alelos presentes en el mismo (Roewer et al., 2013). El análisis de redes se basó en los marcadores DYS19, DYS389I, DYS389II, DYS390, DYS391, DYS392, DYS393, DYS437, DYS438, DYS439, DYS448, DYS456, DYS458, DYS635 y GATA-H4 excluyendo los *loci* DYS385a/b debido a que presentan un *locus* STR duplicado. Se aplicaron las opciones de contracción, unión media y máxima parsimonia en el análisis de los datos.

5.9 Comparación con otras poblaciones

Con la intención de comparar nuestros datos con otras poblaciones, se emplearon los datos de informes previos tanto en población mexicana como en 44 diferentes poblaciones parentales (Nativos Americanos, Europeos y Africanos (Tabla Suplementaria S1). La comparación entre poblaciones se analizó por medio de distancias genéticas R_{ST} , el cual es un modelo de mutación gradual generalizado, utilizando ARLEQUIN 3.5 (Excoffier and Lischer. 2010), para posteriormente analizar estas distancias a través de un escalamiento multidimensional (MDS por sus siglas en inglés *Multidimensional Scaling*) utilizando el software SPSS (por sus siglas en inglés *Statistical Package for the Social Sciences*) V.20 (IBM, Armonk, NY, EE.UU). Finalmente se llevó a cabo un análisis de varianza molecular (AMOVA por sus siglas en inglés *Analysis of molecular variance*) para conocer la estructura genética de las poblaciones.

5.10 Asociación entre CP y alelos de los 17 marcadores Y-STRs

La asociación entre el linaje ancestral y el riesgo de CP se evaluaron mediante una prueba de razón de verosimilitud basada en un modelo de regresión logística no condicionada, ajustada por edad al momento del diagnóstico con el software de STATA v12.0. Con la finalidad de evaluar la asociación entre los linajes encontrados y la

agresividad del cáncer de próstata se corrieron tres modelos independientes de regresión logística no condicionada, en los que cada grupo de casos: bien diferenciado (Gleason ≤ 6), moderadamente diferenciado (Gleason=7) y pobremente diferenciado (Gleason ≥ 8), se compararon de manera independiente con todos los controles. Del mismo modo para evaluar la asociación entre los linajes con la edad al momento del diagnóstico se corrieron modelos independientes en los que se compararon a los casos de CP de inicio temprano < 60 años y CP de inicio tardío > 60 años con todos los controles.

Para realizar las asociaciones se consideraron tres linajes ancestrales categorizados como Nativo Americano (Q), Europeo (R) y Afroascendente (E), tomando como categoría de referencia el linaje Nativo Americano por su alta frecuencia en la población de estudio.

Con el fin de disminuir asociaciones espurias, las frecuencias alélicas de nuestros datos fueron comparados con las frecuencias reportadas previamente por Santana y colaboradores (Carla Santana et al., 2014). Estas frecuencias corresponden a la zona geográfica del Centro del país, la que incluye los estados de Querétaro, Puebla y Guanajuato.

RESULTADOS

6.1 Características de las poblaciones de estudio.

Debido al diseño del estudio, no se encontraron diferencias significativas ($p=0.24$) entre los casos y los controles con respecto a la variable edad. Respecto a la escolaridad, el contar con un grado de estudios de licenciatura o incluso mayor aumentó marginalmente las posibilidades de presentar CP (RM= 1.75, 95% IC=0.95-3.22, $p= 0.07$); no se observaron diferencias significativas en relación al estado civil. Comparado con los sujetos nacidos en la Ciudad de México, aquellos que nacieron en la región Central-Este tuvieron 2 veces más posibilidades de CP (RM= 2.14, 95% IC=1.20-3.83, $p= 0.01$). Finalmente, los hombres con antecedentes familiares de CP tuvieron hasta cuatro veces más posibilidades de tener CP (RM= 4.24, 95% IC=1.93-9.3, $p \leq 0.01$) en comparación con aquellos que no tenían este antecedente (Tabla 9).

Tabla 9. Características seleccionadas en la población de estudio de acuerdo con la condición de caso-control.

Características seleccionadas en la población de estudio de acuerdo con la condición de caso-control.

| Variables | Casos n= 118 | Controles n= 321 | RM ^a | IC 95% | Valor <i>p</i> |
|----------------------------------|-----------------|---------------------|-----------------|------------------|------------------|
| Edad (años) | | | | | |
| Media ± DE | 65.45 ± 9.12 | 66.58 ± 8.85 | 0.99 | 0.96-1.01 | 0.24 |
| Escolaridad (%) | | | | | |
| Primaria | 52 (44.07) | 147 (45.80) | 1 | - | - |
| Secundaria | 17 (14.41) | 77 (23.99) | 0.59 | 0.32-1.11 | 0.1 |
| Preparatoria | 25 (21.19) | 60 (18.69) | 1.11 | 0.62-1.97 | 0.73 |
| Licenciatura | 24 (20.34) | 37 (11.53) | 1.75 | 0.95-3.22 | 0.07 |
| Estado Civil (%) | | | | | |
| Unidos vs. No unidos | 88 (74.58) | 251 (78.19) | 0.81 | 0.50-1.33 | 0.41 |
| Lugar de nacimiento ^b | | | | | |
| Ciudad de México | 70 (59.32) | 208 (64.80) | 1 | -- | -- |
| Sur | 10 (8.47) | 30 (9.35) | 1.08 | 0.50-2.35 | 0.84 |
| Central-Oeste | 4 (3.39) | 25 (7.80) | 0.49 | 0.16-1.46 | 0.2 |
| Central-Este | 26 (22.03) | 40 (12.44) | 2.14 | 1.20-3.83 | 0.01 |
| Norte | 1 (0.85) | 7 (2.18) | 0.46 | 0.05-3.85 | 0.48 |
| Este | 7 (5.93) | 11 (3.43) | 1.96 | 0.73-5.30 | 0.18 |
| Antecedentes familiares de CP | | | | | |
| Si vs. No | 14 (11.8) | 8 (2.50) | 4.24 | 1.93-9.30 | < 0.01 |

Los datos en negritas corresponden a aquellos que presentaron diferencias estadísticamente significativas con valores de $p \leq 0.01$

^a Razón de momios (RM) ajustada por edad al momento de la entrevista.

^b Lugar de nacimiento: Ciudad de México. Sur: Campeche, Chiapas, Guerrero, Oaxaca, Quintana Roo y Yucatán. Central Oeste: Aguascalientes, Colima, Guanajuato, Jalisco y Michoacán. Central Este: Hidalgo, Estado de México, Morelos, Puebla, Querétaro y Tlaxcala. Norte: Chihuahua, Coahuila, Durango, San Luis Potosí, Zacatecas, Baja California, Baja California Sur, Sinaloa, Sonora, Nayarit, Nuevo León y Tamaulipas. Este: Veracruz y Tabasco.

La [Figura 11A](#) muestra la distribución de los casos de acuerdo con la agresividad basados en la escala de *Gleason*. La mayoría de los casos fueron clasificados como moderadamente diferenciados (*Gleason* = 7) y pobremente diferenciados (*Gleason* > 7) y 67% de ellos tenían 60 años de edad o más al momento del diagnóstico ([Figura 11B](#)).

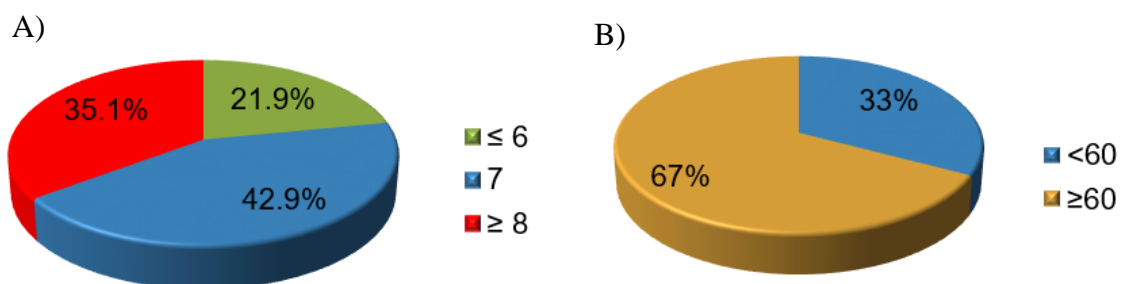


Figura 11. Distribución de los casos de acuerdo con la agresividad y a la edad al momento del diagnóstico. **A.** Clasificación de Gleason: bien diferenciado (*Gleason* ≤ 6), moderadamente diferenciado (*Gleason* = 7), pobremente diferenciado (*Gleason* ≥ 8). **B.** Edad al momento del diagnóstico: Casos de inicio temprano (<60 años de edad), casos de inicio tardío (> 60 años de edad).

Una vez descritas las características poblacionales desde el punto de vista epidemiológico, se procedió a la descripción genética poblacional y su relación con el CP a tres niveles: alélico, por linajes y filogenético.

6.2 Diversidad genética poblacional global.

Para conocer las características genéticas generales de la población, se procedió a calcular el número de alelos diferentes (k), la diversidad alélica (D) y la diversidad haplotípica (h) de cada uno de los diecisiete *loci* Y-STR estudiados. Los detalles en cuanto a la frecuencia de los diferentes alelos encontrados en los diecisiete *loci* estudiados se reportan en la [Tabla S1](#). En general, los *loci* más informativos -con base en el número de alelos presentes en cada marcador, así como, en la distribución de las frecuencias alélicas- fueron: DYS385b ($k=11$; $D=0.833$), DYS385a ($k=9$; $D=0.815$), DYS392 ($k=9$; $D=0.766$), DYS389-II ($k=8$; $D=0.747$), y DYS458 ($k=7$; $D=0.785$).

6.2.1 Diversidad alélica.

La [Tabla 10](#) muestra las frecuencias alélicas de acuerdo a la condición de caso y control. Los *loci* DYS19 (13,15) DYS385b (17), DYS389-I (12), DYS389-II (28,32), DYS391 (9), DYS392 (11, 12,14), DYS393 (14), DYS448 (19,20), DYS456 (15,19) mostraron diferencias de frecuencias de distribución entre los casos y los controles. No obstante, al realizar la corrección por Bonferroni ($p = 0.0029$), solo el alelo 14 del *loci* DYS393 (0.18 vs 0.09; $p = 0.0022$) y el 19 del *loci* DYS448 (0.11 vs 0.08; $p = 0.0004$), fueron significativamente más frecuente entre los casos que entre los controles.

6.2.2 Diversidad haplotípica global.

Con respecto a la diversidad haplotípica, de los 533 haplotipos obtenidos 457 presentaron haplotipos únicos (*singletons*) sugiriendo una alta diversidad ($h = 0.860 \pm 0.027$). El resto de los haplotipos ($n=75$), presentaron frecuencias compartidas y el número máximo de individuos que compartieron un mismo haplotipo fue de 3. Esta diversidad fue soportada por los valores de *median pairwise difference* (MPD); valor que nos indica las diferencias encontradas en cada uno de los haplotipos comparándolo en pares, así como por el parámetro de la varianza genética intra-poblacional (Vp); valor que sirvió para determinar que tanta variación existía entre un haplotipo y otro dentro del mismo linaje

6.2.3 Diversidad de linajes.

A partir de los haplotipos de cada individuo (Datos no mostrados), se determinó el linaje ancestral. Globalmente, los resultados mostraron la presencia de dieciocho haplogrupos, de los cuales los más frecuentes fueron: Q (42.21%), R1b (30.01%) y E1b1b (7.50%). La [Tabla Suplementaria S2](#) muestra con mayor detalle las frecuencias de todos los linajes presentes en los grupos de estudio. Al interior de cada linaje ancestral o haplogrupo también se observó un gran diversidad haplotípica. La mayor diversidad se observó en el haplogrupo E1b1b seguido por Q; mientras que el haplogrupo R1b las diferencias entre los haplotipos fue menos notable. ([Tabla 11](#)).

Tabla 10. Frecuencias alélicas de los 17 marcadores Y-STR encontradas entre los casos (Ca) incidentes de CP y controles (Co) poblacionales de varones mestizos mexicanos.

| Alelo | DYS19 | | DYS385-a | | DYS385-b | | DYS389-I | | DYS389-II | | DYS390 | | DYS391 | | DYS392 | | DYS393 | | DYS437 | | DYS438 | | DYS439 | | DYS448 | | DYS458 | | DYS456 | | DYS635 | | Y GATA H4 | | | | |
|-------|-------|-------|----------|-------|----------|-------|----------|-------|-----------|-------|--------|-------|--------|-------|--------|-------|-------------|-------------|--------|-------|--------|-------|--------|-------|--------|-------------|-------------|-------|--------|-------|--------|-------|-----------|-------|--|--|--|
| | Ca | Co | Ca | Co | Ca | Co | Ca | Co | Ca | Co | Ca | Co | Ca | Co | Ca | Co | Ca | Co | Ca | Co | Ca | Co | Ca | Co | Ca | Co | Ca | Co | Ca | Co | Ca | Co | | | | | |
| 8 | | | | | | | | | | | | | | | | | | | | | | 0.020 | 0.000 | | | | | | | | | | | | | | |
| 9 | | | | | | | | | | | | | 0.120 | 0.071 | | | | | | | | | 0.093 | 0.065 | | | | | | | | | 0.007 | 0.000 | | | |
| 10 | | | 0.000 | 0.013 | | | | | | | | | 0.567 | 0.668 | 0.013 | 0.011 | 0.000 | 0.005 | | | | 0.193 | 0.185 | 0.080 | 0.050 | | | | | | | | 0.020 | 0.021 | | | |
| 11 | | | 0.260 | 0.264 | 0.007 | 0.010 | 0.000 | 0.005 | | | | | 0.300 | 0.248 | 0.253 | 0.195 | 0.013 | 0.008 | | | | 0.380 | 0.407 | 0.293 | 0.305 | | | | | | | | 0.399 | 0.353 | | | |
| 12 | 0.013 | 0.005 | 0.080 | 0.071 | 0.033 | 0.018 | 0.260 | 0.175 | | | | | 0.013 | 0.013 | 0.067 | 0.018 | 0.147 | 0.138 | | | | 0.293 | 0.321 | 0.447 | 0.457 | | | | | | | | 0.527 | 0.550 | | | |
| 13 | 0.380 | 0.465 | 0.147 | 0.162 | 0.047 | 0.057 | 0.560 | 0.601 | | | | | | | 0.360 | 0.350 | 0.647 | 0.752 | 0.027 | 0.010 | | 0.020 | 0.021 | 0.160 | 0.172 | | | 0.007 | 0.013 | 0.047 | 0.042 | | 0.047 | 0.076 | | | |
| 14 | 0.380 | 0.371 | 0.253 | 0.227 | 0.240 | 0.264 | 0.173 | 0.217 | | | | | | | 0.160 | 0.253 | 0.18 | 0.09 | 0.640 | 0.655 | | | | 0.020 | 0.010 | | | 0.053 | 0.055 | 0.127 | 0.196 | | | | | | |
| 15 | 0.167 | 0.107 | 0.140 | 0.154 | 0.147 | 0.112 | 0.007 | 0.003 | | | | | | | 0.073 | 0.066 | 0.013 | 0.010 | 0.300 | 0.290 | | | 0.000 | 0.005 | | | 0.513 | 0.554 | 0.273 | 0.270 | | | | | | | |
| 16 | 0.040 | 0.039 | 0.080 | 0.086 | 0.160 | 0.120 | | | | | | | | | 0.073 | 0.097 | | | 0.033 | 0.044 | | | | | | | 0.320 | 0.300 | 0.280 | 0.277 | | | | | | | |
| 17 | 0.020 | 0.013 | 0.040 | 0.021 | 0.167 | 0.238 | | | | | | | | | 0.000 | 0.003 | | | | | | | | | | 0.000 | 0.003 | 0.093 | 0.065 | 0.180 | 0.165 | | | | | | |
| 18 | | | 0.000 | 0.003 | 0.147 | 0.123 | | | | | | | | | 0.000 | 0.008 | | | | | | | | | | 0.000 | 0.010 | 0.013 | 0.013 | 0.080 | 0.045 | | | | | | |
| 19 | | | | | 0.040 | 0.037 | | | | | | | | | | | | | | | | | | | | 0.11 | 0.08 | | | 0.013 | 0.005 | | | | | | |
| 20 | | | | | 0.007 | 0.016 | | | | | | | | | | | | | | | | | | | | 0.275 | 0.432 | | | | | 0.047 | 0.024 | | | | |
| 21 | | | | | 0.007 | 0.005 | | | | | 0.027 | 0.016 | | | | | | | | | | | | | | 0.463 | 0.353 | | | | | 0.173 | 0.141 | | | | |
| 22 | | | | | | | | | | | 0.060 | 0.060 | | | | | | | | | | | | | | 0.128 | 0.115 | | | | | 0.380 | 0.465 | | | | |
| 23 | | | | | | | | | | | 0.253 | 0.245 | | | | | | | | | | | | | | 0.013 | 0.008 | | | | | 0.307 | 0.298 | | | | |
| 24 | | | | | | | | | | | 0.540 | 0.551 | | | | | | | | | | | | | | 0.007 | 0.000 | | | | | 0.093 | 0.063 | | | | |
| 25 | | | | | | | | | | | 0.093 | 0.104 | | | | | | | | | | | | | | | | | | | 0.000 | 0.010 | | | | | |
| 26 | | | | | | | | | | 0.007 | 0.010 | 0.027 | 0.013 | | | | | | | | | | | | | | | | | | | | | | | | |
| 27 | | | | | | | | | | 0.013 | 0.018 | 0.000 | 0.008 | | | | | | | | | | | | | | | | | | | | | | | | |
| 28 | | | | | | | | | | 0.213 | 0.117 | 0.000 | 0.003 | | | | | | | | | | | | | | | | | | | | | | | | |
| 29 | | | | | | | | | | 0.327 | 0.319 | | | | | | | | | | | | | | | | | | | | | | | | | | |
| 30 | | | | | | | | | | 0.313 | 0.339 | | | | | | | | | | | | | | | | | | | | | | | | | | |
| 31 | | | | | | | | | | 0.113 | 0.131 | | | | | | | | | | | | | | | | | | | | | | | | | | |
| 32 | | | | | | | | | | 0.013 | 0.063 | | | | | | | | | | | | | | | | | | | | | | | | | | |
| 34 | | | | | | | | | | 0.000 | 0.003 | | | | | | | | | | | | | | | | | | | | | | | | | | |
| k= | 6 | | 9 | | 11 | | 5 | | 8 | | 8 | | 4 | | 9 | | 6 | | 4 | | 6 | | 6 | | 8 | | 6 | | 7 | | 6 | | 5 | | | | |

Los números en negritas representan los alelos de los marcadores Y-STR que fueron significativamente diferentes entre casos y controles ($p < 0.01$). , k = Número de alelos por *loci*

Tabla 11. Parámetros globales de diversidad haplotípica encontrada en cada haplogrupo

| Haplogrupo | N | <i>h</i> | Diversidad Haplotípica | MPD | <i>V_p</i> |
|----------------|-----|----------|------------------------|--------------------|----------------------|
| E1b1a | 8 | 6 | 0.96 ± 0.08 | 6.50 ± 3.44 | 0.67 |
| E1b1b | 40 | 32 | 0.99 ± 0.01 | 6.77 ± 3.21 | 0.87 |
| G2a | 11 | 11 | 1.00 ± 0.04 | 8.60 ± 4.31 | 0.69 |
| H | 3 | 1 | 1.00 ± 0.28 | 6.00 ± 3.93 | 0.44 |
| I1 | 6 | 4 | 0.93 ± 0.12 | 4.07 ± 2.36 | 0.22 |
| I2a1 | 9 | 9 | 1.00 ± 0.05 | 7.67 ± 3.95 | 0.38 |
| I2b (xI2b1) | 1 | 1 | 1.00 ± 0.00 | 0.00 ± 0.00 | 0.00 |
| I2b1 | 7 | 7 | 1.00 ± 0.08 | 6.52 ± 3.51 | 0.68 |
| J1 | 15 | 13 | 0.99 ± 0.03 | 5.27 ± 2.70 | 0.46 |
| J2a1 x J2a1-bh | 6 | 6 | 1.00 ± 0.09 | 6.80 ± 3.73 | 0.55 |
| J2a1b | 6 | 6 | 1.00 ± 0.09 | 5.73 ± 3.20 | 0.50 |
| J2a1h | 9 | 5 | 0.94 ± 0.07 | 5.08 ± 2.72 | 0.35 |
| J2b | 3 | 3 | 1.00 ± 0.27 | 6.33 ± 4.13 | 0.40 |
| L | 7 | 7 | 1.00 ± 0.08 | 10.24 ± 5.33 | 0.93 |
| Q | 225 | 191 | 0.99 ± 0.00 | 7.66 ± 3.58 | 0.78 |
| R1a | 6 | 6 | 1.00 ± 0.13 | 6.20 ± 3.55 | 0.38 |
| R1b | 160 | 140 | 0.99 ± 0.00 | 6.63 ± 3.15 | 0.42 |
| T | 11 | 9 | 0.98 ± 0.05 | 8.14 ± 4.09 | 0.11 |
| Total | 533 | 457 | - | - | - |

Los datos en negritas representan las haplogrupos más diversos haplotípicamente, confirmados por *MDP* y *V_p*. N= cantidad de individuos que comparten el mismo haplogrupo, *h*= haplotipos diferentes encontrados en la población, *MPD*= *median pairwise difference*, *V_p*= Varianza intra-poblacional.

Con la finalidad de simplificar la variabilidad dentro de los linajes y tomando en cuenta la convergencia con el ancestro común, los haplogrupos R1a y R1b fueron englobados en el linaje R, el haplogrupo E incluyó a los linajes E1b1a y E1b1b. Los linajes G2a, I1, I2a1, I2b, I2b1, J1a, J2a1b, J2a1h, J2b al estar filogenéticamente relacionados con el macrohaplogrupo F, fueron agrupados para su estudio presentando una frecuencia conjunta del 13.69% (Tabla 12).

Tabla 12. Distribución de haplogrupos o linajes ancestrales de acuerdo con casos y controles.

| Haplogrupo | N | Frecuencia % |
|------------|------------|--------------|
| E | 48 | 9% |
| GIJ | 73 | 14.3% |
| H | 3 | 0.6% |
| L | 7 | 1.3% |
| Q | 225 | 42.2% |
| R | 166 | 31.2% |
| T | 11 | 2.1% |

Los datos en negritas corresponden a aquellos haplogrupos que mostraron una mayor frecuencia.

A las frecuencias de los linajes ancestrales entre casos y controles (Tabla 13), se encontró que los haplogrupos E (0.19 vs 0.11; $p < 0.01$) y GIJ (0.19 vs 0.11; $p < 0.01$) fueron significativamente más frecuente en los casos que en los controles; R1a (0.03 vs 0.01; $p = 0.07$) esta diferencia fue marginal ($p = 0.07$). En contraste, el haplogrupo Q, fue menos frecuente en los casos (0.36 vs. 0.45; $p = 0.003$) que en los controles (Tabla 13).

Tabla 13. Distribución de haplogrupos o linajes ancestrales de acuerdo con casos y controles.

| Haplogrupo | Casos | Controles | <i>p</i> |
|------------|-------------|-------------|--------------|
| E | 0.19 | 0.11 | 0.009 |
| GIJ | 0.19 | 0.11 | 0.009 |
| H | 0 | 0.01 | 0.26 |
| L | 0 | 0.01 | 0.26 |
| Q | 0.36 | 0.45 | 0.003 |
| R1a | 0.03 | 0.01 | 0.07 |
| R1b | 0.29 | 0.3 | 0.48 |
| T | 0.01 | 0.03 | 0.24 |

Los datos en negritas corresponden a aquellos haplogrupos que mostraron diferencias estadísticamente significativas con un valor de $p \leq 0.01$.

6.2.4 Análisis filogenético

La construcción de las redes filogenéticas se realizó para todos los linajes. Los linajes Q y R mostraron la mayor complejidad filogenética mostrando una arquitectura estrellada (Material Suplementario Figuras S1 y S2). Aquellos linajes en los que se encontraron diferencias significativas entre los casos y los controles se presentan a continuación.

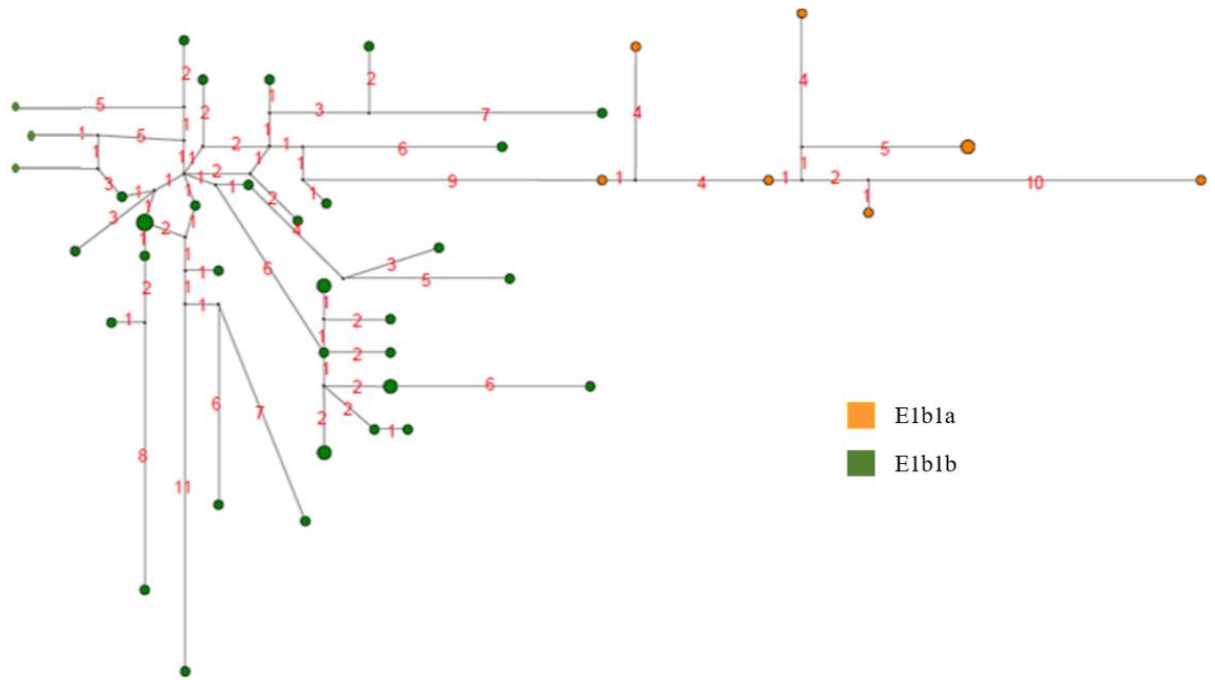
- Linaje E

Presenta una arquitectura estrellada, con un elevado número de *singletons*, aunque algunos haplotipos son compartidos. El sublinaje E1b1a aparece en un rearrreglo separado de los haplotipos E1b1b (Figura 12A). Al comparar entre casos y controles (Figura 12B), se observa que entre los casos destaca el sublinaje E1b1b; no obstante los casos no se distribuyen de forma separada de los controles, estando insertados en toda la red.

- Linaje GIJ

Este macrohaplogrupo muestra una arquitectura estrellada, con una elevada diversidad representada por el elevado número de *singletons* (Figura 13). En general, los diferentes linajes se distribuyen de forma diferencial pero algunos de ellos se distribuyen indistintamente lo que podría relacionarse con los sublinajes dentro de cada linaje (haplogrupo I).

A)



B)

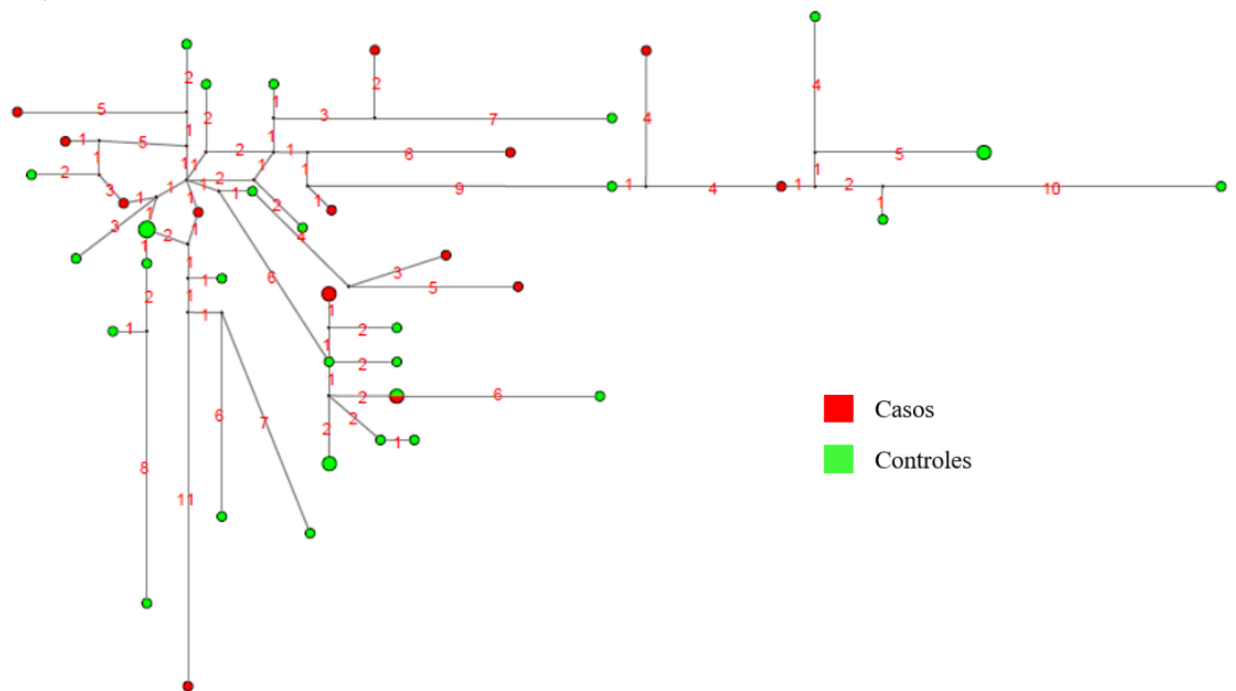


Figura 12. Red filogenética del haplogrupo E. A. Distribución de los linajes E1b1a y E1b1b dentro de la población mexicana, haciendo uso de 15 Y-STRs. B. Red filogenética del linaje E de acuerdo a la distribución de casos y controles. La longitud de las líneas representa la distancia entre un haplotipo y otro. Los números en color rojo representan las mutaciones entre los haplotipos.

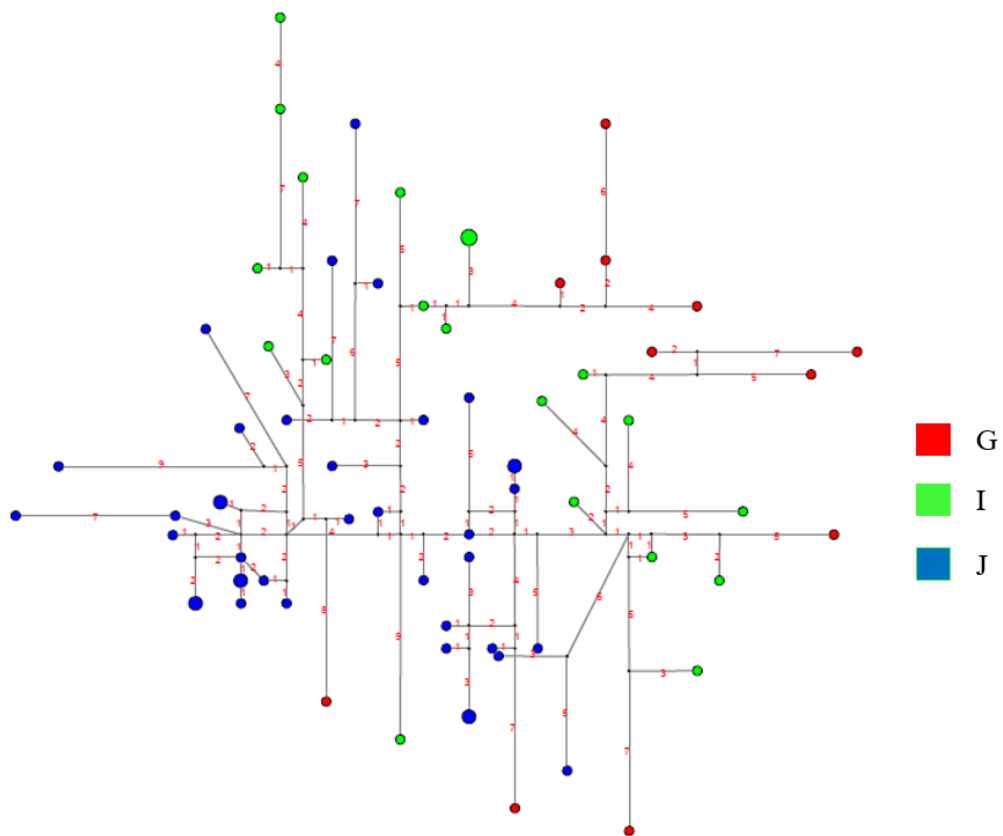


Figura 13. Red filogenética del macrohaplogrupo F (haplogrupos G,I, y J) en la población de estudio utilizando 15 Y-STRs. La distancia entre cada haplotipo está representada por la longitud de la línea así como el número de mutaciones (Números rojos).

Al realizar la construcción de la red filogenética del haplogrupo I por separado, se observa que cada sublinaje tiene un comportamiento particular. El sublinaje I1 aparece en el extremo inferior derecho de la red, el sublinaje I2a1 en la parte central separado de I2b1 que se localiza en la región izquierda de la red (Figura 14A). Al comparar los portadores de este linaje entre casos y controles se observa que la mayor frecuencia de los casos portan el sublinaje I2b1(Figura 14B).

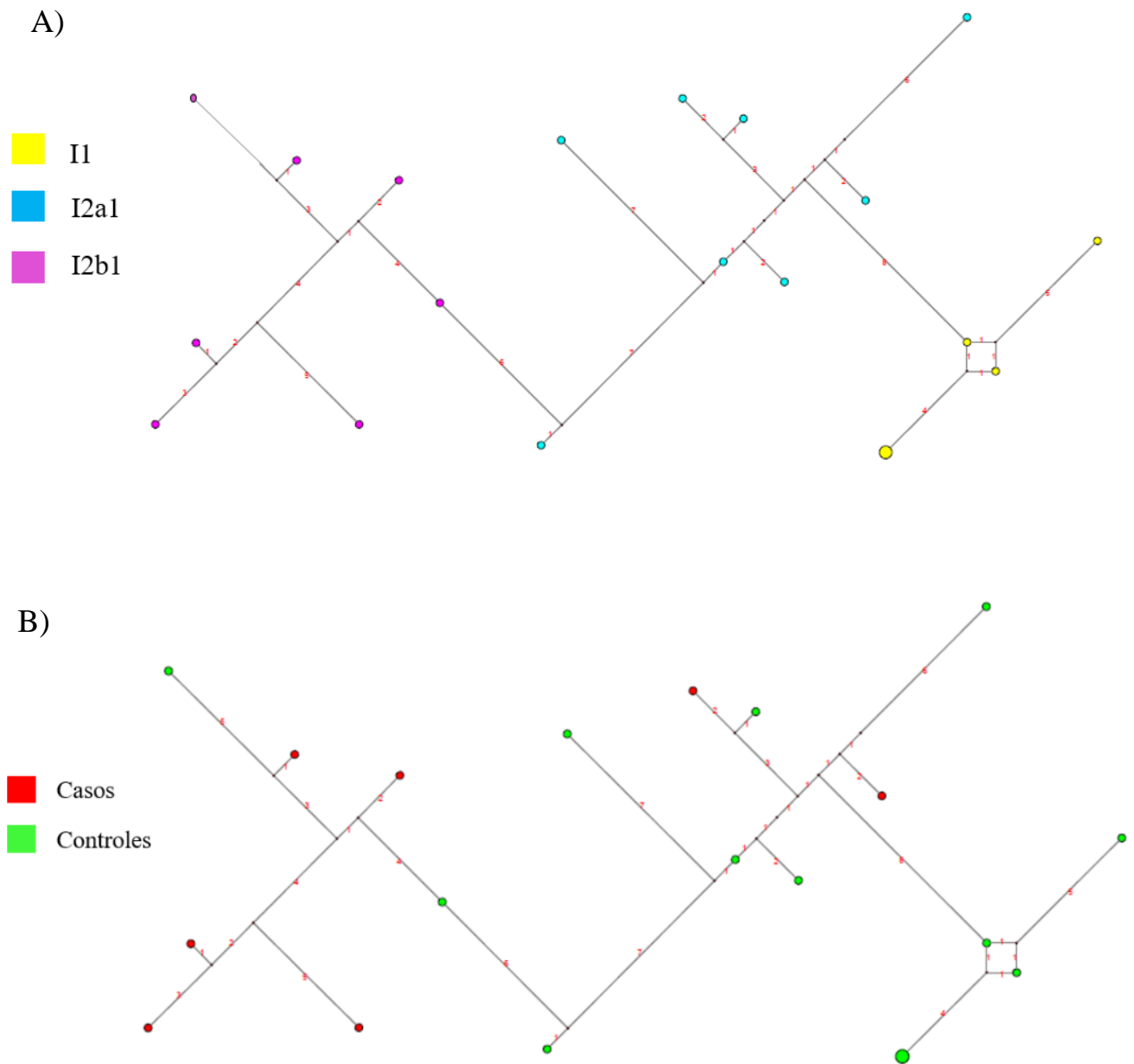


Figura 14. Red filogenética del haplogrupo I.A. Se muestra la red filogenética de acuerdo a los sub-haplogrupos encontrados en el linaje I. **B.** Red filogenética del haplogrupo I de acuerdo a la distribución de casos y controles. El largo de las líneas representa la distancia (medida en número de mutaciones-números rojos) entre los haplotipos.

La red del linaje J, muestra una separación de sublinajes (Figura 15A), en donde la mayoría son *singletons*. El comportamiento de los casos y los controles dentro del haplogrupo J, fue relativamente diferencial (Figura 15B).

Particularmente, la mayor cantidad de casos se concentró en el sublinaje J2a1h, el resto presentó una distribución ubicua, pero interesantemente en los casos el número de mutaciones fue mayor que en los controles.

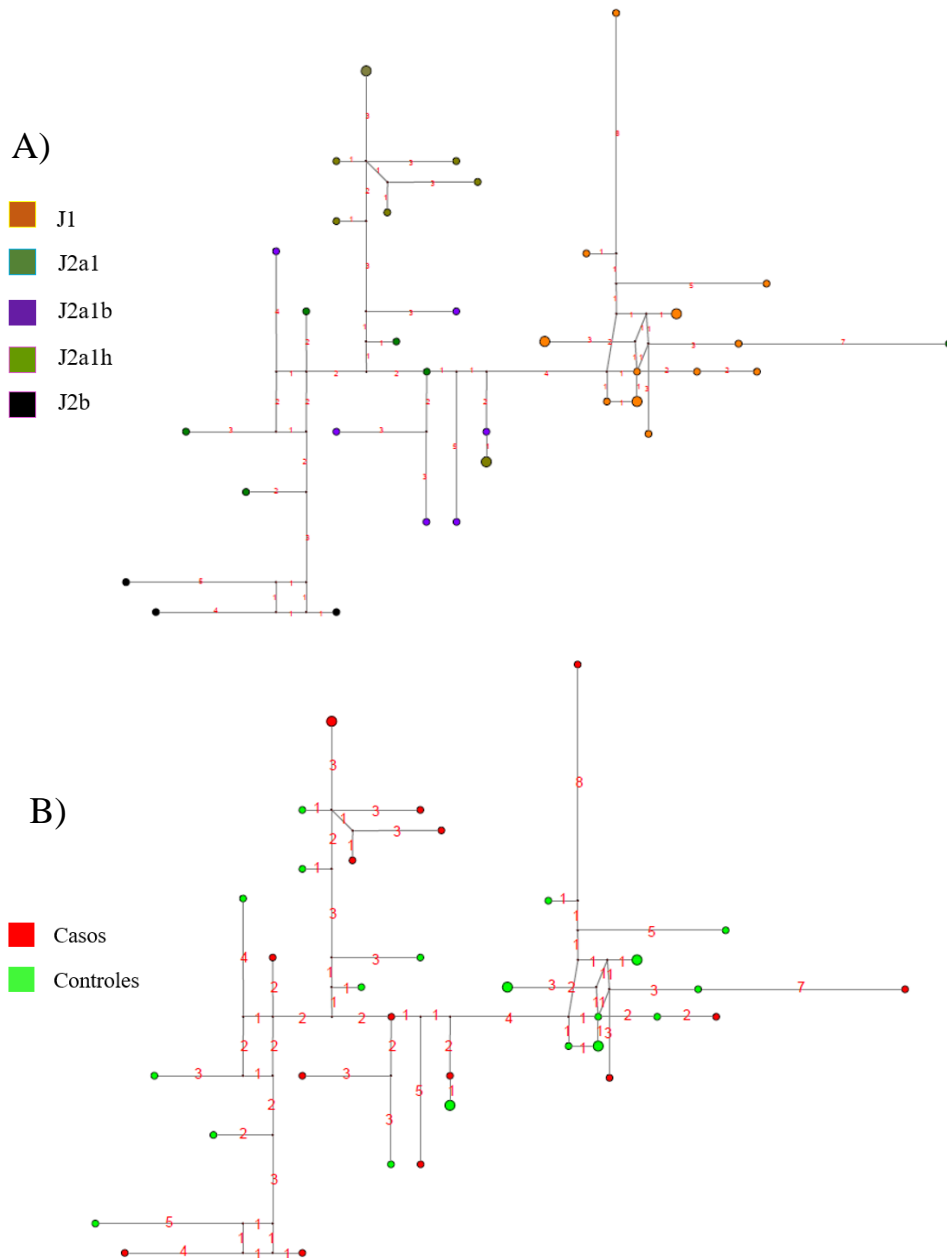


Figura 15. Red filogenética del haplogrupo J.A. Se muestra la red filogenética de acuerdo a los linajes derivados de J. **B.** Distribución de la red filogenética del haplogrupo J de acuerdo a casos y controles. El largo de las líneas representa la distancia (medida en número de mutaciones-números rojos) entre los haplotipos.

6.3 Análisis Molecular de Varianza (AMOVA).

El AMOVA mostró una diferencia significativa ($p \leq 0.0001$) entre las poblaciones, siendo las diferencias intrapoblacionales las que mostraron el mayor porcentaje de variación (98.46%).

6.4. Análisis de escala multidimensional (MDS).

Con el fin de evaluar las distribuciones espaciales entre los casos y controles, se construyó un MDS en aquellos haplogrupos en los que se encontraron diferencias destacables. Se incluyeron los datos de una población del centro de México, publicada previamente por nuestro grupo de investigación (Santana et al., 2014), para conocer la distribución de los haplotipos en una población general, por lo que fue usada como un control genómico (CG).

La [Figura 16](#) muestra el MDS usando los dieciocho haplogrupos en conjunto. La primera dimensión separa a los casos de los controles y el CG, mientras que la segunda dimensión agrupa a los controles con el CG, esta agrupación no es muy íntima probablemente representando la diversidad de haplotipos de ambas poblaciones. El valor de *stress* (0.00031) sugiere una configuración con un alto grado de verosimilitud, permitiendo una detallada inspección de los datos. Los MDSs de aquellos linajes en los que se encontraron diferencias significativas entre los casos y los controles (E, G, I, y J) mostraron un comportamiento similar al descrito en la [Figura 16](#) (Datos no mostrados).

Por lo que respecta al linaje más prominente (haplogrupo Q), el MDS mostró una íntima relación entre los casos y los controles, mostrando una sutil diferencia con el CG ([Figura Suplementaria S3](#)). Los tres grupos (casos, controles y CG) mostraron una conexión ancestral con las poblaciones Nativas Mexicanas (Otomíes, Mayas y Tarahumaras; Regueiro et al., 2013), así como con las poblaciones del Norte de América (Alaska y Canadá; Dulik et al., 2012; Schurr et al., 2012).

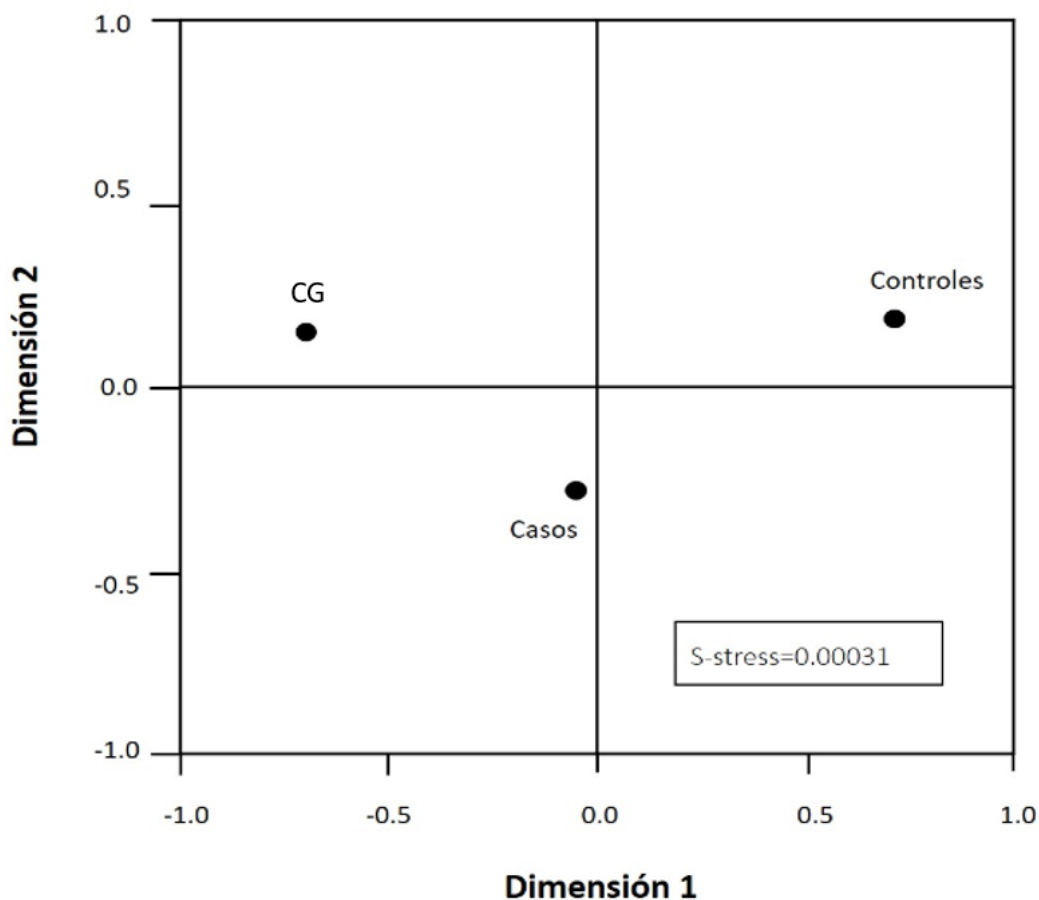


Figura 16. Análisis de escalamiento multidimensional entre los grupos de estudio (Casos y controles) y un CG (Control Genómico) empleando 15 Y-STRs. La gráfica muestra el análisis usando los dieciocho haplogrupos reportados en el presente estudio.

6.4.1 Comparación con otras poblaciones.

Con el fin de relacionar a los portadores de los linajes que mostraron diferencias entre casos y controles con poblaciones portadoras de estos linajes, se realizaron diferentes MDS (Figuras 17-19).

De manera general, los linajes E, G, I y J mostraron separaciones entre la posición de los casos *versus* los controles y el CG. Interesantemente, los casos estuvieron más relacionados con las poblaciones Medio Orientales, mientras que los controles y el CG se aproximaron a las poblaciones Europeas. El haplogrupo J

mostró una cercanía entre los casos y la población de Italia, mientras que los controles se asemejaron a las poblaciones del norte de Italia.

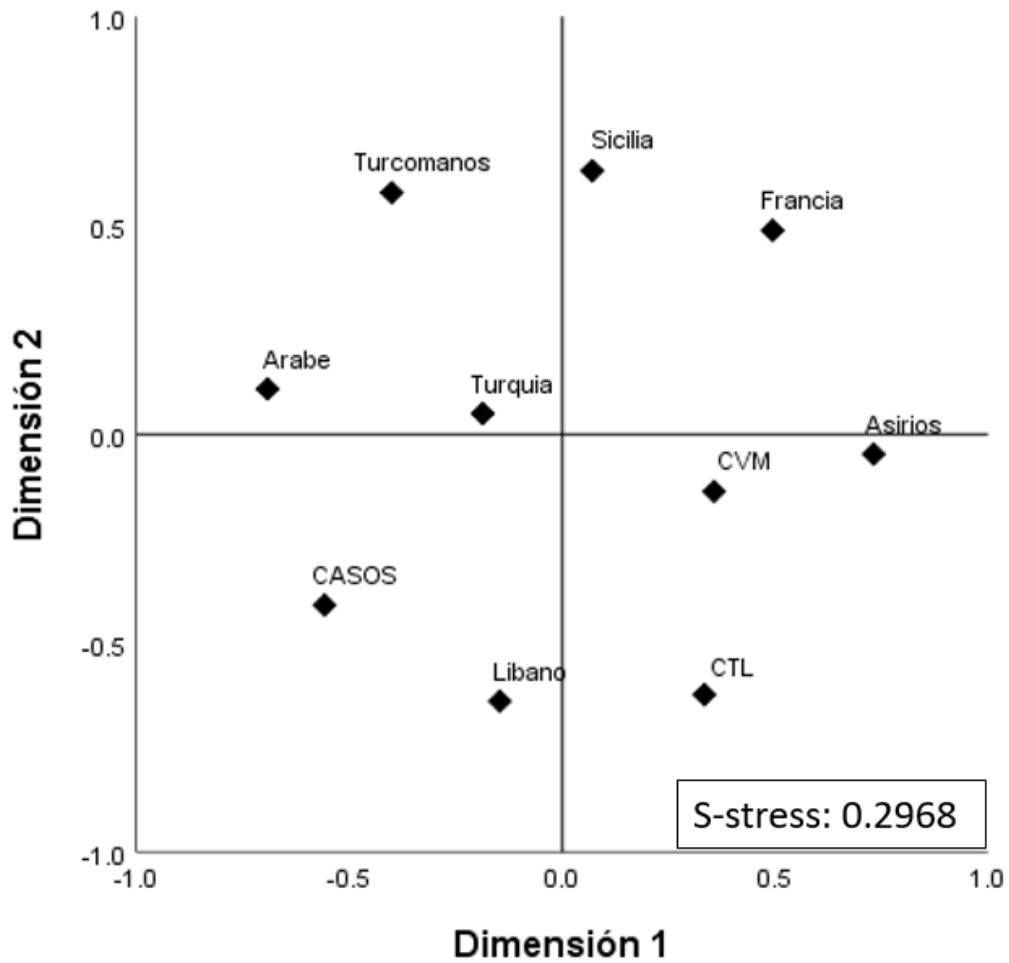


Figura 17. Análisis de escalamiento multidimensional del linaje ancestral E entre los grupos de estudio (Casos y controles) y un CG (Control Genómico), empleando 15 Y-STRs. La figura muestra el análisis usando únicamente los resultados de los haplogrupos E1b1a y E1b1a. CTL: Control; CVM: Valle Central de México (Santana et al., 2014); Turcomanos, Asirios (Dogan et al., 2017); Árabes (); Libano (Pierre et al., 2008); Turquía (Dönbak et al., 2008); Sicilia, Francia (Martínez et al., 2012).

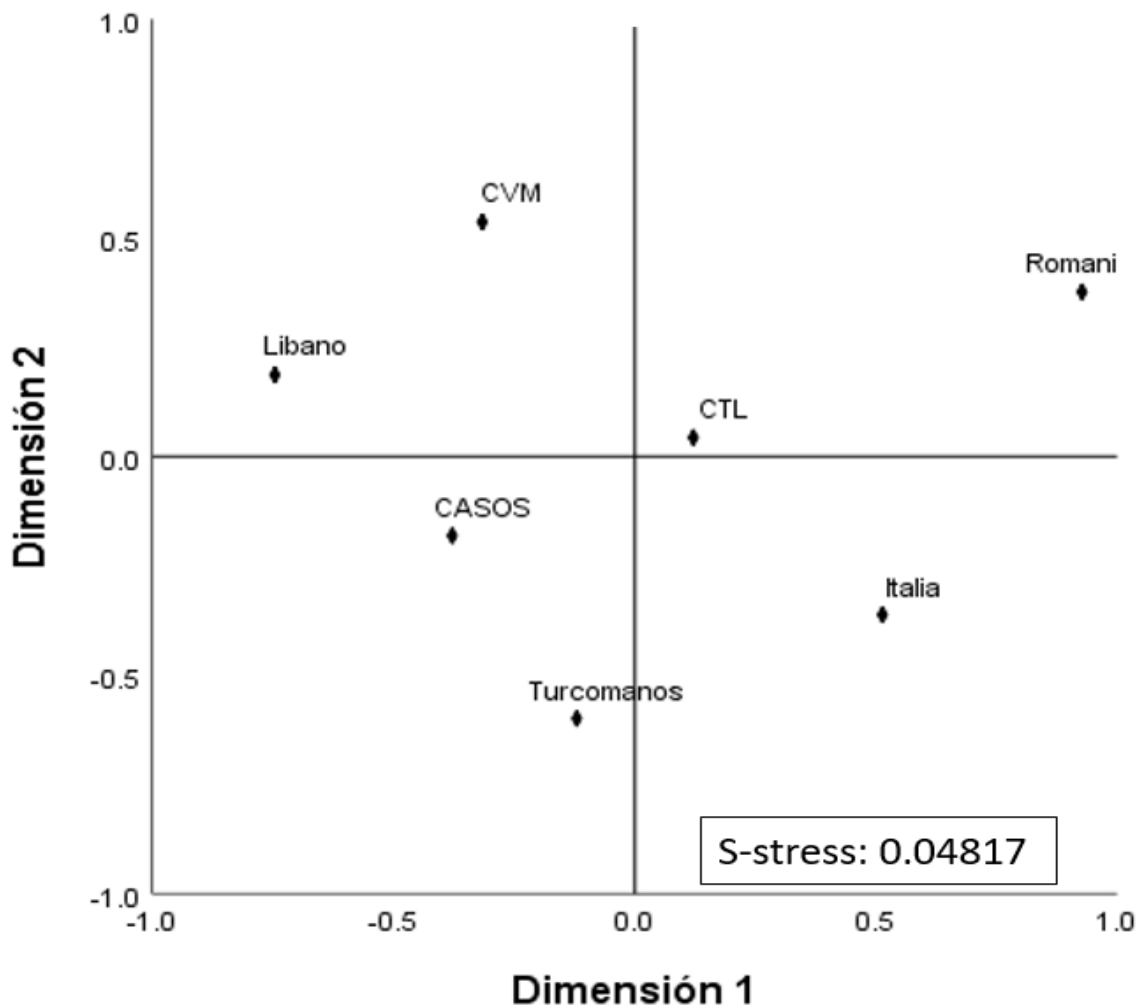


Figura 18. Análisis de escalamiento multidimensional del linaje ancestral I entre los grupos de estudio (Casos y controles) y un CG (Control Genómico), empleando 15 Y-STRs. Este análisis únicamente empleó las muestras que tuvieran haplogrupos derivados del linaje I (I1, I2a1, I2a1b, I2a1h, J2b). CTL: Control; CVM: Valle Central de México (Santana et al., 2014); Turcomanos (Dogan et al., 2017); Libano (Pierre et al., 2008); Romani (Bogoña- Cruz et al., 2016); Italia (Martínez et al., 2012).

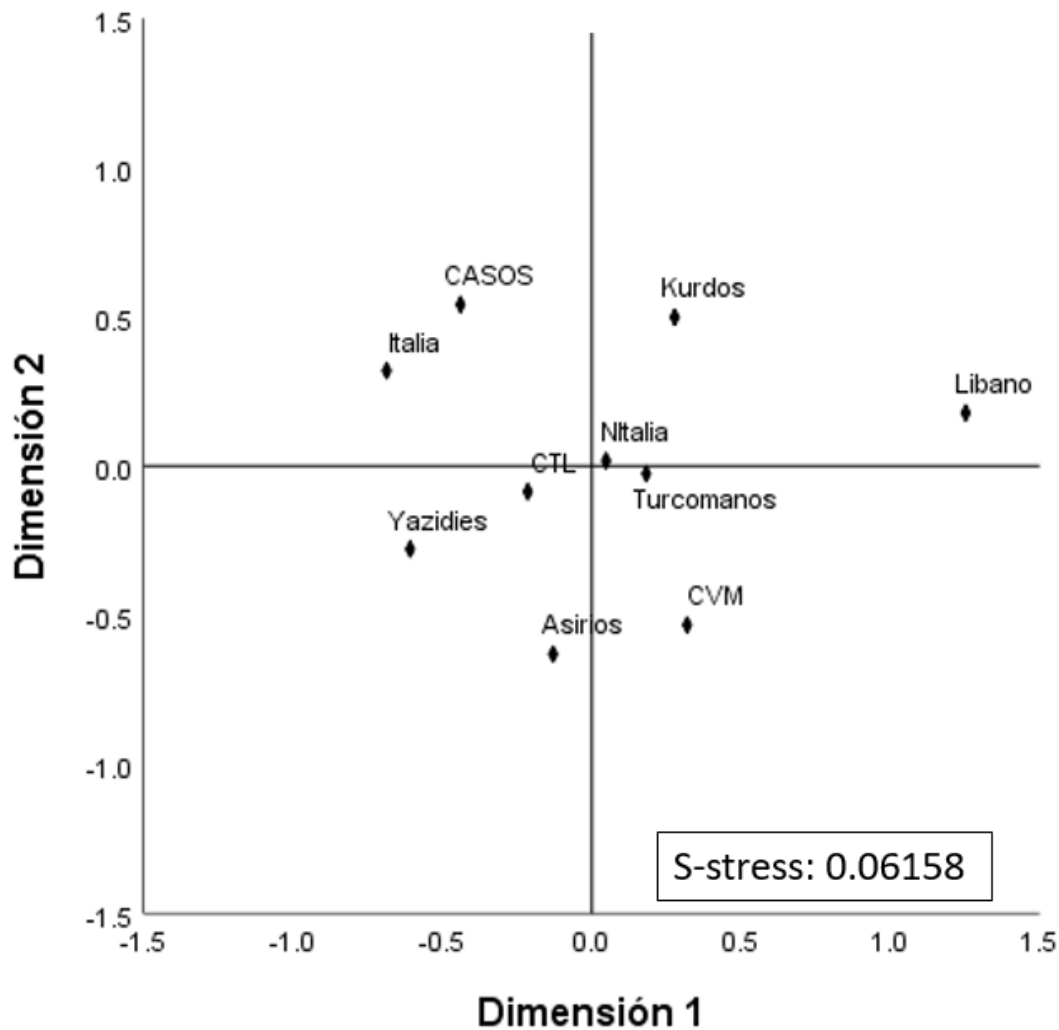


Figura 19. Análisis de escalamiento multidimensional del linaje J entre los grupos de estudio (Casos y controles) y un CG (Control Genómico), empleando 15 Y-STRs. Este análisis únicamente empleó las muestras que tuvieran haplogrupos derivados del linaje J (J1, J2a1, J2a1b, J2a1h, J2b). CTL: Control; CVM: Valle Central de México (Santana et al., 2014); Turcomanos, Yazidies, Kurdos, Asirios (Dogan et al., 2017); Libano (Pierre et al., 2008); Italia, Norte de Italia (Martínez et al., 2012).

6.5 Asociación entre el linaje ancestral y el cáncer de próstata.

Para evaluar las asociaciones entre los linajes ancestrales Afroscendente (E1b1a, E1b1b), Europeo (R1a, R1b) con el riesgo de CP, consideramos el linaje ancestral Nativo Americano (Q) como referencia, debido a su alta frecuencia en nuestra población de estudio. En comparación con los hombres con un linaje ancestral nativo americano, la frecuencia de CP entre los portadores de los linajes Europeo (RM=1.31; IC: 0.83-2.06) y Afroscendientes (RM=1.63; IC: 0.83-3,21) fue mayor pero no estadísticamente significativa (Tabla 14). En los haplogrupos G, I y J no se realizó el análisis dado el limitado número de muestra.

Tabla 14. Asociación entre el clan ancestral y el riesgo de cáncer de próstata.

| Clan ancestral | Casos n= 322 (%) | Controles n= 628 (%) | RM ^a | IC 95% | Valor de <i>p</i> |
|------------------|---------------------|-------------------------|-----------------|-----------|----------------------|
| Nativo Americano | 160 (49.7) | 314 (50.0) | 1.0 | -- | |
| Europeo | 123 (38.2) | 236 (37.6) | 1.31 | 0.83-2.06 | 0.25 |
| Afroscendente | 39 (12.1) | 78 (12.4) | 1.63 | 0.83-3.21 | 0.16 |

Nativo Americano: incluye al haplogrupo Q; Europeo: incluye a los haplogrupos R1a y R1b; Afroscendente: incluye a los haplogrupos E1b1a y E1b1b.

a: Razón de momios (RM) modelo ajustado por edad al momento de la entrevista y antecedente familiar de CP.

6.6 Asociación entre linaje ancestral y la agresividad del cáncer de próstata.

Con la finalidad de evaluar una posible contribución entre la carga ancestral y la agresividad del CP, se estratificaron a los casos de CP con respecto al grado de agresividad de acuerdo con la escala de *Gleason* al momento del diagnóstico. En comparación con los sujetos que tenían un linaje ancestral nativo americano, los sujetos con linaje afroscendientes tuvieron más posibilidades de tener un Gleason menor o igual a 6 o mayor o igual a 8; no obstante, esta última asociación no alcanzó significancia estadística. Entre los que contaban con un linaje europeo se observó un incremento no significativo de todos los tipos de Gleason. (Tabla 15).

Tabla 15.

Tabla 15. Asociación entre el clan ancestral y la agresividad del cáncer de próstata evaluada por la escala de *Gleason* al momento del diagnóstico.

| Clan Ancestral | Controles (n=321) | <i>Gleason</i> ≤ 6 | | | <i>Gleason</i> sólo 7 | | | <i>Gleason</i> ≥ 8 | | |
|------------------|----------------------|--------------------|-----------------|-------------------|-----------------------|-----------------|-----------|--------------------|-----------------|------------------|
| | | Casos (n=82) | RM ^a | IC 95% | Casos (n=49) | RM ^a | IC 95% | Casos (n=116) | RM ^a | IC 95% |
| Nativo Americano | 171 | 9 | 1 | - | 25 | 1 | - | 17 | 1 | - |
| Europeo | 118 | 9 | 1.44 | 0.55-3.75 | 22 | 1.22 | 0.65-2.29 | 15 | 1.42 | 0.66-3.07 |
| Afroascendente | 32 | 7 | 4.02 | 1.38-11.65 | 2 | 0.47 | 0.10-2.10 | 7 | 2.67 | 0.98-7.29 |

Los valores en negritas corresponden a aquellos que mostraron diferencias estadísticamente significativas con un valor de $p < 0.05$.

Nativo Americano: incluye el haplogrupo Q; Europeo: incluye los haplogrupos R1a y R1b; Afroascendente: incluye los haplogrupos E1b1a y E1b1b

a: Razón de momios (RM) modelo ajustado por edad al momento de la entrevista y antecedente familiar de CP.

6.7 Asociación entre linaje ancestral y la edad de presentación del cáncer de próstata

Para evaluar si la carga ancestral se asociaba con la edad al momento del diagnóstico de CP, se llevó a cabo una estratificación en CP de inicio temprano (CP-IT < 60 años) y CP de inicio tardío (CP-ITO ≥ 60 años). La frecuencia de CP-ITO fue dos veces mayor entre el clan ancestral Afroascendente que entre el clan ancestral Nativo Americano (RM= 2.15; IC_{95%}= 1.02-4.54 $p = 0.04$) (Tabla 16). No obstante, con la finalidad de determinar si el resultado era consecuencia de que el rango de edad entre los controles era más amplio que entre los casos (45-90 años), realizamos una segunda comparación incluyendo en el modelo de inicio temprano sólo a aquellos controles que tenían máximo 65 años y el modelo de inicio tardío aquellos controles con una edad mínima de 55 años. Los resultados obtenidos se mantuvieron sin cambio con respecto al análisis anterior (Tabla 17).

Tabla 16. Asociación entre clan ancestral y edad al momento del diagnóstico.

| Clan ancestral | Inicio temprano <60 | | | | Inicio Tardío ≥60 | | | |
|------------------|---------------------|-------------------|-----------------|-----------|-------------------|-------------------|-----------------|------------------|
| | Casos (n=35) | Controles (n=119) | RM ^a | IC 95% | Casos (n=83) | Controles (n=287) | RM ^a | IC 95% |
| Nativo Americano | 19 | 41 | 1 | - | 35 | 149 | 1 | - |
| Europeo | 14 | 23 | 1.15 | 0.51-2.61 | 34 | 108 | 1.3 | 0.76-2.22 |
| Afroscendente | 2 | 4 | 0.67 | 0.12-3.63 | 14 | 30 | 2.07 | 0.94-4.37 |

Los valores en negritas corresponden a aquellos que mostraron diferencias estadísticamente significativas con un valor de $p < 0.04$.

Nativo Americano incluye el haplogrupo Q; Europeo incluye los haplogrupos R1a y R1b; Afroscendente incluye los haplogrupos E1b1a y E1b1b

a: Razón de momios (RM) modelo ajustado por edad al momento de la entrevista y antecedente familiar de CP

Tabla 17. Asociación entre clan ancestral y edad al momento del diagnóstico restringido por edad.

| Clan ancestral | Inicio temprano <60 | | | | Inicio Tardío ≥60 | | |
|------------------|---------------------|--------------|-----------------|-----------|-------------------|-----------------|------------------|
| | Controles (n=321) | Casos (n=35) | RM ^a | IC 95% | Casos (n=83) | RM ^a | IC 95% |
| Nativo Americano | 171 | 19 | 1 | - | 35 | 1 | - |
| Europeo | 118 | 14 | 1.29 | 0.55-2.98 | 34 | 1.32 | 0.78-2.28 |
| Afroscendente | 32 | 2 | 0.35 | 0.05-2.37 | 14 | 2.15 | 1.02-4.54 |

Los valores en negritas corresponden a aquellos que mostraron diferencias estadísticamente significativas con un valor de $p < 0.05$.

Nativo Americano incluye el haplogrupo Q; Europeo incluye los haplogrupos R1a y R1b; Afroscendente incluye los haplogrupos E1b1a y E1b1b

a: Razón de momios (RM) modelo ajustado por edad al momento de la entrevista y antecedente familiar de cáncer.

DISCUSIÓN

En el presente trabajo, se determinó la contribución de la carga ancestral patrilineal, en específico el linaje afroascendente en el riesgo para desarrollar cáncer de próstata en población mestiza mexicana. Para ello, se obtuvieron los haplotipos de 533 hombres mestizos mexicanos no relacionados entre sí residentes de la Ciudad de México (150 casos incidentes de cáncer de próstata y 383 controles), los que fueron analizados con diecisiete marcadores hipervariables presentes en la región no recombinante del cromosoma Y.

Los haplotipos encontrados mostraron una compleja diversidad en la que más del 85% presentaron combinaciones únicas, cada una de ellas con una elevada variación alélica. Consecuentemente, la diversidad de linajes también fue notable, reflejando la multietnicidad y complejidad étnica de los linajes paternos. Esta diversidad es consecuencia de las diferentes olas migratorias que han tenido lugar en México en los últimos siglos (Luna-Vásquez et al., 2005; Beezley, 2012) las que han enriquecido constantemente el pool genético de la población, conformando una compleja arquitectura y un origen paterno multicultural, el que estuvo mayoritariamente constituido por los haplogrupos Q, R1b y E1b1b.

Contrario a los reportes previos en donde la ancestría Europea ha sido la más frecuente, en el presente trabajo la ancestría Nativa Americana -representada por el haplogrupo Q-(~42%) mostró ser la más prominente (Santana et al., 2014; Salazar-Flores et al., 2010, Martínez-Cortes et al., 2012; Luna-Vásquez et al., 2008; Martínez- Fierro et al., 2009; Martínez- Marnignac et al.,2007). Lo anterior sugiere una sobre-representación de linajes indígenas en nuestra muestra de estudio, lo que pudo haberse debido a que en el presente trabajo sólo se consideró la residencia en la Ciudad de México del participante, más no así el de sus ancestros paternos de primer grado. Consecuentemente, algunos individuos podrían haber reportado un determinado lugar de nacimiento, aunque sus orígenes ancestrales pudieron haber estado en otro estado de la República. Particularmente, la distribución de haplogrupos está sesgada por el lugar de residencia, como es el caso de otras ciudades del centro del país donde se han realizado estudios previos; la Ciudad de México tiene una migración que proviene especialmente de zonas rurales donde la presencia Nativa Americana es más prominente. (González-Sobrino et al., 2016).

Consecuentemente, otro aspecto a considerar fue el número de individuos analizados de cada región del país, lo que probablemente pudo sobre representar a este linaje. Como ya fue mencionado, el haplogrupo Q está relacionado con los pueblos Nativos Mexicanos (v.gr., Mayas, Tarahumaras y Triquis) en los que este linaje los conecta con las poblaciones parentales Asiáticas, sin que ello justifique su alta frecuencia en la población Mestiza.

Con respecto al linaje R1b, éste representa la ancestría Europea producto de la conquista del Nuevo Mundo, mostrando un claro sesgo en el patrón de mezcla genética en donde los hombres Europeos se mezclaban con las mujeres Nativas Americanas dando origen a la población mestiza (Guardado- Estrada et al., 2009; Santana et al., 2014; Beezley, 2011). Los estudios poblacionales realizados en diferentes regiones de la Península Ibérica reflejan la elevada frecuencia de este linaje, soportando su presencia en la población Mexicana a partir de las migraciones trasatlánticas durante la colonización del Nuevo Mundo (Ambrosio et al., 2012). La alta prominencia de R1b en la población mestiza, ha sido soportada por estudios previos en los que este linaje resulta ser el haplogrupo mayoritario de la población masculina (Martínez-Cortes et al., 2012; Santana et al., 2014). Sin embargo, en el presente trabajo, el linaje R1b resultó ser el segundo más frecuente, posicionándose por debajo del haplogrupo Q. En este sentido, algunos reportes señalan que la distribución de R1b puede presentar fluctuaciones dado que existe un aumento de la contribución genética europea en el norte del país y un fondo predominante de Nativos Americanos en el sur del país, pudiendo explicar su posición de frecuencia en el presente estudio dado que la muestra estudiada incluyó a diferentes zonas geográficas, contrario a lo que reportan las investigaciones previas que solo se concentran determinadas regiones (Martínez-Cortes et al., 2012; Moreno-Estrada et al., 2014).

Por lo que respecta a los haplogrupos E1b1a y E1b1b, estos presentaron una mayor frecuencia en nuestra población de estudio (8-23 %) en comparación a la reportada por otros grupos de investigación (~ 5-8%) (Santana et al., 2012; Gorodezky et al., 2001; Martínez-Cortes et al., 2012). Como se mencionó anteriormente, el muestreo que se realizó en el presente trabajo pudo tener una mayor representación de las diferentes zonas geográficas de la República, incluyendo a las regiones costeras, en las que la ancestría Africana es más evidente.

En este sentido, las frecuencias de estos haplogrupos presentan un aumento relativamente mayor en algunas regiones costeras de Jalisco, Guerrero, Oaxaca y Veracruz, sugiriendo un flujo limitado de genes africanos a las ciudades circundantes (Rangel-Villalobos et al., 2008). No obstante, es importante tener en mente que este linaje también presenta una amplia distribución en las poblaciones Medio Orientales, cuya impronta genética pudo llegar con la conquista española, en la que individuos del Medio Oriente (Árabes, Libaneses y Cripto Judíos portadores del haplogrupo E) y de algunas regiones de África (Senegal, Angola y Argelia) viajaron a la Nueva España en búsqueda de oportunidades económicas (Aguirre-Beltra'n, 1989; Gorodezky et al., 2001; Beezley, 2011). Otro aspecto a considerar es que los conquistadores Europeos también pudieron haber sido portadores de linajes Medio Orientales y no exclusivamente del linaje R1b, pues la Península Ibérica vivió bajo el yugo Moro previo a la Conquista del Nuevo Mundo (Kamier et al., 2010). Más recientemente, las migraciones Libanesas a México a finales del siglo XIX también pueden verse reflejadas a través de la frecuencia de este y otros linajes Medio Orientales como G, I y J, los que también mostraron cierta presencia en la nuestra población de estudio (Aida et al., 2016; Ambrosio et al., 2009).

Con respecto a la contribución de la carga ancestral en el desarrollo de CP, los linajes Africanos y Medio orientales representados por el haplogrupo E1b1b y por el macrohaplogrupo F (haplogrupos G, I y J), respectivamente, mostraron diferencias significativas entre casos y controles.

Particularmente, la presencia mayoritaria del haplogrupo E1b1b en casos sugiere que las características genéticas de este linaje ancestral pudieran tener una participación directa en el desarrollo del CP. Estos resultados son consistentes con los realizados por otros grupos de investigación, los que han sugerido que el macrohaplogrupo DE (en donde E1b1b está integrado), o bien algunos subhaplogrupos dentro del linaje E (E3a) presentan un mayor riesgo a desarrollar CP (Ewis et al., 2006; Lee et al., 2018; Wang et al., 2012). Estos hallazgos, muestran que el trasfondo genético podría tener un impacto importante en la incidencia y mortalidad del CP y que incluso la incidencia podría variar entre los portadores de diferentes linajes (Ewis et al., 2006). En este mismo tenor y convencidos de la importancia del estudio de la carga ancestral en el desarrollo del

CP, se han empleado marcadores informativos de ancestría (AIMs por sus siglas en inglés Ancestry Informative Markers) en hombres pertenecientes a diferentes etnicidades. Los resultados soportan que existe una mayor susceptibilidad al CP en hombres con ascendencia Afroamericana en comparación con hombres europeos y asiáticos (Boyle & Severi, 1999; Eeles et al., 2013; Grönberg, 2003; Paracchini et al., 2016). Las explicaciones para entender esta disparidad étnica en el desarrollo del CP podrían estar relacionadas con la estructura genética del cromosoma Y en cada individuo. Las diferencias entre los diversos haplogrupos radican en el número de repetidas en tándem entre los diferentes *loci*, los que de alguna manera constituyen una firma genética del haplogrupo (Cannon-Albright et al., 2014a). Cabe resaltar que los microsatélites han sido relacionados con la regulación genética, por lo que esta característica podría modificar los patrones de expresión de ciertas proteínas de importancia para el desarrollo, riesgo y agresividad del CP (Aranguren-Méndez, Román-Bravo, Isea, Villasmil, & Jordana, 2005). Diversos estudios han reportado que el gen SRY, ubicado en la región no recombinante del cromosoma Y, es un regulador negativo del receptor de andrógenos (AR), lo que sugiere que la pérdida de este gen puede aumentar la actividad de AR y, en consecuencia, estimular el crecimiento del cáncer (Kregel et al., 2013). El AR es un gen que encierra un trinucleótido [(CAG)*n*], el cual se ha relacionado con la actividad del receptor en donde una repetición más corta se asocia con mayor riesgo de CP (Chamberlain, Driver, & Miesfeld, 1994; Gómez, 2016). Interesantemente, un menor número de repetidas en tándem del (CAG)*n* es menor en las poblaciones Afro-ascendientes en comparación con las poblaciones Europeas pudiendo soportar nuestros hallazgos (Jernberg et al., 2017; Banerjee et al., 2018).

Con respecto a la asociación entre la identidad étnica y el grado de agresividad del CP, nuestros resultados mostraron que la ancestría afroascendente se asociaba significativamente con los cánceres menos agresivos (Gleason ≤ 6) y marginalmente con los cánceres altamente agresivos (Gleason ≥ 8). Estos hallazgos han sido encontrados en otros grupos de investigación en donde los casos de agresividad baja (Gleason 1-4) e intermedia (Gleason 5-6) mostraron un mayor grado de ascendencia Africana (Berndt et al., 2015; Powell et al., 2010), sugiriendo que el linaje ancestral también podría ser pronóstico de la progresión del CP. Sin embargo, debido a la baja frecuencia de sujetos en el grupo de Gleason 7 que tenían

linaje afroascendente, no es posible descartar que el linaje está relacionado con la enfermedad en general y no con la agresividad. Se requieren estudios con un tamaño de muestra más amplio para poder así, evaluar de forma más precisa la asociación entre etnicidad y la agresividad del CP.

Con respecto a otros haplogrupos (ej. macrohaplogrupo F) que por su limitada frecuencia no fueron extensamente analizados, son pocos los estudios que relacionan a los haplogrupos que lo integran (G, I y J) con el desarrollo de CP. No obstante, los análisis de redes filogenéticas así como los MDS muestran una separación entre la arquitectura genética de los casos versus los controles, sugiriendo una aportación por parte de estos linajes. En este sentido, el haplogrupo I, particularmente el sublinaje I2b1, concentró el mayor número de casos dentro de este linaje. Un dato similar fue encontrado por un grupo Sueco en donde el sublinaje I1c mostró contribuir con casi tres veces al riesgo de desarrollar CP (Lindstrom et al., 2008). Desafortunadamente, el bajo número de portadores dentro de este linaje nos impidió realizar análisis más profundos al respecto. Sin embargo, los MDS muestran un claro acercamiento entre las poblaciones Medio Orientales (Líbano) y los casos de CP de este estudio, pudiendo sugerir un efecto fundador. Interesantemente, estudios en poblaciones Medio-Orientales muestran un claro incremento en la incidencia de CP tanto en poblaciones de esta zona geográfica como en poblaciones de la región norte de África, con las que presentan una conexión genética a través de los linajes J1 y J2 (Pilleron et al., 2012; Badro et al., 2013).

De acuerdo a lo anterior esto podría estar relacionado con las diferentes migraciones ya que estudios relacionados con la salud de las poblaciones migrantes señalan que el aumento en la incidencia del CP podría estar íntimamente relacionado con los diversos factores de riesgo que existen en el país anfitrión, en donde los factores de exposición como el estilo de vida al que se está acostumbrado se modifican (Al-Abdin, Rabah, Badr, Kotb, & Aprikian, 2013; Moussawi et al., 2013). Consecuentemente, el conocimiento sobre los factores ambientales, la arquitectura genética de la población y el componente ancestral pueden ser aspectos cuyo conocimiento pueda ser empleado como estrategias de control para prevenir la incidencia del CP en un país de ancestría patrilineal multiétnica como México. En este sentido, las variaciones genéticas dentro del cromosoma Y (responsables de

originar los diferentes linajes paternos), tienen un gran impacto en la salud humana, ya que pueden albergar alelos medicamente relevantes que participen en la susceptibilidad a ciertas patologías (infertilidad, cáncer de testículo, cáncer de vejiga) o que tengan una respuesta diferencial a tratamientos farmacológicos (Acuña- Alonso et al., 2010). Algunas asociaciones de haplogrupo-enfermedad son compatibles con múltiples estudios que presentan diseños estadísticos rigurosos, sin embargo, una gran cantidad de estudios también presentan tamaños de muestra pequeños y controles inapropiados para la estratificación de la población, por lo que algunas asociaciones carecen del suficiente soporte estadístico (Via, Ziv, & Burchard, 2009). Dentro de las ventajas que presenta estudiar poblaciones multiétnicas, es que al ser genéticamente diversas se puede aumentar el poder para detectar una mayor cantidad de *loci* medicamente relevantes, confiriendo una mejor resolución del mapeo fino de las variantes causales mediante el aprovechamiento de las diferencias entre los grupos étnicos, lo que permitiría, a su vez, dar información sobre otras poblaciones que compartan ancestría en común con la población de estudio (como el resto de las poblaciones Latinas) (Asimit, Hatzikotoulas, McCarthy, Morris, & Zeggini, 2016).

Por otro lado y con respecto a las fortalezas de nuestro estudio, éste incluyó un tamaño de muestra superior al de estudios previos, además de una relación de 2.55 controles por cada caso, proporcionando un mayor poder estadístico a nuestro estudio. Sobre las limitaciones de este estudio fue la carencia de información sobre el lugar de nacimiento de sus familiares, lo que podría presentar un sesgo en las frecuencias de los linajes ancestrales encontrados las que se relacionan íntimamente con las zonas geográficas; el reducido número de muestras de individuos con agresividad intermedia (*Gleason* = 7), así como de edades menores de 60 años, limitó que pudiéramos ver un panorama completo de las asociaciones con el linaje ancestral.

Dentro de las áreas de oportunidad en este estudio, se encuentra la realización de análisis de alta definición de cada uno de los linajes lo que podría por un lado relacionar sub-linajes con el desarrollo del CP, además de apoyar el estudio de los movimientos migratorios que dieron origen a nuestra población fomentando áreas de investigación como la antropología médica y genética. El presente estudio cierra las brechas en cuanto a la relación de la influencia ancestral

en el desarrollo de enfermedades complejas (Helgason et al., 2010). No obstante, aún hace falta mucha más investigación para poder implementar el uso de marcadores uniparentales como marcadores de diagnóstico.

Hasta nuestro conocimiento, el presente trabajo es el primero en explorar el papel de la herencia patrilineal y su relación con la susceptibilidad al desarrollo del CP en la población mestiza mexicana, además de ser el mayor estudio que ha evaluado el papel de los linajes ancestrales en la etiología del CP en la población Latina a partir de marcadores tipo microsatélites.

CONCLUSIONES

Tal como esta investigación ha demostrado, la población mexicana presenta una elevada diversidad, representada por los linajes paternos Nativos Americanos (Q), Europeos (R) , Afroacendentes (E) y del Medio Oriente (G, I, J) , viéndose reflejada en los diferentes parámetros poblaciones establecidos, confirmando así la multiétnicidad de la población mexicana.

Respecto a la relación entre los casos incidentes de CP y la carga ancestral se pudo observar una distribución diferencial de acuerdo a los linajes ancestrales, siendo los haplogrupos E y los del macrohaplogrupo F los que presentaron la mayor cantidad de casos de CP respecto a los controles.

De acuerdo al análisis filogenético, los casos portadores de los linajes I y J presentaron un comportamiento relativamente diferencial, centrándose la mayor cantidad de casos en los sub-linajes I2b1 y J2a1h.

Respecto al análisis de escalamiento multidimensional se pudo observar similitudes ancestrales entre los casos y las poblaciones de Medio Oriente, principalmente Líbano, apoyando los hallazgos filogenéticos descritos en el punto anterior.

Nuestro hallazgo más importante sugiere la participación del linaje E con un mayor riesgo a desarrollar CP de tipo poco agresivo y de inicio tardío.

Finalmente, estos resultados son pioneros en revelar el papel potencial de los marcadores del cromosoma Y, asociados a diferentes orígenes ancestrales y su relación con la susceptibilidad al desarrollo del CP en la población mestiza mexicana.

PERSPECTIVAS

Realización de los análisis de alta definición a partir de SNPs, para poder observar la estructura genética dentro de cada linaje ancestral y saber la posible contribución de los sub-linaje en el desarrollo del CP.

Realizar análisis estadísticos para buscar la posible relación entre los diferentes alelos reportados en la población mexicana y su relación con el CP.

Estudiar con detenimiento la asociación entre el macrohaplogrupo F y el desarrollo del CP.

Realizar un estudio piloto en la búsqueda de genes dentro del cromosoma Y que tengan una participación relevante en el desarrollo de la enfermedad

BIBLIOGRAFIA

- Ackerman, C. M., Lowe, L. P., Lee, H., Hayes, M. G., Dyer, A. R., Metzger, B. E., ... Hapo Study Cooperative Research Group, for the H. S. C. R. (2012). Ethnic variation in allele distribution of the androgen receptor (AR) (CAG)_n repeat. *Journal of Andrology*, 33(2), 210–5. <https://doi.org/10.2164/jandrol.111.013391>
- Aimola, P., Carmignani, M., Volpe, A. R., Di Benedetto, A., Claudio, L., Waalkes, M. P., ... Claudio, P. P. (2012). Cadmium induces p53-dependent apoptosis in human prostate epithelial cells. *PloS One*, 7(3), e33647. <https://doi.org/10.1371/journal.pone.0033647>
- Al-Abdin, O. Z., Rabah, D. M., Badr, G., Kotb, A., & Aprikian, A. (2013). Differences in prostate cancer detection between Canadian and Saudi populations. *Brazilian Journal of Medical and Biological Research = Revista Brasileira de Pesquisas Medicas E Biologicas*, 46(6), 539–45. <https://doi.org/10.1590/1414-431X20132757>
- Aranguren-Méndez, J. A., Román-Bravo, R., Isea, W., Villasmil, Y. y J. J. (2004). Los microsatélites (STR's), marcadores moleculares de ADN por excelencia para programas de conservación: una revisión. *Arch Latinoam Prod Anim*, 13(1), 30–42. Retrieved from <http://www.rac.uab.cat/bibliografia/articles/RAC/ALPA1.pdf>
- Aranguren-Méndez, J. A., Román-Bravo, R., Isea, W., Villasmil, Y., & Jordana, J. (2005). Los microsatélites (STR's), marcadores moleculares de ADN por excelencia para programas de conservación: una revisión (Microsatellites (STR's), ADN Molecular Markers for Excellency for conservation programs: A review). Retrieved from <https://tspace.library.utoronto.ca/handle/1807/7085>
- Asimit, J. L., Hatzikotoulas, K., McCarthy, M., Morris, A. P., & Zeggini, E. (2016). Trans-ethnic study design approaches for fine-mapping. *European Journal of Human Genetics*, 24(9), 1330–1336. <https://doi.org/10.1038/ejhg.2016.1>
- Athey, T. W. (2005). Haplogroup Prediction from Y-STR Values Using an Allele-Frequency Approach. *Journal of Genetic Genealogy*, 2(2), 34–39.
- Athey, T. W. (2006). Haplogroup Prediction from Y-STR Values Using a Bayesian-

- Allele- Frequency Approach. *Journal of Genetic Genealogy*, 2(2), 34–39.
- Attia, J., Ioannidis, J. P. A., Thakkinstian, A., McEvoy, M., Scott, R. J., Minelli, C., ... Guyatt, G. (2009). How to use an article about genetic association: A: Background concepts. *Jama*, 301(1), 74–81. <https://doi.org/10.1001/jama.2008.901>
- Bangsi, D., Zhou, J. Y., Sun, Y., Patel, N. P., Darga, L. L., Heilbrun, L. K., ... Everson, R. B. (2006). Impact of a genetic variant in CYP3A4 on risk and clinical presentation of prostate cancer among white and African-American men. *Urologic Oncology: Seminars and Original Investigations*, 24(1), 21–27. <https://doi.org/10.1016/j.urolonc.2005.09.005>
- Barrot, C., Simili, C., Sánchez, C., Brandt-Casadevall, C., González-Martín, A., Xifró, A., ... Gené, M. (2007). Haplotype Frequencies of Eight Y-Chromosome Short Tandem Repeats Loci in Four Amerindian Populations (State of Hidalgo, Mexico). *Journal of Forensic Sciences*, 52(2), 504–506. <https://doi.org/10.1111/j.1556-4029.2007.00398.x>
- Beuten, J., Gelfond, J. A. L., Franke, J. L., Shook, S., Johnson-Pais, T. L., Thompson, I. M., & Leach, R. J. (2010). Single and multivariate associations of MSR1, ELAC2, and RNASEL with prostate cancer in an ethnic diverse cohort of men. *Cancer Epidemiology, Biomarkers & Prevention: A Publication of the American Association for Cancer Research, Cosponsored by the American Society of Preventive Oncology*, 19(2), 588–99. <https://doi.org/10.1158/1055-9965.EPI-09-0864>
- Bolaños, P., & Chacon, C. (2017). Escala patológica de Gleason para el cáncer de próstata y sus modificaciones. *Medicina Legal de Costa Rica*, 34(1), 237–243. Retrieved from http://www.scielo.sa.cr/scielo.php?pid=S1409-00152017000100237&script=sci_arttext
- Bonilla, C., Lewis, S. J., Martin, R. M., Donovan, J. L., Hamdy, F. C., Neal, D. E., ... consortium, P. (2016). Pubertal development and prostate cancer risk: Mendelian randomization study in a population-based cohort. *BMC Med*, 14(1), 66. <https://doi.org/10.1186/s12916-016-0602-x>
- Bostwick, D. G., Burke, H. B., Djakiew, D., Euling, S., Ho, S., Landolph, J., ... Timms, B. (2004). Human prostate cancer risk factors. *Cancer*, 101(S10), 2371–2490. <https://doi.org/10.1002/CNCR.20408>

- Boyle, P., & Severi, G. (1999). Epidemiology of prostate cancer chemoprevention. *European Urology*, *35*(5–6), 370–376. <https://doi.org/10.1159/000019911>
- Brandt, A., Sundquist, J., & Hemminki, K. (2012). Risk for incident and fatal prostate cancer in men with a family history of any incident and fatal cancer. *Annals of Oncology*, *23*(1), 251–256. <https://doi.org/10.1093/annonc/mdr056>
- Cannon-Albright, L. A., Farnham, J. M., Bailey, M., Albright, F. S., Teerlink, C. C., Agarwal, N., ... Thomas, A. (2014a). Identification of specific Y chromosomes associated with increased prostate cancer risk. *The Prostate*, *74*(9), 991–8. <https://doi.org/10.1002/pros.22821>
- Cannon-Albright, L. A., Farnham, J. M., Bailey, M., Albright, F. S., Teerlink, C. C., Agarwal, N., ... Thomas, A. (2014b). Identification of specific Y chromosomes associated with increased prostate cancer risk. *The Prostate*, *74*(9), 991–8. <https://doi.org/10.1002/pros.22821>
- Carson, C., & Rittmaster, R. (2003). The role of dihydrotestosterone in benign prostatic hyperplasia. *Urology*, *61*(4), 2–7. [https://doi.org/10.1016/S0090-4295\(03\)00045-1](https://doi.org/10.1016/S0090-4295(03)00045-1)
- Carvalho, R., Pinheiro, M. F., & Medeiros, R. (2010). Localization of Candidate Genes in a Region of High Frequency of Microvariant Alleles for Prostate Cancer Susceptibility: The Chromosome Region *Yp11.2* Genetic Variation. *DNA and Cell Biology*, *29*(1), 3–7. <https://doi.org/10.1089/dna.2009.0905>
- CDC. (2014). CDC - Tasas de cáncer de próstata por raza y grupo étnico. Retrieved March 5, 2018, from <https://www.cdc.gov/spanish/cancer/prostate/statistics/race.htm>
- Chamberlain, N. L., Driver, E. D., & Miesfeld, R. L. (1994). The length and location of CAG trinucleotide repeats in the androgen receptor N-terminal domain affect transactivation function. *Nucleic Acids Research*, *22*(15), 3181–3186. <https://doi.org/10.1093/nar/22.15.3181>
- Chang, E. T., & Delzell, E. (2016). Systematic review and meta-analysis of glyphosate exposure and risk of lymphohematopoietic cancers. *Journal of Environmental Science and Health, Part B*, *51*(6), 402–428. <https://doi.org/10.1080/03601234.2016.1142748>
- Chen, Y.-C., Page, J. H., Chen, R., & Giovannucci, E. (2008). Family history of prostate and breast cancer and the risk of prostate cancer in the PSA era. *The*

- Prostate*, 68(14), 1582–91. <https://doi.org/10.1002/pros.20825>
- Cheong, A., Zhang, X., Cheung, Y.-Y., Tang, W.-Y., Chen, J., Ye, S.-H., ... Ho, S.-M. (2016). DNA methylome changes by estradiol benzoate and bisphenol A links early-life environmental exposures to prostate cancer risk. *Epigenetics*, 11(9), 674–689. <https://doi.org/10.1080/15592294.2016.1208891>
- Colaco, S., & Modi, D. (2018). Genetics of the human Y chromosome and its association with male infertility. *Reproductive Biology and Endocrinology: RB&E*, 16(1), 14. <https://doi.org/10.1186/s12958-018-0330-5>
- Crawford, D. (2015). Epidemiology of prostate cancer. *Elsevier*, 16(13), 5137–5141. <https://doi.org/10.7314/APJCP.2015.16.13.5137>
- da Silva, H. B., Amaral, E. P., Nolasco, E. L., de Victo, N. C., Atique, R., Jank, C. C., ... Correa, R. G. (2013). Dissecting Major Signaling Pathways throughout the Development of Prostate Cancer. *Prostate Cancer*, 2013, 920612. <https://doi.org/10.1155/2013/920612>
- Darwish-Yassine, M., & Wing, ; Diane. (2005). *D. CANCER EPIDEMIOLOGY IN ARAB AMERICANS AND ARABS OUTSIDE THE MIDDLE EAST SELECTED LITERATURE REVIEW*. Retrieved from <https://www.ethndis.org/priorsuparchives/ethn-15-01s-0005.pdf>
- Dasari, V. K., Goharderakhshan, R. Z., Perinchery, G., Li, L. C., Tanaka, Y., Alonzo, J., & Dahiya, R. (2001). Expression analysis of Y chromosome genes in human prostate cancer. *The Journal of Urology*, 165(4), 1335–41. Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/11257712>
- Diaz, L. F. (2010). Analisis de 17 loci de str de cromosoma Y en las poblaciones de Bogota y Santander con fines genético, poblacionales y forenses, 1–158. <https://doi.org/10.1017/CBO9781107415324.004>
- Ding, D., Xu, L., Menon, M., Reddy, G. P. V., & Barrack, E. R. (2004). Effect of a Short CAG (Glutamine) Repeat on Human Androgen Receptor Function. *Prostate*, 58(1), 23–32. <https://doi.org/10.1002/pros.10316>
- Eeles, R., Goh, C., Castro, E., Bancroft, E., Guy, M., Olama, A. A. Al, ... Kote-Jarai, Z. (2013). The genetic epidemiology of prostate cancer and its clinical implications. *Nature Reviews Urology*, 11(1), 18–31. <https://doi.org/10.1038/nrurol.2013.266>
- Elhaik, E., Tatarinova, T., Chebotarev, D., Piras, I. S., Maria Calò, C., De Montis,

- A., ... Genographic Consortium, T. G. (2014). Geographic population structure analysis of worldwide human populations infers their biogeographical origins. *Nature Communications*, 5, 3513. <https://doi.org/10.1038/ncomms4513>
- Ewis, A. A., Lee, J., Naroda, T., Sano, T., Kagawa, S., Iwamoto, T., ... Nakahori, Y. (2006). Prostate cancer incidence varies among males from different Y-chromosome lineages. *Prostate Cancer and Prostatic Diseases*, 9(3), 303–309. <https://doi.org/10.1038/sj.pcan.4500876>
- Ferlay, J., Soerjomataram, I., Dikshit, R., Eser, S., Mathers, C., Rebelo, M., ... Bray, F. (2015). Cancer incidence and mortality worldwide: Sources, methods and major patterns in GLOBOCAN 2012. *International Journal of Cancer*, 136(5), E359–E386. <https://doi.org/10.1002/ijc.29210>
- Forsberg, L. A. (2017). Loss of chromosome Y (LOY) in blood cells is associated with increased risk for disease and mortality in aging men. *Human Genetics*, 136(5), 657–663. <https://doi.org/10.1007/s00439-017-1799-2>
- Goldstein, D. B., & Schlötterer, C. (1999). *Microsatellites: evolution and applications*. Oxford University Press.
- Gómez, R., Torres-Sánchez, L., Camacho-Mejorado, R., Burguete-García, A. I., Vázquez-Salas, R. A., Martínez-Nava, G. A., ... Noris, G. (2016). Androgen receptor CAG polymorphism and sporadic and early-onset prostate cancer among Mexican men. *Journal of Human Genetics*, 61(9), 781–786. <https://doi.org/10.1038/jhg.2016.49>
- González-Sobrino, B. Z., Pintado-Cortina, A. P., Sebastián-Medina, L., Morales-Mandujano, F., Contreras, A. V., Aguilar, Y. E., ... Medrano-González, L. (2016). Genetic Diversity and Differentiation in Urban and Indigenous Populations of Mexico: Patterns of Mitochondrial DNA and Y-Chromosome Lineages. *Biodemography and Social Biology*, 62(1), 53–72. <https://doi.org/10.1080/19485565.2015.1117938>
- González, E. (2003). Microsatélites: sus aplicaciones en la conservación de la biodiversidad. *Revista Científicas Del CSIC*, 59, 377–388. <https://doi.org/10.3989/graellsia.2003.v59.i2-3.253>
- Grönberg, H. (2003). Prostate cancer epidemiology. *The Lancet*, 361(9360), 859–864. [https://doi.org/10.1016/S0140-6736\(03\)12713-4](https://doi.org/10.1016/S0140-6736(03)12713-4)

- Hanson, E. K., Ballantyne, J., Chakraborty, R., Allen, M., Parsons, T. J., Brown, L. G., & Al., E. (2006). Comprehensive annotated STR physical map of the human Y chromosome: Forensic implicatioHanson, Erin K et al. "Comprehensive Annotated STR Physical Map of the Human Y Chromosome: Forensic Implications." *Legal medicine (Tokyo, Japan)* 8.2 (2006): 110–20. Web. *Legal Medicine (Tokyo, Japan)*, 8(2), 110–20. <https://doi.org/10.1016/j.legalmed.2005.10.001>
- Harada, N., Mitani, T., Higashimura, Y., Yamaji, R., Okamoto, K., Nakano, Y., & Inui, H. (2010). Involvement of three glutamine tracts in human androgen receptor transactivation. *Journal of Steroid Biochemistry and Molecular Biology*, 118(1–2), 77–84. <https://doi.org/10.1016/j.jsbmb.2009.10.003>
- Hasanain, K. S., Noran, J. I., Imad, H. H., & Muhanned, A. K. (2015). DYS459, DYS391, DYS388 and DYS19 genetic loci have high allelic frequency in patients with prostate cancer. *International Journal of Genetics and Molecular Biology*, 7(6), 47–58. <https://doi.org/10.5897/IJGMB2015.0110>
- Hermabessière, J., & Taillandier, J. (2002). Fisiología de la próstata. *EMC - Urología*, 34(1), 1–6. [https://doi.org/10.1016/S1761-3310\(02\)72340-2](https://doi.org/10.1016/S1761-3310(02)72340-2)
- Humphrey, P. A. (2004). Gleason grading and prognostic factors in carcinoma of the prostate. *Modern Pathology*, 17(3), 292–306. <https://doi.org/10.1038/modpathol.3800054>
- Jordan, J. J., Hanlon, A. L., Al-Saleem, T. I., Greenberg, R. E., & Tricoli, J. V. (2001). Loss of the short arm of the Y chromosome in human prostate carcinoma. *Cancer Genetics and Cytogenetics*, 124(2), 122–126. [https://doi.org/10.1016/S0165-4608\(00\)00340-X](https://doi.org/10.1016/S0165-4608(00)00340-X)
- Kayser, M. (2017). Forensic use of Y-chromosome DNA: a general overview. *Human Genetics*, 136(5), 621–635. <https://doi.org/10.1007/s00439-017-1776-9>
- Kido, T., & Lau, Y.-F. C. (2015). Roles of the Y chromosome genes in human cancers. *Asian Journal of Andrology*, 17(3), 373–80. <https://doi.org/10.4103/1008-682X.150842>
- Lee, J. S., Cheong, H. S., Kim, L. H., Kim, J. O., Seo, D. W., Kim, Y. H., ... Shin, H. D. (2013). Screening of Genetic Polymorphisms of CYP3A4 and CYP3A5 Genes. *The Korean Journal of Physiology & Pharmacology : Official Journal*

- of the Korean Physiological Society and the Korean Society of Pharmacology, 17(6), 479–84. <https://doi.org/10.4196/kjpp.2013.17.6.479>
- Lindström, S., Adami, H. O., Adolfsson, J., & Wiklund, F. (2008). Y chromosome haplotypes and prostate cancer in Sweden. *Clinical Cancer Research*, 14(20), 6712–6716. <https://doi.org/10.1158/1078-0432.CCR-08-0658>
- Litman, H. J., Bhasin, S., Link, C. L., Araujo, A. B., & McKinlay, J. B. (2006). Serum Androgen Levels in Black, Hispanic, and White Men. *The Journal of Clinical Endocrinology & Metabolism*, 91(11), 4326–4334. <https://doi.org/10.1210/jc.2006-0037>
- Lopez, D. S., Peskoe, S. B., Joshu, C. E., Dobs, A., Feinleib, M., Kanarek, N., ... Platz, E. A. (2013). Racial/ethnic differences in serum sex steroid hormone concentrations in US adolescent males. *Cancer Causes & Control: CCC*, 24(4), 817–26. <https://doi.org/10.1007/s10552-013-0154-8>
- López, J. I., & Angulo, J. (2000). Desarrollo embriológico y postnatal de la próstata. In *Fundamentos de Urología*. Madrid. Retrieved from https://www.researchgate.net/profile/Jose_I_Lopez/publication/280546173_Desarrollo_embriologico_y_postnatal_de_la_prostata/links/55b8853008ae9289a08d60d1/Desarrollo-embriologico-y-postnatal-de-la-prostata.pdf
- Lorente, D., Mateo, J., Zafeiriou, Z., Smith, A. D., Sandhu, S., Ferraldeschi, R., & de Bono, J. S. (2015). Switching and withdrawing hormonal agents for castration-resistant prostate cancer. *Nature Reviews. Urology*, 12(1), 37–47. <https://doi.org/10.1038/nruro.2014.345>
- Lorenz, T. C. (2012). Polymerase chain reaction: basic protocol plus troubleshooting and optimization strategies. *Journal of Visualized Experiments: JoVE*, (63), e3998. <https://doi.org/10.3791/3998>
- Magaña, J., Arenas-Sordo, M. de la L., & Gómez, R. (2009). La electroforesis capilar como una nueva estrategia en la medicina y el diagnóstico clínico. *Revista Médica de Chile*, 137(7), 946–956. <https://doi.org/10.4067/S0034-98872009000700014>
- Makaya, T. (n.d.). *Embryology Prostate Development*. Retrieved from https://embryology.med.unsw.edu.au/embryology/index.php/Prostate_Development
- Moussawi, A. H., Yassine, M., Dey, S., & Soliman, A. S. (2013). Clinical profile,

- quality of care, and recurrence in Arab-American and Caucasians prostate cancer patients in Michigan. *Journal of Immigrant and Minority Health*, 15(4), 803–9. <https://doi.org/10.1007/s10903-012-9662-y>
- Mustafa, M., Salih, A., Illzam, E., Sharifa, A., Suleiman, M., & Hussain, S. (2016). Prostate Cancer: Pathophysiology, Diagnosis, and Prognosis. *IOSR Journal of Dental and Medical Sciences Ver. II*, 15, 2279–861. <https://doi.org/10.9790/0853-1506020411>
- Nargesi, M. M., Ismail, P., Razack, A. H. A., Pasalar, P., Nazemi, A., Oshkoor, S. A., & Amini, P. (2011). Linkage between prostate cancer occurrence and Y-chromosomal DYS loci in Malaysian subjects. *Asian Pacific Journal of Cancer Prevention: APJCP*, 12(5), 1265–8. Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/21875279>
- Nasseri, K., Mills, P. K., & Allan, M. (2007). Cancer incidence in the Middle Eastern population of California, 1988-2004. *Asian Pacific Journal of Cancer Prevention: APJCP*, 8(3), 405–11. Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/18159978>
- Nelles, J. L., Hu, W.-Y., & Prins, G. S. (2011). Estrogen action and prostate cancer. *Expert Review of Endocrinology & Metabolism*, 6(3), 437–451. <https://doi.org/10.1586/eem.11.20>
- Packer, J. R., & Maitland, N. J. (2016). The molecular and cellular origin of human prostate cancer. *Biochimica et Biophysica Acta (BBA) - Molecular Cell Research*, 1863(6), 1238–1260. <https://doi.org/10.1016/j.bbamcr.2016.02.016>
- Parent, M.-E., Désy, M., & Siemiatycki, J. (2009). Does exposure to agricultural chemicals increase the risk of prostate cancer among farmers? *McGill Journal of Medicine: MJM: An International Forum for the Advancement of Medical Sciences by Students*, 12(1), 70–77.
- Paris, P. L., Kupelian, P. A., Hall, J. M., Williams, T. L., Levin, H., Klein, E. A., ... Witte, J. S. (1999). Association between a CYP3A4 Genetic Variant and Clinical Presentation in African-American Prostate Cancer Patients Association between a CYP3A4 Genetic Variant and Clinical Presentation in African-American Prostate Cancer Patients 1. *Cancer Epidemiology, Biomarkers and Prevention*, 8(October), 901–905.
- Park, J. W., Lee, J. K., Phillips, J. W., Huang, P., Cheng, D., Huang, J., & Witte,

- O. N. (2016). Prostate epithelial cell of origin determines cancer differentiation state in an organoid transformation assay. *Proceedings of the National Academy of Sciences*, *113*(16), 4482–4487. <https://doi.org/10.1073/pnas.1603645113>
- Pierorazio, P. M., Walsh, P. C., Partin, A. W., & Epstein, J. I. (2013). Prognostic Gleason grade grouping: Data based on the modified Gleason scoring system. *BJU International*, *111*(5), 753–760. <https://doi.org/10.1111/j.1464-410X.2012.11611.x>
- Ragavan, N., Hewitt, R., Cooper, L. J., Ashton, K. M., Hindley, A. C., Nicholson, C. M., ... Martin, F. L. (2004). CYP1B1 expression in prostate is higher in the peripheral than in the transition zone. *Cancer Letters*, *215*(1), 69–78. <https://doi.org/10.1016/j.canlet.2004.06.051>
- Rey, R., & Copelli, S. (2010). DIFERENCIACIÓN SEXUAL EMBRIOFETAL | Tratado de endocrinología pediátrica, 4e | AccessMedicina | McGraw-Hill Medical. Retrieved February 20, 2018, from <http://accessmedicina.mhmedical.com/content.aspx?bookid=1508§ionid=102965192>
- Rey, R., Josso, N., & Racine, C. (2000). *Sexual Differentiation*. Endotext. MDText.com, Inc. Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/25905232>
- Ross, M. T., Bentley, D. R., & Tyler-Smith, C. (2006). The sequences of the human sex chromosomes. *Current Opinion in Genetics & Development*, *16*(3), 213–218. <https://doi.org/10.1016/J.GDE.2006.04.006>
- Rybak, A. P., Bristow, R. G., & Kapoor, A. (2015). Prostate cancer stem cells: deciphering the origins and pathways involved in prostate tumorigenesis and aggression. *Oncotarget*, *6*(4), 1900–19. <https://doi.org/10.18632/oncotarget.2953>
- Salazar-Flores, J., Dondiego-Aldape, R., Rubi-Castellanos, R., Anaya-Palafox, M., Nuño-Arana, I., Canseco-Ávila, L. M., ... Rangel-Villalobos, H. (2010). Population structure and paternal admixture landscape on present-day Mexican-mestizos revealed by Y-STR haplotypes. *American Journal of Human Biology*, *22*(3), 401–409. <https://doi.org/10.1002/ajhb.21013>
- Santana, C., Noris, G., Meraz-Ríos, M. A., Magaña, J. J., Calderon-Aranda, E. S.,

- De Lourdes Muñoz, M., & Gómez, R. (2015). Genetic analysis of 17 y-STRs in a mestizo population from the central valley of Mexico. *Human Biology*, 86(4), 289–312.
- Santana, C., Noris, G., Meraz-Ríos, M. A., Magaña, J. J., Calderon-Aranda, E. S., Muñoz, M. de L., & Gómez, R. (2014). Genetic Analysis of 17 Y-STRs in a Mestizo Population from the Central Valley of Mexico. *Human Biology*, 86(4), 289–312. <https://doi.org/10.13110/humanbiology.86.4.0289>
- Severi, G., Morris, H. A., Macinnis, R. J., English, D. R., Tilley, W., Hopper, J. L., ... Giles, G. G. (2006). Circulating Steroid Hormones and the Risk of Prostate Cancer. *Cancer Epidemiol Biomarkers Prev*, 15(January), 86–91. <https://doi.org/10.1158/1055-9965.EPI-05-0633>
- Skaletsky, H., Kuroda-Kawaguchi, T., Minx, P. J., Cordum, H. S., Hillier, L., Brown, L. G., ... Page, D. C. (2003). The male-specific region of the human Y chromosome is a mosaic of discrete sequence classes. *Nature*, 423(6942), 825–837. <https://doi.org/10.1038/nature01722>
- Sucheston, L. E., Bensen, J. T., Xu, Z., Singh, P. K., Preus, L., Mohler, J. L., ... Taylor, J. A. (2012). Genetic ancestry, self-reported race and ethnicity in African Americans and European Americans in the PCaP cohort. *PloS One*, 7(3), e30950. <https://doi.org/10.1371/journal.pone.0030950>
- Tamay de Dios, L., Ibarra, C., & Velasquillo, C. (2013). Fundamentos de la reacción en cadena de la polimerasa (PCR) y de la PCR en tiempo real. *Enfermedades Infecciosas Y Microbiología Clínica*, 2(5), 70–78. <https://doi.org/10.1157/13059826>
- Taylor, T. (2017). Prostate - Anatomy of the prostate. Retrieved August 2, 2017, from http://www.innerbody.com/image_repmov/repo09-new4.html#full-description
- Tiirikka, T., & Moilanen, J. S. (2015). Human Chromosome Y and Haplogroups; introducing YDHS Database. *Clinical and Translational Medicine*, 4(1), 20. <https://doi.org/10.1186/s40169-015-0060-7>
- Timms, B. G. (2008). Prostate development: A historical perspective. *Differentiation*, 76(6), 565–577. <https://doi.org/10.1111/j.1432-0436.2008.00278.x>

- Tóth, G., Gáspári, Z., & Jurka, J. (2000). Microsatellites in different eukaryotic genomes: survey and analysis. *Genome Research*, 10(7), 967–81. <https://doi.org/10.1101/GR.10.7.967>
- Velcheti, V., Karnik, S., Bardot, S. F., & Prakash, O. (2008). Pathogenesis of prostate cancer: lessons from basic research. *The Ochsner Journal*, 8(4), 213–8. <https://doi.org/10.1038/sj.bjc.6690407>
- Via, M., Ziv, E., & Burchard, E. (2009). Recent advances of genetic ancestry testing in biomedical research and direct to consumer testing. *Clinical Genetics*, 76(3), 225–235. <https://doi.org/10.1111/j.1399-0004.2009.01263.x>
- Wigle, D. T., Turner, M. C., Gomes, J., & Parent, M.-É. (2008). Role of Hormonal and Other Factors in Human Prostate Cancer. *Journal of Toxicology and Environmental Health, Part B*, 11(3–4), 242–259. <https://doi.org/10.1080/10937400701873548>
- Wu, Y., Godoy, A., Azzouni, F., Wilton, J. H., Ip, C., & Mohler, J. L. (2013). Prostate cancer cells differ in testosterone accumulation, dihydrotestosterone conversion, and androgen receptor signaling response to steroid 5 α -reductase inhibitors. *Prostate*, 73(13), 1470–1482. <https://doi.org/10.1002/pros.22694>

INFORMACIÓN SUPLEMENTARIA

Tabla Suplementaria S1. Frecuencias alélicas globales presentes en las 533 muestras analizadas (150 casos y 383 controles) de la población mexicana, haciendo uso de los 17 Y-STRs.

| Alelo | DYS19 | DYS385-a | DYS385-b | DYS389-I | DYS389-II | DYS390 | DYS391 | DYS392 | DYS393 | DYS437 | DYS438 | DYS439 | DYS448 | DYS458 | DYS456 | DYS635 | Y GATA H4 |
|-------|--------------|--------------|--------------|--------------|--------------|--------------|--------------|--------------|--------------|--------------|--------------|--------------|--------------|--------------|--------------|--------------|--------------|
| 8 | | | | | | | | | | | 0.006 | | | | | | |
| 9 | | | | | | | 0.084 | | | | 0.073 | | | | | | 0.002 |
| 10 | | 0.009 | | | | | 0.640 | 0.011 | 0.004 | | 0.188 | 0.058 | | | | | 0.021 |
| 11 | | 0.263 | 0.009 | 0.004 | | | 0.263 | 0.211 | 0.009 | | 0.400 | 0.302 | | | | | 0.366 |
| 12 | 0.008 | 0.073 | 0.023 | 0.199 | | | 0.013 | 0.032 | 0.141 | | 0.313 | 0.454 | | | | | 0.544 |
| 13 | 0.441 | 0.158 | 0.054 | 0.589 | | | | 0.353 | 0.722 | 0.015 | 0.021 | 0.169 | | | 0.011 | | 0.068 |
| 14 | 0.373 | 0.235 | 0.257 | 0.205 | | | | 0.226 | 0.113 | 0.651 | | 0.013 | | 0.043 | 0.054 | | |
| 15 | 0.124 | 0.150 | 0.122 | 0.004 | | | | 0.068 | 0.011 | 0.293 | | 0.004 | | 0.177 | 0.542 | | |
| 16 | 0.039 | 0.084 | 0.131 | | | | | 0.091 | | 0.041 | | | | 0.271 | 0.306 | | |
| 17 | 0.015 | 0.026 | 0.218 | | | | | 0.002 | | | | | 0.008 | 0.278 | 0.073 | | |
| 18 | | 0.002 | 0.129 | | | | | 0.006 | | | | | 0.089 | 0.169 | 0.013 | | |
| 19 | | | 0.038 | | | | | | | | | | 0.388 | 0.055 | | | |
| 20 | | | 0.013 | | | | | | | | | | 0.386 | 0.008 | | 0.030 | |
| 21 | | | 0.006 | | | 0.019 | | | | | | | 0.119 | | | 0.150 | |
| 22 | | | | | | 0.060 | | | | | | | 0.009 | | | 0.441 | |
| 23 | | | | | | 0.248 | | | | | | | | | | 0.300 | |
| 24 | | | | | | 0.548 | | | | | | | 0.002 | | | 0.071 | |
| 25 | | | | | | 0.101 | | | | | | | | | | 0.008 | |
| 26 | | | | | 0.009 | 0.017 | | | | | | | | | | | |
| 27 | | | | | 0.017 | 0.006 | | | | | | | | | | | |
| 28 | | | | | 0.144 | 0.002 | | | | | | | | | | | |
| 29 | | | | | 0.321 | | | | | | | | | | | | |
| 30 | | | | | 0.332 | | | | | | | | | | | | |
| 31 | | | | | 0.126 | | | | | | | | | | | | |
| 32 | | | | | 0.049 | | | | | | | | | | | | |
| 34 | | | | | 0.002 | | | | | | | | | | | | |
| k= | 6 | 9 | 11 | 5 | 8 | 8 | 4 | 9 | 6 | 4 | 6 | 6 | 7 | 7 | 6 | 6 | 5 |
| D= | 0.649 | 0.815 | 0.833 | 0.572 | 0.747 | 0.624 | 0.514 | 0.766 | 0.446 | 0.491 | 0.701 | 0.671 | 0.680 | 0.785 | 0.604 | 0.688 | 0.568 |

Los números en negritas representan los alelos más frecuentes de cada uno de los 17 marcadores Y-STR; k= Número de alelos por loci; D= Diversidad alélica.

Tabla Suplementaria 2. Frecuencia de los linajes encontrados en la población mexicana.

| Haplogrupo | Individuos | Frecuencia |
|----------------|------------|------------|
| E1b1a | 8 | 1.50 |
| E1b1b | 41 | 7.69 |
| G2a | 11 | 2.06 |
| H | 3 | 0.56 |
| I1 | 6 | 1.13 |
| I2a1 | 9 | 1.69 |
| I2b (xI2b1) | 1 | 0.19 |
| I2b1 | 7 | 1.31 |
| J1 | 14 | 2.63 |
| J2a1 x J2a1-bh | 6 | 1.13 |
| J2a1b | 6 | 1.13 |
| J2a1h | 9 | 1.69 |
| J2b | 3 | 0.56 |
| L | 7 | 1.31 |
| Q | 225 | 42.21 |
| R1a | 6 | 1.13 |
| R1b | 160 | 30.02 |
| T | 11 | 2.06 |
| Total | 533 | 100 |

Figura Suplementaria S1. Red haplotípica del linaje Q utilizando 15 Y-STRs. Cada haplotipo está representado por un círculo, cuyo tamaño corresponde a la frecuencia de ese haplotipo en la población.

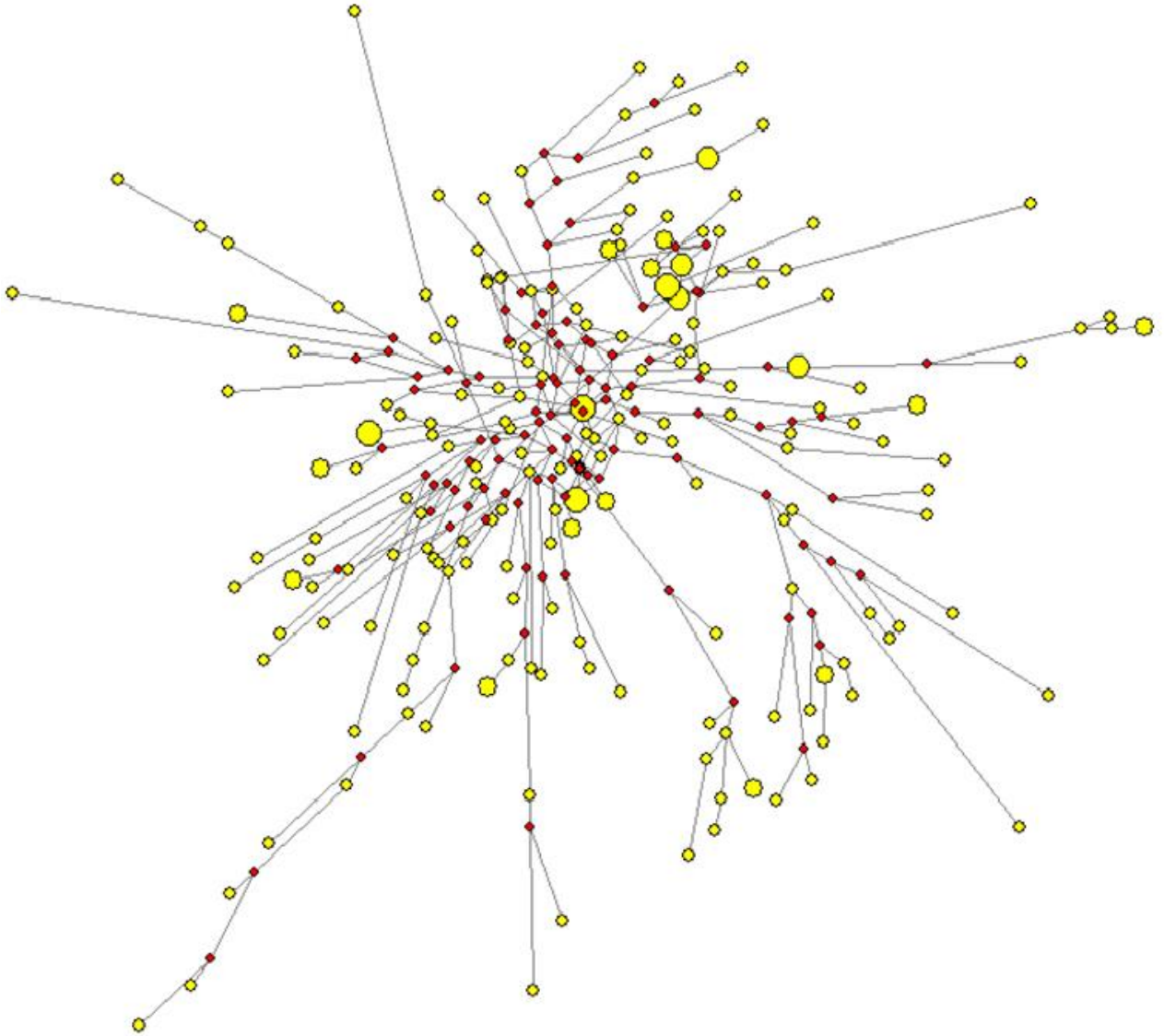


Figura Suplementaria S2. Red haplotípica del linaje R (R1a y R1b) utilizando 15 Y-STRs. Los haplotipos pertenecientes al linaje R1a se presentan enmarcados en el recuadro rojo. Cada haplotipo está representado por un círculo, cuyo tamaño corresponde a la frecuencia de ese haplotipo en la población.

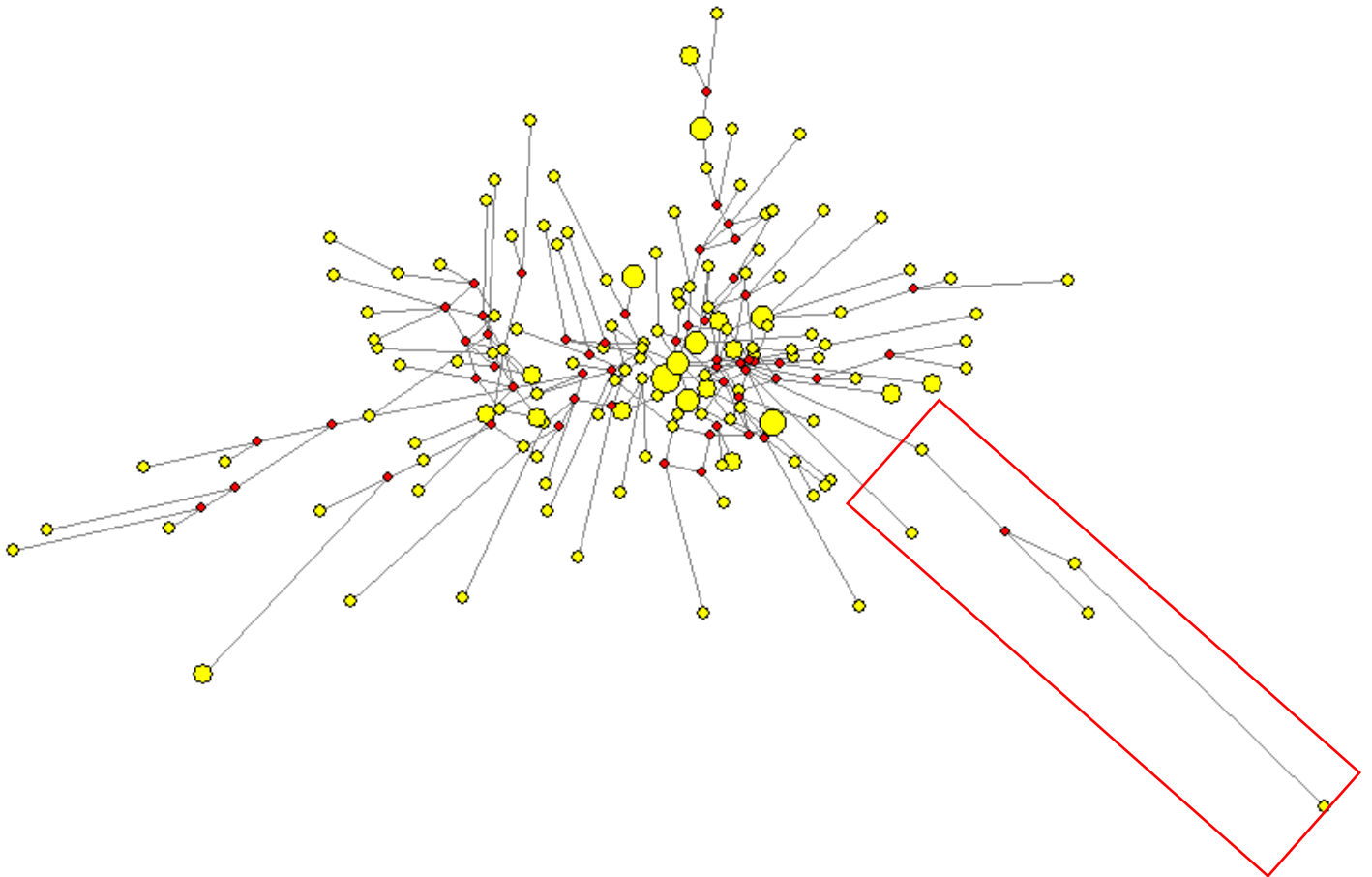
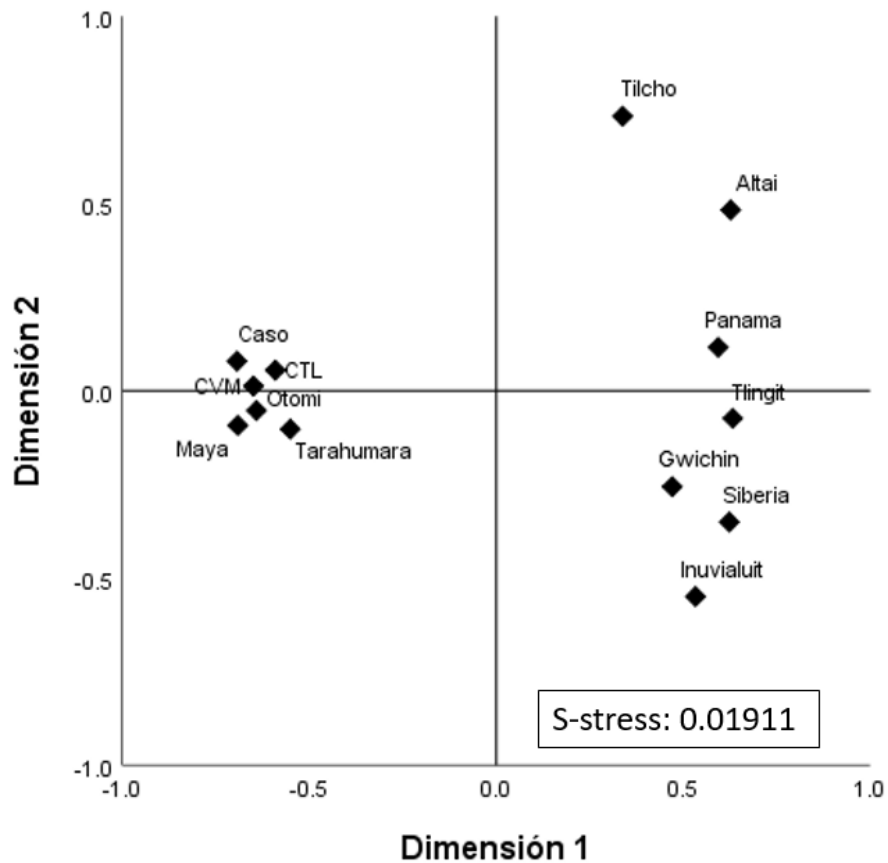
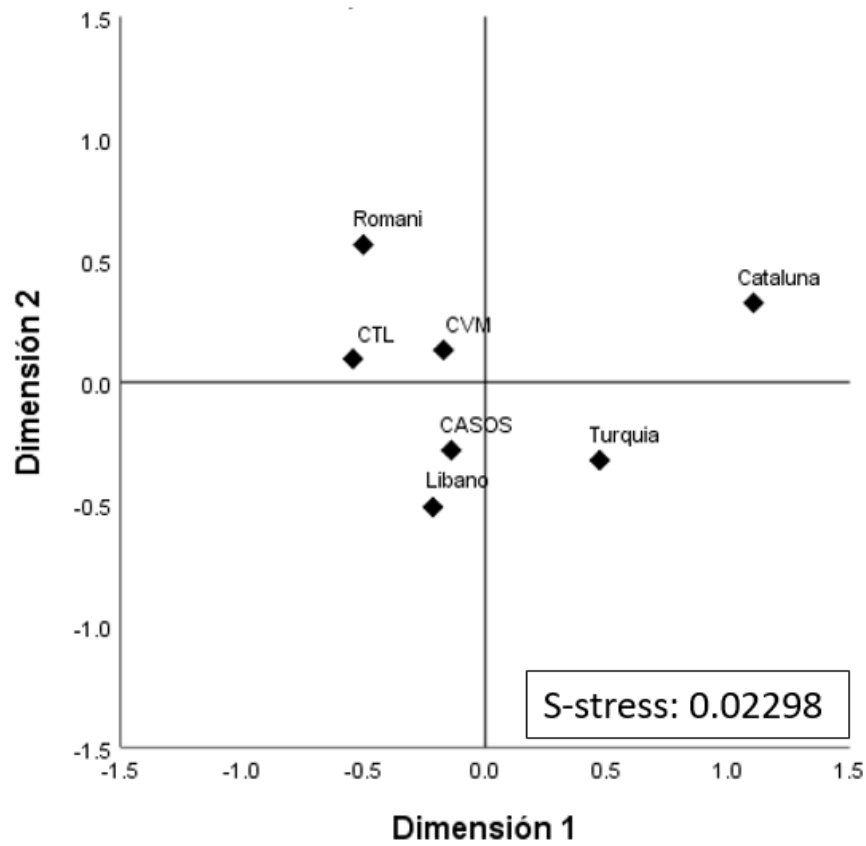


Figura Suplementaria S3. Análisis de escalamiento multidimensional empleando portadores del haplogrupo Q empleando 15 Y-STRs. La gráfica muestra el análisis empleando diferentes poblaciones de comparación.



CTL: Control; CVM: Valle Central de México (Santana et al., 2012); Otomies, Maya, Tarahumara (Regueiro et al., 2013); Tilcho, Gwichin, Inuvialuit (Dulik et al., 2012a); Altai (Dulik et al., 2012b); Panamá (Purps et al., 2014); Tlingit (Schurr et al., 2012); Siberia (Regueiro et al., 2013).

Figura Suplementaria S4. Análisis de escalamiento multidimensional empleando portadores del haplogrupo G empleando 15 Y-STRs. La gráfica muestra el análisis empleando diferentes poblaciones de comparación.



CTL: Control; CVM: Valle Central de México (Santana et al., 2014); Turquía (Dönbak et al., 2008); Libano (Pierre et al., 2008); Romaní (Bogoña- Cruz et al., 2016); Cataluña (Martínez et al., 2012).